

Chapter 2

Theoretical Background

2.1 Basics of Crystallography

2.1.1 Introduction to the general features of crystals

This and the next sections aim at equipping the reader with the essential knowledge of crystallography and X-ray diffraction deemed necessary to gain a better understanding and appreciation of the various crystal structures considered in the context of the research presented in this thesis. For a full introduction to the theory of crystals, the reader is referred to one of the many textbooks on Condensed Matter physics [1] and X-ray diffraction [2].

Materials in the crystalline state are widespread and play a role of paramount importance in many everyday applications. A solid is considered to be crystalline if at the microscopic level its atoms are arranged in a three-dimensional periodic array. This was directly confirmed through the experimental work of W. and L. Bragg on X-ray diffraction by crystals. Crystals are generally-speaking very

regular in shape and exhibit a significant amount of symmetry. This is apparent from both the point of view of the external morphology of crystals, which is found to be bounded by planar surfaces, and from that of the interior structure, which can be understood from the preferred directions along which crystals cleave when fractured. This suggests that the macroscopic form of a crystal depends on the underlying structure at the atomic or molecular level, and that the fundamental factor which controls crystal formation is the way in which atoms and molecules pack together. Examples of this regular shape are the well-formed crystals of alum, which have the shape of a perfect octahedron, and quartz crystals, which have cross-sections that are regular hexagons.

2.1.2 Bravais Lattices

The most fundamental concept in the description of crystalline solids is the *Bravais Lattice*, which explains the geometry of the underlying periodic structure of atoms in a crystal, regardless of what the actual chemical composition of the crystal may be. There are two equivalent definitions of a Bravais Lattice:

1. A Bravais lattice is an infinite periodic array of abstract discrete lattice points with an arrangement and orientation that appears exactly the same from whichever of the points the array is viewed. This is to say that all lattice points are equivalent, and the array is therefore periodic.
2. A three-dimensional Bravais lattice consists of the set of all lattice points with position vectors \vec{R} relative to one of the points of the form:

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 \quad (2.1)$$

which are known as Bravais Lattice Vectors. The three \vec{a} vectors must not lie in the same plane, and the n coefficients are integers. The \vec{a} vectors are known as primitive translation vectors and are said to generate or span the Bravais Lattice. \vec{R} corresponds to the displacement vectors between any two Bravais lattice points, that is the sum or difference of any two Bravais lattice vectors, including the Bravais lattice vectors themselves.

An important property of all Bravais Lattices is that they have translational symmetry, and therefore are invariant under a translation by a displacement vector \vec{R} between any two Bravais Lattice points. \vec{R} therefore represents the periodicity of the Bravais lattice. As a result, any local physical property of the crystal is invariant (periodic) under \vec{R} , such as the charge concentration, electron number density, or magnetic moment density. For any given Bravais lattice the set of primitive vectors is not unique, and there are infinitely many nonequivalent choices, just like basis sets in linear algebra for vector spaces. Hence the most convenient set is typically chosen. There are three Bravais lattices, in addition to the simple cubic (sc) lattice, that are of enormous importance since a great variety of solids crystallise in these forms with a repeating unit like a molecule or atom at each lattice site, and serve to illustrate many of the general features of Bravais lattices which have just been outlined. These are the body-centred cubic (bcc) and face-centred cubic (fcc) Bravais Lattices:

1. bcc lattices are formed by adding to the sc lattice an additional point at the centre, as shown in Fig. 2.1. The bcc lattice can be easily shown to have more than one set of primitive vectors. If the original sc lattice is generated by primitive vectors $a\hat{x}, a\hat{y}, a\hat{z}$, where $\hat{x}, \hat{y}, \hat{z}$ form an orthonormal basis set

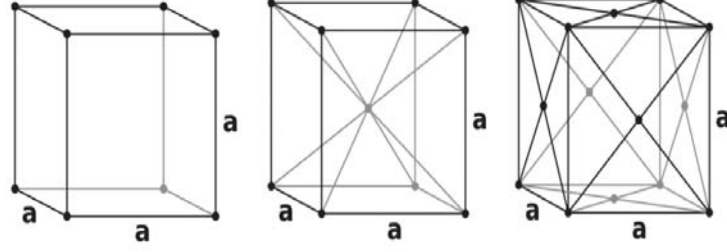


Figure 2.1: From left to right: the sc, bcc and fcc Bravais lattices. The solid lines indicate the bonds between the atoms. The lattice constant a (same on all sides) is also indicated.

and a is the length between points in the lattice, then a set of primitive vectors for the bcc lattice is:

$$\vec{a}_1 = a\hat{x} \quad \vec{a}_2 = a\hat{y} \quad \vec{a}_3 = \frac{a}{2}(\hat{x} + \hat{y} + \hat{z}) \quad (2.2)$$

A more symmetric set is:

$$\vec{a}_1 = \frac{a}{2}(\hat{y} + \hat{z} - \hat{x}) \quad \vec{a}_2 = \frac{a}{2}(\hat{z} + \hat{x} - \hat{y}) \quad \vec{a}_3 = \frac{a}{2}(\hat{x} + \hat{y} - \hat{z}) \quad (2.3)$$

2. fcc lattices are obtained by adding to the sc lattice an additional point in the centre of each square face, as shown in Fig. 2.1. A symmetric set of primitive vectors for the fcc lattice is:

$$\vec{a}_1 = \frac{a}{2}(\hat{y} + \hat{z}) \quad \vec{a}_2 = \frac{a}{2}(\hat{z} + \hat{x}) \quad \vec{a}_3 = \frac{a}{2}(\hat{x} + \hat{y}) \quad (2.4)$$

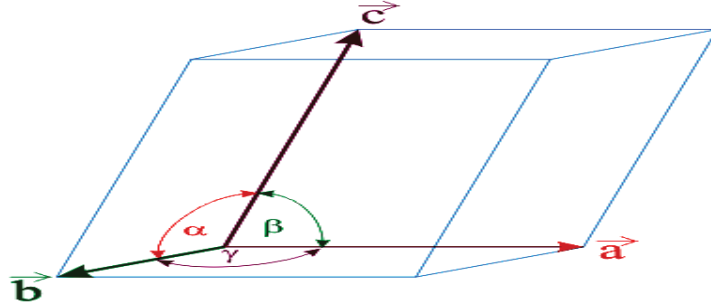


Figure 2.2: Notation for lattice parameters and angles between crystal axes.

The points in a Bravais lattice that are closest to a given point are called its nearest neighbours. Because of the periodic nature of a Bravais Lattice, each point has the same number of nearest neighbours. This number is thus a property of the Bravais lattice, and is referred to as the coordination number of the lattice. The sc lattice has coordination number of 6, bcc lattice of 8, and a fcc lattice of 12.

2.1.3 Unit Cells

The unit cell of a Bravais lattice is defined as a parallelepiped-shaped volume which, when reproduced by close-packing in 3D, gives the whole Bravais lattice. The shape of a parallelepiped is completely defined by the lengths of the three sides a , b and c , known as the crystal axes (or lattice constants) of the unit cell, and the values of the three angles α , β and γ , such that α lies between the b and c axes, β between the a and c axes, and γ between the a and b axes, as shown in Fig. 2.2. Most Bravais lattices can have many possibilities for a unit cell, which differ between them in the set of translations necessary to fill the whole space without either overlapping or leaving voids.

As explained below, there are two main choices for unit cells, the Primitive

unit cell and the Conventional unit cell. Both types of unit cells are always parallelepipeds, with lattice points at all eight corners. The main difference is that primitive unit cells contain only one lattice point, whereas the conventional ones contain more than one lattice point. In particular, in primitive unit cells each lattice point is shared between eight unit cells, whereas in conventional unit cells, in addition to the eight lattice points at the corners of the parallelepiped, there is a lattice point in the interior or on the surface of the cell. It can be deduced that the representations of the bcc and fcc lattices in Fig. 2.1 correspond to conventional unit cells, while that of the sc lattice can be thought of as being either primitive or conventional, since the two are equivalent in this special case. In general, conventional unit cells are more useful since they convey more readily the overall symmetry of the underlying lattice. It follows that the conventional unit cell is normally the most symmetrical among all the possible choices for a unit cell, such that they contain all the symmetry elements of the lattice that they represent. Primitive unit cells on the other hand are always minimum-volume cells, which is a property of the Bravais lattice. Even with this definition however there is still no unique way of choosing a primitive unit cell for a given Bravais lattice. The obvious primitive unit cell to associate with a particular Bravais lattice is the parallelepiped whose sides are delineated by the three primitive translation vectors, used to define the Bravais lattice points, applied to the lattice point occupying one of the eight corners. The Primitive unit cell has often the disadvantage of not displaying the full symmetry of the Bravais lattice. For example the primitive unit cells of the bcc and fcc lattices for the choice of primitive vectors of Eq. (2.3) and Eq. (2.4) respectively are oblique parallelepipeds as shown in Fig. 2.3, and therefore do not have the full cubic

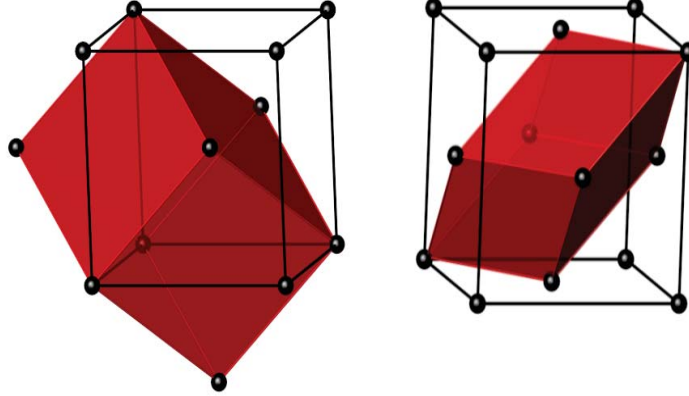


Figure 2.3: Conventional and primitive (shaded in red) unit cells for the bcc (left) and fcc (right) Bravais lattices. For the bcc lattice, the primitive cell has half the volume of the conventional, whereas for the fcc lattice it has one quarter the volume of the cube. In both cases the primitive cells have obviously much less symmetry than the conventional.

symmetry of the lattices in which they are embedded.

In general there are three main types of Conventional unit cells which together suitably represent the symmetry of all possible Bravais lattices. The distinction between these is based on the position of the lattice points within the unit cell. They are known as Side-centred, Face-centred and Body-centred, and are conventionally referred to with letters as explained below:

1. Side-centred unit cells (letter designation A,B or C): apart from the lattice points on the eight corners of the unit cell, they have lattice points in the middle of the pair of faces perpendicular to one of the crystal axes. The letter designation corresponds to the perpendicular crystal axis a , b or c .
2. Face-centred unit cells (letter designation F): effectively equivalent to simultaneous A,B and C-face centring. fcc lattices belong to this category.

Table 2.1: Number of lattice points contained in the various classes of unit cells presented so far in this section.

Unit cell type	Z
Primitive	1 (by definition)
Conventional - Side centred	2
Conventional - Body centred	2
Conventional - Face centred	4

3. Body-centred unit cells (letter designation I): apart from the points on the eight corners of the unit cell, they have a lattice point in the middle of the unit cell. bcc lattices belong to this category.

Table 2.1 summarises the number of lattice points Z in each type of unit cell.

2.1.4 Miller index labelling system for crystal planes

The orientation of a crystal plane is determined by three points in the plane, provided that they are not collinear. We can specify the orientation of a plane by the indices determined according to the following rules:

1. We first find the intercepts of the plane with the three crystal axes a , b and c , in the units of the crystal axes themselves. The axes may be those of a primitive or non-primitive unit cell, as long as a consistent set of axes is used to interpret the indices unambiguously. When the Bravais lattice is cubic for example, the simple cubic crystal axes of the conventional unit cell are invariantly used.
2. We then take the reciprocals of these intercepts and reduce them to the smallest three integers with no common factors. The result, enclosed in parentheses (hkl), are called the *Miller indices* of the plane.

-
3. For an intercept of infinity, that is if the plane is parallel to a given axis, the corresponding index is zero.
 4. If a plane crosses the origin, then it suffices to take any parallel plane not passing through the origin to find its indices.
 5. The indices may denote a single plane or a whole family of parallel lattice planes (since parallel planes in general have proportional intercepts).
 6. If a plane cuts an axis on the negative side of the origin, the corresponding index is negative, indicated by placing a bar above the index like so: $(h\bar{k}l)$.
 7. The Miller indices depend on the choice of crystal axes and hence are related to a particular unit cell. They are therefore not uniquely defined for a given crystal face. However, no matter which unit cell is chosen, one can always find a set of Miller indices for each face.

The indices of some important planes in a cubic crystal are illustrated in Fig. 2.4 for representative purposes.

There is also a notation for planes that are equivalent by virtue of the symmetry of the crystal, which can be denoted by curly brackets around the indices. For example, the faces of a cubic crystal are (100) , (010) , (001) , $(\bar{1}00)$, $(0\bar{1}0)$ and $(00\bar{1})$. Collectively, these planes are referred to as the $\{100\}$ planes. In general, one uses $\{hkl\}$ to refer to the (hkl) planes and all those that are equivalent to them by virtue of the crystal symmetry.

An index system also exists to label directions in crystals. The indices $[uvw]$ of a direction in a crystal are the set of the smallest integers that have the ratio of the components of a vector in the desired direction. A particular choice of crystal

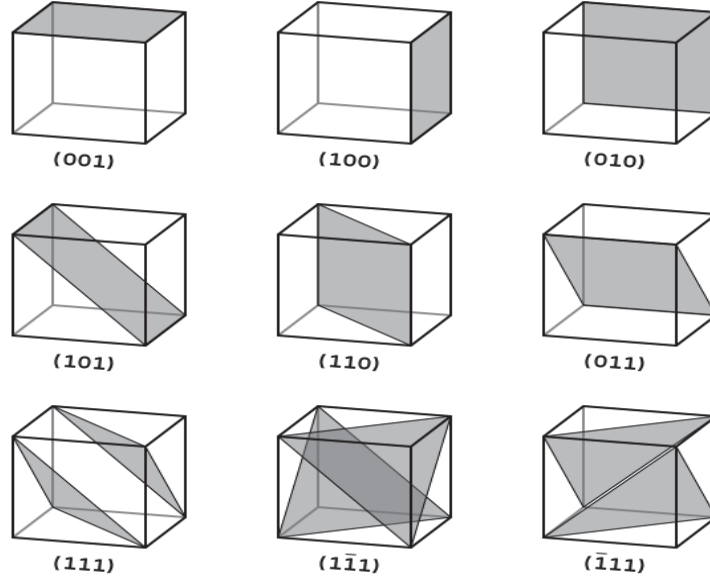


Figure 2.4: The Miller indices (hkl) of the symmetry planes in a cubic crystal.

axes \vec{a}_1 , \vec{a}_2 and \vec{a}_3 serves as the basis set, such that they lie in the $[100]$, $[010]$ and $[001]$ directions respectively. Thus in general the lattice point $n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$ lies in the direction $[n_1n_2n_3]$ from the origin. In cubic crystals the direction $[hkl]$ is perpendicular to a plane (hkl) having the same indices, but this is not generally true in other crystal systems. The notation used to label directions related by symmetry collectively (as is done for lattice planes) is to enclose the indices in triangular brackets $\langle uvw \rangle$. Hence for example the $[100]$, $[010]$, $[001]$, $[\bar{1}00]$, $[0\bar{1}0]$ and $[00\bar{1}]$ directions in a cubic crystal are referred to, collectively, as the $\langle 100 \rangle$ directions.

2.1.5 Lattice with a basis: crystal structures

A physical crystal is described by its underlying Bravais lattice together with its basis (or repeated structural unit). The basis describes the arrangement and

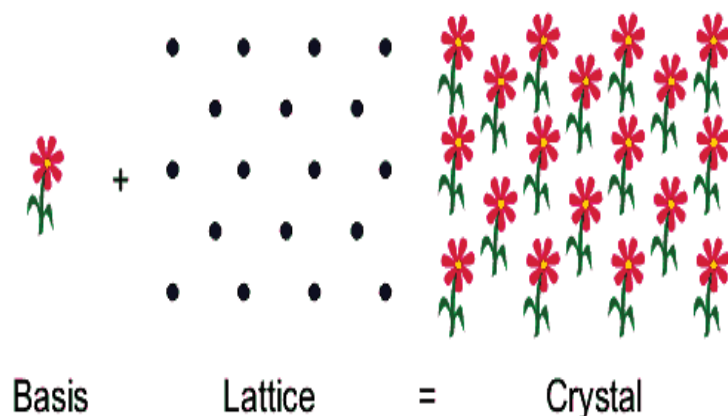


Figure 2.5: The crystal structure of a solid is formed by the convolution of the basis to every lattice point in the Bravais lattice.

composition of the chemical unit (be it an atom or molecule) associated with each lattice point in the Bravais lattice. A basis of atoms is attached to every lattice point, with every basis identical in composition and orientation. The basis does not have to be centred on the lattice point, as it can also be displaced relative to it. In general, the basis can consist of single atoms (as in metals), groups of atoms, molecules (as in organic crystals), groups of molecules or ions. The atoms or molecules in the basis in particular can be different from one another. The fact that many atoms can be associated with a lattice point means that even a primitive unit cell can contain many atoms. The basis convolved with the underlying Bravais lattice therefore gives the physical *crystal structure* of the crystalline solid, consisting in identical copies of the same basis located at the locations of all the points of a Bravais lattice. The concept of crystal structure is illustrated in Fig. 2.5.

It is possible for one compound to adopt more than one crystal structure under different conditions of crystallisation (such as different rates of crystallisa-

tion, different temperatures or pressures, different solvents). This phenomenon is called polymorphysism, and is important because different crystal structures can have drastically different physical and chemical properties. In the context of the research presented in this thesis, an important example of crystal structure which will be considered is the cubic diamond crystal structure portrayed in Fig. 2.6. The cubic diamond structure consists of two interpenetrating fcc Bravais lattices, displaced along the body diagonal of the cubic cell by one quarter the length of the diagonal. Hence it corresponds to a single fcc lattice with a two atom basis at atomic positions (000) and $(\frac{1}{4}\frac{1}{4}\frac{1}{4})$ relative to the crystal axes of the conventional cubic unit cell. Its coordination number is consequently 4, with the four nearest neighbours of each point forming the vertices of a regular tetrahedron. The diamond structure is not a Bravais lattice, because the environment of any point differs in orientation from the environments of its nearest neighbours. The structure bears its name from the crystal structure of diamond, where the two-atom basis consists of Carbon atoms.

2.1.6 Symmetry transformations of Bravais Lattices and Crystal Structures

There are two kinds of symmetry in crystallography: external, which relates to crystal macroscopic shapes, and internal, which relates to the atomic arrangement in the underlying unit cells. Clearly the symmetry of a macroscopic crystal as a whole is inherently connected with the symmetry of the constituent unit cells. All the symmetry elements of the crystal are in fact guaranteed to be always found in its unit cells. From the point of view of symmetry, a Bravais Lattice/Crystal

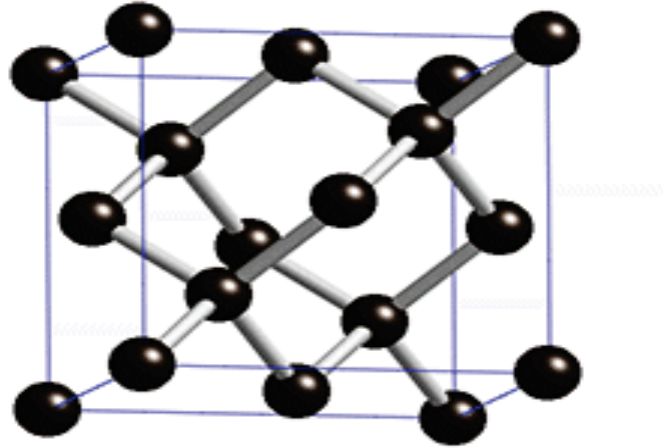


Figure 2.6: Conventional cubic unit cell of the cubic diamond crystal structure, with the nearest-neighbour bonds drawn between the atoms.

Structure is characterised by all rigid operations that transform the lattice into itself, and under which the lattice is therefore invariant . In this context rigid means operations that preserve the distance between all lattice points. When a rigid operation transforms the lattice into itself, it is known as a symmetry transformation or symmetry element of the Bravais lattice/Crystal structure. The set of all symmetry transformations of a Bravais lattice/Crystal structure is known as the *symmetry group* or *space group* of the Bravais lattice/Crystal structure. Even though symmetry elements apply to the whole lattice, we can consider just a single unit cell that exhibits the full symmetry of the Bravais lattice and that therefore has the same symmetry elements of the full space group of the Bravais lattice. This is known as a symmetry unit cell, and often consists in a conventional and not primitive unit cell. The symmetry transformations in the

space group of a Bravais lattice consist in: ¹

1. Translations: described by a translation vector which is applied equally to all lattice points. A translation is a symmetry operation of the Bravais lattice if and only if the translation vector \hat{T} has the form $\hat{T} = \sum_i n_i \vec{a}_i$, and therefore corresponds to any of the possible Bravais lattice vectors (or their sums or differences). In a translation all points are translated, and so no point stays fixed.
2. Rotations: described by a rotation axis passing through at least one lattice point, and a rotation angle. For a rotation to be a symmetry transformation of a Bravais lattice, the rotation angle must be an integral multiple of $2\pi/n$ radians, where n is an integer. The axis is then called an n -fold rotation axis. A Bravais lattice can contain only 1- (trivial case of identity), 2- (diad), 3- (triad), 4- (tetrad), or 6-fold (hexad) axes. Hence for an n -fold rotation axis, you need to rotate the lattice n times to get back to the initial position (2π rotation), with each step being a symmetry operation in its own right. In a rotation, the only points that stay fixed are those that lie on the rotation axis.
3. Reflections: described by a mirror plane containing at least a lattice point. The object is replaced with its mirror image with respect to the mirror plane. In a reflection, the only points that stay fixed are those that lie on the mirror plane.

¹These symmetry transformations are applicable to both Bravais lattices and Crystal structures, unless otherwise stated. This does not however mean that a crystal structure has necessarily the same set of symmetry elements as its underlying Bravais lattice.

-
4. Inversions: described by an inversion point corresponding to a lattice point P . This inversion point takes the point with coordinates \vec{r} (with respect to P as the origin) into $-\vec{r}$. The inversion operation is composed of a rotation of π followed by reflection in a plane normal to the rotation axis. In an inversion, the inversion point is the only point that stays fixed. The inversion is a symmetry transformation for the lattice if for every point at vector position \vec{r} there is an equivalent point located at $-\vec{r}$, in which case the inversion point is called centre of symmetry and the lattice is said to be centro-symmetric with respect to that point.

Following the four symmetry operations outlined above, there are three more operations that can be obtained by combining two of the above transformations.

1. Rotation-Inversions (Improper rotations): sometimes a rotation through $2\pi/n$ followed by an inversion with respect to a point lying on the rotation axis is a symmetry transformation, even though the rotation by itself is not. The axis is then called an n -fold inversion axis, and is labelled as \bar{n} . Since $\bar{1}$ and $\bar{2}$ are equivalent to a centre of symmetry and mirror plane respectively, they are not included here as inversion axes (the only inversion axes are $\bar{3}$, $\bar{4}$, $\bar{6}$). In a rotation-inversion, the only point that stays fixed is the inversion point on the inversion axis.
2. Screw Axes (Crystal structures only): a crystal structure with a screw axis is transformed into itself by translation through a non-Bravais lattice vector, followed by a rotation about an axis parallel to the translation vector (the screw axis). In general the symmetry element R_d along the a direction involves a rotation $2\pi/R$ about the a axis followed by a displacement

$$(D/R) \vec{a}.$$

3. Glide planes (Crystal structures only): a lattice with a glide plane is transformed into itself by translation through a non-Bravais lattice vector, followed by a reflection with respect to a plane parallel to the translation vector (the glide plane). The description of this symmetry element is simplified by reference to the vectors \vec{a} , \vec{b} and \vec{c} which define the edges of the unit cell. One may have an a -glide plane parallel to the a crystal axis, and also b - and c - glide planes parallel to b and c respectively. An n -glide plane is one which, if perpendicular to c , gives a displacement component $1/2a + 1/2b$.

Fig. 2.7 summarises all the symmetry operations and gives their graphical and written symbols used to represent them.

In general any symmetry operation of a Bravais lattice can be compounded out of two other symmetry operations: a translation through a Bravais lattice vector \vec{R} and a rigid operation leaving at least one lattice point fixed (the translation by itself leaves no point fixed). These are known as point operations, and include all symmetry operations that are not translations. As explained above, the types of point operations are Inversion, Reflection, Rotation and Rotation-Inversion. The subset of all point symmetry operations is known as the point group of the Bravais lattice. There turn out to be only seven distinct point groups that a Bravais lattice can have, known as crystal systems. When one relaxes the restriction to point operations and considers the full symmetry group of the Bravais lattice (including therefore translational symmetries), there turn out to be 14 distinct space groups that a Bravais lattice can have. Thus from the point of view of symmetry (space

Type of symmetry element	Written symbol	Graphic symbol	
Centre of symmetry	$\bar{1}$	\circ	
		Perpendicular to paper	In plane of paper
Mirror plane	m		
Glide planes	$a\ b\ c$		
		Glide in plane of paper	Arrow shows glide direction
Rotation	n		
	2		
	3		
	4		
	6		
Screw axes	2_1		
	$3_1, 3_2$		
	$4_1, 4_2, 4_3$		
	$6_1, 6_2, 6_3, 6_4, 6_5$		
Inversion axes	$\bar{3}$		
	$\bar{4}$		
	$\bar{6}$		

Figure 2.7: Symmetry operations in Bravais lattices and Crystal structures, together with the symbols used to represent them.

groups) there are 14 different kinds of Bravais lattices. A summary of the names and structural properties of the conventional unit cells (which also correspond to the symmetry unit cells) of these 14 Bravais lattices is presented in Fig. 2.8. Lines connect the lattice points to clarify the relationships between them. In the absence of any symmetry, the three axes of these conventional unit cells have different lengths and the three angles are different from each other and from 90

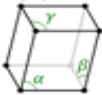
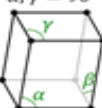
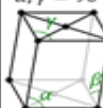
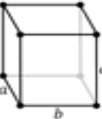

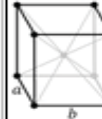
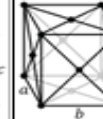


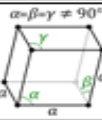

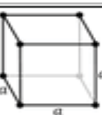
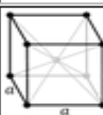
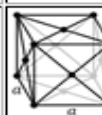
The 7 lattice systems		The 14 Bravais lattices				
Triclinic	P					
	$\alpha, \beta, \gamma \neq 90^\circ$ 					
Monoclinic	P	C				
	$\beta \neq 90^\circ$ $\alpha, \gamma = 90^\circ$ 	$\beta \neq 90^\circ$ $\alpha, \gamma = 90^\circ$ 				
Orthorhombic	P	C	I	F		
	$a \neq b \neq c$ 	$a \neq b \neq c$ 	$a \neq b \neq c$ 	$a \neq b \neq c$ 		
Tetragonal	P	I				
	$a \neq c$ 	$a \neq c$ 				
Rhombohedral	P					
	$\alpha = \beta = \gamma \neq 90^\circ$ 					
Hexagonal	P					
						
Cubic	P (bcc)	I (bcc)	F (fcc)			
						

Figure 2.8: The seven crystal systems and 14 Bravais lattices, as defined from the symmetry elements that they possess. The geometry and type of conventional unit cell is indicated for each Bravais lattice. The symmetry of the crystal systems increases moving down the table.

°. In the presence of symmetry elements, there are restrictions and special values of the unit cell parameters.

Table 2.2: Number of point and space groups of Bravais lattices and crystal structures.

	Bravais lattice (Basis of spherical symmetry)	Crystal Structure (Basis of arbitrary symmetry)
No. point groups:	7 (the seven crystal systems)	32 (the 32 crystallographic point groups)
No. space groups:	14 (the 14 Bravais lattices)	230 (the 230 Crystal structures)

We now perform a similar analysis not on Bravais lattices but on general crystal structures. We take as the basis of the crystal structure an arbitrary object, and try to classify the symmetry groups of the crystal structure so obtained. The symmetry transformations that transform the crystal structure into itself now depend both on the symmetry of the object (internal symmetry of the basis) and the symmetry of the underlying Bravais lattice. Because the objects no longer have the maximum spherical symmetry like the individual lattice points of a Bravais lattice, the number of symmetry transformations and possible symmetry groups is greatly increased. There turn out to be 230 different space groups that a lattice with a basis can have, and therefore from the point of view of symmetry there are 230 distinct crystal structures. The possible point groups of a general crystal structure have also been enumerated. These contain the symmetry operations that take the crystal structure into itself while leaving one point fixed (non-translational point operations). There are 32 distinct point groups that a crystal structure can have, known as the 32 crystallographic point groups. A summary of the number of point and space groups for both Bravais lattices and crystal structures is presented in table 2.2. Table 2.3 instead shows how the numbers of space and point groups are split across the seven crystal systems.

Table 2.3: Numbers of point and space groups in each crystal system.

Crystal system	No. of point groups	No. of Bravais lattices	No. of space groups
Triclinic	2	1	2
Monoclinic	3	2	13
Orthorhombic	3	4	59
Tetragonal	7	2	68
Rhombohedral	5	1	25
Hexagonal	7	1	27
Cubic	5	3	36
Total	32	14	230

2.1.7 Nomenclature for crystallographic point groups and space groups

Due to the sheer number of crystallographic point groups (32) and crystal structure space groups (230), it is useful to introduce a nomenclature system, instead of resorting to general names like in the case of Bravais lattices. Here we shall consider only the international nomenclature system (also known as Hermann-Mauguin notation). The conventions used to list the symmetry elements in the group are the same for both space groups and point groups. We shall consider first the space group nomenclature as it is more general than the one for point groups. Space groups are labelled according to a general name structure of the form L_{ijk} , where in general L denotes the type of symmetry unit cell of the underlying Bravais lattice and ijk denote the symmetry elements of the space group for the different symmetry directions.

The first thing that has to be specified in the name of a space group of a crystal structure is the type of symmetry unit cell of the corresponding Bravais lattice. As explained before, for each of the 14 Bravais lattices the symmetric unit cell can either be Primitive or Conventional. The meaning of the capital letters

(P,A,B,C,F,I) assigned to the various types of symmetry unit cells has already been explained in section 2.1.3. The symmetry unit cells for all 14 Bravais lattices were presented in Fig. 2.8.

The letters and numbers that follow the initial capital letter are the sequence of symmetry elements of the space group. The symbols for the symmetry elements follow the convention of Fig. 2.7. Their position in the sequence relative to the capital letter tells in what symmetry direction they are either orthogonal to or coincident with. From the point of view of space group nomenclature, it is convenient to divide symmetry operations in the following two groups:

1. Operations with reflections (mirror planes and glide planes), which are always noted by small letters (a,b,c,n or m), are positioned according to which direction the mirror plane is orthogonal to.
2. Operations with rotation axes (simple rotations, screw axes, inversion axes) are positioned according to which direction the rotation axis lies in.
3. The centre of symmetry operation has no reference to directions.

Finally, to include multiple symmetry elements along the same direction one uses a "/" symbol. Hence, for example, the notation $2/m$ means that for a given symmetry direction there is a diad rotation axis along that direction and a mirror plane perpendicular to it. The symmetry directions i, j, k do not necessarily have to be along the crystal axes a, b, c . The directions of the i, j, k symmetry axes in each of the crystal systems is explained in table 2.4

Hence for the monoclinic and orthorhombic systems (where all three crystal axes are different in length), the symmetry directions and the cell axes are coincident. In crystal systems where at least one axis is equal to another axis

Table 2.4: Directions of the symmetry labels ijk in the space group nomenclature relative to the crystal axes.

System	Symmetry directions
Triclinic	none
Monoclinic	i - along a axis j - along b axis k - along c axis
Orthorhombic	i - along a axis j - along b axis k - along c axis
Tetragonal	i - along c axis (4-fold) j - along a and b axes k - along diagonal between a and b
Trigonal/Hexagonal	i - along c axis (3-fold) j - along a and b axes k - along diagonal between a and b
Cubic	i - along a and b and c j - along diagonal between a and b , b and c , and a and c k - along diagonal between a and b and c

(tetragonal, cubic, trigonal, hexagonal), the crystal axes now are not coincident with the symmetry directions. In general, if a symmetry operation exists along/perpendicular one axis, then a similar operation will exist down the equal axis. Operations along/perpendicular to equivalent axes are not repeated to avoid redundancy. In general, only the fundamental symmetry elements are included in space group notation. We know that if two symmetry operations meet in space, a third operation is formed, and elements that arise directly as a result of the presence of other elements are discarded to avoid redundancy. For example two orthogonal mirror planes automatically generate a diad axis along the line of their intersection, which need not be stated to avoid redundancy. The space group notation does not therefore necessarily need to note all symmetry operations in the unit cell. Some of them can be implied from the notation itself.

The same exact conventions apply to the naming of crystallographic point groups, except that the capital letter labelling the unit cell type is not mentioned. Of course, point groups include only point operations for the crystal structure, so that glide planes and screw axes never appear.

2.1.8 The reciprocal lattice

The lattice reciprocal to a Bravais lattice (the reciprocal lattice) plays a fundamental role in the analysis of how periodic structures behave under diffraction by, for example, X-rays. The definition of reciprocal lattice stems from considering a set of Bravais lattice vectors \vec{R} constituting a Bravais lattice, and a plane wave $e^{i\vec{k}\cdot\vec{r}}$. For a general wave vector \vec{k} such a plane wave will not of course have the periodicity of the Bravais lattice, but for certain special choices it will. The reciprocal lattice of a Bravais lattice embodies the set of all wave vectors \vec{K} (or \vec{G}) that yield plane waves with the periodicity of a given Bravais lattice. The \vec{K} vectors are then known as the reciprocal lattice vectors. Analytically, \vec{K} belongs to the reciprocal lattice of a Bravais lattice of points \vec{R} provided that the relation

$$e^{i\vec{K}\cdot(\vec{r}+\vec{R})} = e^{i\vec{K}\cdot\vec{r}} \quad (2.5)$$

holds for any \vec{r} , and for all \vec{R} in the Bravais lattice. Factoring out $e^{i\vec{K}\cdot\vec{r}}$, we can characterize the reciprocal lattice as the set of wave vectors \vec{K} satisfying

$$e^{i\vec{K}\cdot\vec{R}} = 1 \quad (2.6)$$

for all \vec{R} in the Bravais lattice. It must be emphasised that a reciprocal lattice is always defined with reference to a particular Bravais lattice. The Bravais lattice

that determines a given reciprocal lattice is then referred to as the direct lattice, when viewed in relation to its reciprocal. It must also be noted that, given the above definition, the reciprocal lattice lies in the Fourier (reciprocal) space associated with the crystal. \vec{r} and \vec{k} therefore constitute a pair of reciprocal (conjugated) coordinates. In particular, if vectors in the direct lattice have the dimensions of [length], vectors in the reciprocal lattice have the dimensions of [1/length].

The reciprocal lattice is itself a Bravais lattice, according to both definitions of Bravais lattice. The proof of this provides us with an explicit algorithm for constructing the reciprocal lattice. Let \vec{a}_1 , \vec{a}_2 and \vec{a}_3 be a set of primitive translation vectors for the direct lattice. Then it can be shown that the reciprocal lattice can be generated by the following three reciprocal primitive translation vectors:

$$\begin{aligned}\vec{b}_1 &= 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \\ \vec{b}_2 &= 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \\ \vec{b}_3 &= 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}\end{aligned}\tag{2.7}$$

To verify that the above set of equations gives a set of primitive vectors for the reciprocal lattice, one first notes that the \vec{b}_i vectors satisfy:

$$\vec{b}_i \cdot \vec{a}_j = 2\pi \delta_{ij}\tag{2.8}$$

where $\delta_{ij} = 0$ if $i \neq j$, or $\delta_{ij} = 1$ if $i = j$. Since the set of \vec{b}_i vectors form a complete basis set in reciprocal space, any vector \vec{k} in reciprocal space can be

written as a linear combination:

$$\vec{k} = k_1 \vec{b}_1 + k_2 \vec{b}_2 + k_3 \vec{b}_3 \quad (2.9)$$

If \vec{R} is a Bravais lattice vector, then:

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 \quad (2.10)$$

where the n_i coefficients are integers. Hence the dot product of \vec{k} with \vec{R} is given by:

$$\vec{k} \cdot \vec{R} = 2\pi (k_1 n_1 + k_2 n_2 + k_3 n_3) \quad (2.11)$$

where we have made use of the results of Eq. (2.8). For $e^{i\vec{k} \cdot \vec{R}}$ to be unity for all \vec{R} as required for a reciprocal lattice, $\vec{k} \cdot \vec{R}$ must be 2π times an integer for any choices of the integers n_i . This requires the coefficients k_i to be integers. Thus the condition that \vec{K} be a reciprocal lattice vector is satisfied by just those vectors for which:

$$\vec{K} = k_1 \vec{b}_1 + k_2 \vec{b}_2 + k_3 \vec{b}_3 \quad (2.12)$$

where the k_i coefficients are integers. This demonstrates that the reciprocal lattice is a Bravais lattice and that the \vec{b}_i vectors can be taken as the reciprocal primitive translation vectors. Since the reciprocal lattice is itself a Bravais lattice, one can in theory construct its reciprocal, but this turns out to be nothing but the original direct lattice. A further important property of reciprocal lattices is that, if v is the volume of a primitive unit cell in the direct lattice, then the

primitive cell of the reciprocal lattice has a volume of $(2\pi)^3/v$. Some important examples of reciprocal lattices are:

1. Simple Cubic Bravais lattice, with cubic unit cell of side a , has as its reciprocal a simple cubic lattice with cubic unit cell of side $2\pi/a$.
2. The fcc Bravais lattice with a cubic conventional unit cell of side a has as its reciprocal a bcc lattice with cubic conventional unit cell of side $4\pi/a$
3. The bcc lattice with cubic conventional unit cell of side a has as its reciprocal lattice a fcc lattice with cubic conventional unit cell of side $4\pi/a$

There is a close relation between reciprocal lattice vectors and planes of points in the direct lattice. Given a particular Bravais lattice, a lattice plane is defined to be any plane containing at least three non-collinear Bravais lattice points. By a family of lattice planes, we mean a set of parallel, equally spaced lattice planes, which together contain all the points of the three-dimensional Bravais lattice. Any lattice plane is member of such a family. The reciprocal lattice provides a very simple way to classify all possible families of lattice planes through the following theorem: for any family of lattice planes separated by a distance d , there are reciprocal lattice vectors perpendicular to the planes, the shortest of which have a length of $2\pi/d$. Conversely, for any reciprocal lattice vector \vec{K} , there is a family of lattice planes normal to \vec{K} and separated by a distance d , where $2\pi/d$ is the length of the shortest reciprocal lattice vector parallel to \vec{K} .

The correspondence between reciprocal lattice vectors and families of lattice planes is connected to the Miller index notation for the lattice plane. It turns out in fact that the Miller indices of a family of lattice planes are the coordinates

of the shortest reciprocal lattice vector normal to that plane with respect to a set of reciprocal primitive translation vectors. Thus a plane with Miller indices h, k, l is normal to the reciprocal lattice vector $h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3$, and this vector is of length $2\pi/d$.

2.2 X-ray diffraction

2.2.1 Introduction to crystal structure determination

In diffraction, or light scattering, studies, we usually keep the wavelength of light fixed and measure the variation in intensity with direction as a result of the scattering process. From these measurements it is possible to get structural information about the sample like bond length, bond angles and torsion angles, and hence obtain a complete geometrical description of the crystal structure. This technique is known as crystal structure determination. X-rays are the perfect candidates for resolving atoms and molecules in crystal structures through light diffraction due to the similarity between their characteristic wavelength and the typical interatomic spacings found in crystalline solids, both of which are of the order of a few angstroms (10^{-10} m) in size.

The major drawback of the technique of X-ray Diffraction (XRD) is that no lens mechanism exists to physically recombine the diffracted rays to create the realspace image of the crystal structure, as in a normal optical microscope. This procedure therefore has to be performed mathematically through Fourier Transform techniques. One very important consequence of this limitation is that some of the information in the scattered X-rays is lost. When the X-ray diffraction

pattern is recorded, the individual scattered wave amplitudes are retained as relative intensities ¹, but the relative phases are lost. This makes the mathematical reconstruction stage to find the real image of the crystal structure via Fourier Transform much less straightforward. This is known as the phase problem, and there are methods to circumvent it.

The object studied in XRD experiments is typically a small crystalline sample in the form of a singlecrystal ². In a single crystal, all the unit cells of the crystal periodic lattice are aligned in the same orientation with longrange three dimensional order and hence scatter cooperatively to give a clear pattern of discrete diffracted beams. When the method of crystal structure determination is successful, it provides an image of the crystal structure. More precisely, it locates the components of the material which interact with the incident X-rays and scatter them. These are the atomic electrons in the atom. In particular, the picture that results is of a time-averaged (with respect to thermal motion) electron density, with concentration of electron density corresponding to atoms. We shall now see how the distribution of X-rays scattered by a rigid, periodic array of ions in the crystal structure reveals the locations of the ions within the structure. There are two equivalent ways to view the scattering of X-rays by a perfect periodic structure, due to Bragg and Von Laue, both of which are widely used.

2.2.2 The Bragg formulation of X-ray diffraction

The conditions for constructive interference between Xrays scattered from adjoining parallel lattice planes in a perfectly periodic structure of Bravais lattice

¹The intensity is proportional to the square of the amplitude.

²A single molecule would be impossible to hold in the Xray beam for the duration of the experiment and would give an immeasurably weak scattering pattern

points are dictated by Braggs law, given by:

$$n\lambda = 2d \sin \theta \quad (2.13)$$

where, as shown in figure 2.9, θ is the angle between the incident Xray beam and the family of parallel lattice planes, d the spacing between two neighbouring planes, λ the wavelength of radiation and n is known as the order of diffraction for the given family of lattice planes. Whenever equation (2.13) is satisfied for a given wavelength and incident direction, intense peaks of scattered radiation known as Bragg peaks are observed on the diffraction pattern. Reflections are possible in general for any of the infinitely many ways of sectioning the lattice into planes. The overall diffraction pattern of a regular single crystal therefore consists in a series of discrete diffraction peaks (or diffraction spots) arranged in a regular geometrical fashion in reciprocal (Fourier) space, each spot corresponding to an individual scattered Xray beam travelling in a definite direction from the crystal.

2.2.3 The Von Laue formulation of X-ray diffraction

The Von Laue approach differs from the Bragg approach in that no particular sectioning of the crystal into lattice planes is singled out. Instead one regards the crystal structure as composed of identical microscopic objects (set of ions or atoms) placed at the sites R of a Bravais lattice, each of which can re-radiate the incident radiation in all directions. Sharp peaks will be observed only in directions and at wavelengths for which the rays scattered from all lattice points interfere constructively. It can be shown that this condition can be written as:

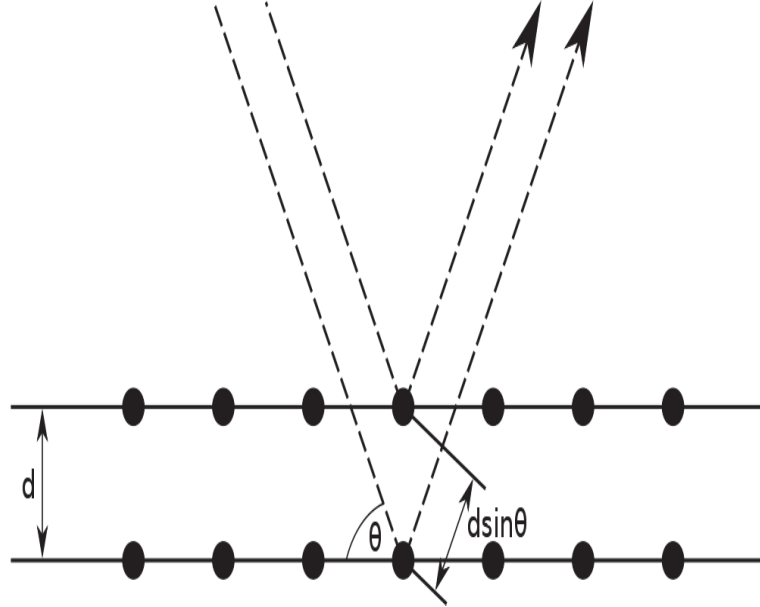


Figure 2.9: A Bragg specular reflection from a particular family of lattice planes, separated by a distance d . Incident and reflected rays are shown for the two neighbouring planes. The path difference between the two rays is $2d \sin \theta$, yielding formula (2.13) for constructive interference.

$$e^{i(\vec{k}' - \vec{k}) \cdot \vec{R}} = 1 \quad (2.14)$$

for all Bravais lattice vectors \vec{R} , where \vec{k} is the incident and \vec{k}' the scattered wavevectors. Comparing this condition with the definition of reciprocal lattice of Eq. (2.6), we arrive at the Laue condition that constructive interference between all scattered rays from all lattice points (Laue diffraction peak) will occur provided that the change in wavevector upon scattering, $\vec{K} = \vec{k}' - \vec{k}$, is a vector of the reciprocal lattice.

It is sometimes convenient to have an alternative formulation of the Laue condition, stated entirely in terms of the incident wavevector \vec{k} . For this we need to combine the above Laue condition with the condition for elastic scattering, that

is that \vec{k} and \vec{k}' have the same magnitude. This yields the following condition for a Laue diffraction peak to appear:

$$\vec{k} \cdot \hat{K} = \frac{1}{2}K \quad (2.15)$$

This means that the component of the incident wavevector \vec{k} along the reciprocal lattice vector \vec{K} must be half the length of \vec{K} . Thus an incident wavevector \vec{k} will satisfy the Laue condition for constructive interference between all scattered rays from all lattice points and give rise to a Laue diffraction peak if and only if the tip of the vector lies in a plane that is the perpendicular bisector of a line joining the origin of k-space to a reciprocal lattice point \vec{K} . Such k-space planes are called Bragg planes, each of which is associated with a particular diffraction peak. These planes are parallel to the family of direct lattice planes responsible for the diffraction peak in the Bragg formulation. Thus the connection between Bragg and Laue conditions for diffraction peaks is the following: a Laue diffraction peak corresponding to a change in wavevector given by the reciprocal lattice vector \vec{K} corresponds to a Bragg reflection from the family of direct lattice planes perpendicular to \vec{K} . The order n of the Bragg reflection is just the length of \vec{K} divided by the length of the shortest reciprocal lattice vector parallel to \vec{K} , \vec{K}_o .

Hence, to summarise, for a Bravais lattice each diffraction peak corresponds to a family of parallel lattice planes in the Bravais lattice (identified by Miller indices hkl), or equivalently to each possible reciprocal lattice vector. The connection between the two conditions for diffraction is expressed by the Miller indices: the Miller indices of a family of lattice planes are the coordinates of the shortest reciprocal lattice vector normal to that plane. In particular, the n -th order peak

from the plane family (hkl) corresponds to the reciprocal lattice vector $\vec{K} = n\vec{K}_o = nh\vec{b}_1 + nk\vec{b}_2 + nl\vec{b}_3$ and is therefore associated with Miller indices $n(hkl)$. Hence each diffraction peak in the diffraction pattern, including different orders from the same plane family, is associated with a distinct reciprocal lattice point and hence a distinct set of Miller indices hkl . The origin of the diffraction pattern has therefore always indices $h = 0, k = 0, l = 0$.

Since the reciprocal lattice associated with a given Bravais lattice is far more easily visualised than the set of all possible planes into which the Bravais lattice can be resolved, the Laue condition for diffraction peaks is more simple to work with than the Bragg condition. According to the Laue formulation, an incident wave vector \vec{k} will lead to a diffraction peak if and only if the tip of the wave vector lies on a k-space Bragg plane. Since the set of all Bragg planes is discrete, the Bragg planes together don't fill all of three-dimensional k-space, and in general the tip of \vec{k} will not lie on a Bragg plane. Thus for a fixed incident wave vector \vec{k} , that is for a fixed X-ray wavelength and fixed incident direction relative to the crystal axes, there will be in general no diffraction peaks. If one wishes to search experimentally for Bragg peaks the constraint of fixed \vec{k} must therefore be relaxed, either by varying the magnitude of \vec{k} (i.e. varying the wavelength of the incident beam) or varying its direction (in practice, varying the orientation of the crystal with respect to the incident direction).

2.2.4 The intensities of the diffracted waves and their relation to the crystal structure

So far we have been discussing the appearance of a diffraction pattern purely in terms of constructive interference between waves scattered by a periodic array of lattice points (the Bravais lattice), and therefore looking only at the relative phases of the waves diffracted by the crystal and using Bragg/Laue laws. This is good in predicting the points where constructive occurs and therefore where diffraction spots should occur in the diffraction pattern, that is to explain the overall geometry and symmetry of the diffraction pattern. However, considering only the underlying Bravais lattice of the crystal tells us nothing about the relative intensities of the diffraction spots. To do that we need to look at the atom composition of the basis associated with each lattice point, and hence at the full crystal structure. In a real crystal structure, with a molecular basis associated with each Bravais lattice point, the cloud of electrons surrounding each atom in the structure is the physical component that interacts with Xrays and gives rise to scattering. As a result, the diffraction spots will exhibit a varying level of intensity. This intensity is given by the direct Fourier Transform (FT) of the electron density distribution in a unit cell ¹ of the realspace crystal structure:

$$F(hkl) = \int_{cell} \rho(xyz) \exp \left[i \left(\vec{k} \cdot \vec{r} \right) \right] dV \quad (2.16)$$

As explained before, each diffraction spot is associated with a reciprocal lattice vector \vec{K} and its Miller indices hkl , so that the reciprocal variable k can in fact

¹In practice, a single crystal contains millions of identical unit cells which scatter cooperatively, so that all diffraction intensities are scaled by the same factor

be restricted to a reciprocal lattice point of the form $\vec{K} = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3$. The exponent in the FT can therefore be rewritten in an explicit form, using also the expression for the atomic coordinates for position within the unit cell in terms of the primitive vectors of the real Bravais lattice $\vec{r} = x\vec{a}_1 + y\vec{a}_2 + z\vec{a}_3$:

$$\vec{K} \cdot \vec{r} = (h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3) \cdot (x\vec{a}_1 + y\vec{a}_2 + z\vec{a}_3) = 2\pi(hx + ky + lz) \quad (2.17)$$

The forward FT therefore is:

$$F(hkl) = \int_{cell} \rho(xyz) \exp[2\pi i(hx + ky + lz)] dV \quad (2.18)$$

where $F(hkl) = |F(hkl)| \exp[i\phi(hkl)]$ is the structure factor of the wave diffracted in the direction hkl , which comprises information about both the phase $\phi(hkl)$ and the amplitude $|F(hkl)|$ of the reflection¹. $\rho(xyz)$ is the electron density at the position xyz in the unit cell. The structure factor (amplitude and phase) for reflection hkl is therefore given by taking the value of the electron density of the real crystal structure at each point in the unit cell $\rho(xyz)$, multiplying it by the complex number $\exp[2\pi i(hx + ky + lz)]$, and integrating these values over the whole unit cell volume. Positions in the unit cell are measured from one corner (the origin) of the unit cell and the atomic coordinates x, y, z are in fractions of the corresponding crystal axes a, b, c . A physical interpretation can be given to this process of FT in terms of adding single diffracted waves together with their correct relative phases to get the total diffracted wave in one direction. Expressing the continuous electron distribution in terms of a more convenient

¹Amplitude and intensity are related by $I(hkl) \propto |F(hkl)|^2$

form involving a finite sum over discrete atoms in the unit cell, this relation can be rewritten as the following discrete Fourier Transform:

$$F(hkl) = \sum_j F_j(hkl) = \sum_j f_j(\theta) \exp[2\pi i(hx_j + ky_j + lz_j)] \quad (2.19)$$

where the sum is over all the waves diffracted in the direction hkl by the individual atoms at position x_j, y_j, z_j in the unit cell, taking into account the relative phase differences. The summation is made over all the atoms forming the basis associated with one lattice point. The above equation therefore represents the combination of many waves each scattered by an individual atom in the unit cell, added with their correct relative amplitudes and phases to give one total resultant diffracted wave in each direction. The intensity of radiation in a given diffraction spot will depend on the extent to which the rays scattered from these atoms interfere constructively with one another. In some directions, complete destructive interference occurs resulting in zero intensity systematic absences at positions where diffraction spots would be expected from purely geometrical considerations. $f_j(\theta)$ is the atomic form factor of the j -th atom describing its scattering power, and in particular the variation of scattered intensity with angle θ away from the forward direction. The atomic form factor is measured in units of electrons, with $f(0)$, the form factor for zero deflection, being equal to the atomic number Z of the atom. This reflects the fact that X-rays interact with crystals through the atomic electrons and therefore the higher the Z of the atom, the higher the strength of the interaction. Hence for a given direction hkl the amplitude of the diffracted wave depends only on the type of atom, while the

phase depends on the positions of the atoms in the unit cell.

The resulting observed intensities in the diffraction pattern hold all the available information about the position of the atoms in the unit cell of the crystal structure. This can be seen by reversing the FT relation between the electron density in the unit cell and the diffracted waves expressed in equation (2.18):

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} |F(hkl)| \exp[i\phi(hkl)] \exp[-2\pi i(hx + ky + lz)] \quad (2.20)$$

where the sum is performed over all the discrete diffraction spots hkl in the pattern. The summation must be carried out for many different atomic coordinates x, y, z in order to show the variation of the electron density in the unit cell and hence to locate the atoms where the electron density is concentrated in peaks. While the diffraction amplitudes $|F(hkl)|$ and the relative phase shift appropriate to each geometrical position in the unit cell $\exp[-2\pi i(hx + ky + lz)]$ are known quantities, the diffraction pattern holds no information about the intrinsic phases $\phi(hkl)$ of each reflection, since all diffraction spots are recorded in the same observation plane with no information on their order of arrival. Without this information the image of the electron density $\rho(xyz)$ cannot be reconstructed from the diffraction pattern. This is the so-called phase problem mentioned earlier.

In summary, the process of crystal structure determination occurs as follows: we record the diffraction pattern from a crystal. Measurement of the diffraction pattern geometry and symmetry tells us the unit cell geometry and gives us some information about the symmetry of arrangement of the molecules in the unit

cell. Then from the individual intensities of the diffraction pattern we work out the positions of the atoms in the unit cell, adding together the individual waves with their correct relative amplitudes and phases. And here we encounter the phase problem, the fact that the measured diffraction pattern provides directly only the amplitudes and not the required phases, without which the FT cannot be made to find the real image of the crystal structure. This phase problem can be circumvented via sophisticated mathematical and statistical experimental techniques, whose description however lies beyond the scope of this Thesis.

2.2.5 Systematic absences and the diffraction pattern of the cubic-diamond crystal structure

According to Laue's law, reflections in the diffraction pattern of a Bravais lattice can occur for all integer hkl combinations. When we consider crystal structures however, things get more complicated because the presence of atoms affects the relative intensities of diffraction spots as explained in the previous section. It turns out that for some arrangements of atoms in the crystal structure, certain diffraction spots corresponding to perfectly good reciprocal lattice vectors \vec{K} disappear. This gives rise to so-called systematic absences in the form of hkl diffraction spots with zero intensity. The indices of systematic absences can be understood in terms of Eq. (2.19) for calculating the structure factor. If, as a result of this addition of waves diffracted by the atoms in the unit cell we have zero intensity at a spot hkl (structure factor $F(hkl) = 0$, then we have a systematic absence. This occurs when the atoms in the unit cell are so arranged that there is a complete destructive interference among all waves in the

hkl direction in question. Exactly which combinations of hkl are allowed in the diffraction pattern, and hence the set of systematic absences, is a property of the space group of the crystal structure, since this destructive interference is a result of the symmetry properties of the crystal structure

Consider for example this very simple crystal structure: a simple cubic Bravais lattice with a two-atom monatomic basis at positions $(0, 0, 0)$ and $(1/2, 1/2, 1/2)$ inside the unit cell. This configuration is equivalent to a bcc Bravais lattice considered in its conventional two-lattice points unit cell. Since we are considering a monatomic crystal structure, atomic form factors can be ignored, and only their relative phase will determine the appearance of missing orders. Therefore to predict missing orders we can take the structure factor to be simply:

$$F(hkl) = \sum_j \exp[2\pi i (hx_j + ky_j + lz_j)] \quad (2.21)$$

Hence, given the two atomic positions, the structure factor associated with each Bragg reflection is:

$$F(hkl) = \sum_j \exp[2\pi i (hx_j + ky_j + lz_j)] = 1 + \exp(\pi i (h + k + l)) = 1 + (-1)^{h+k+l} \quad (2.22)$$

Therefore $F(hkl)$ is zero if $h + k + l$ is odd. This represents the systematic absence for this crystal structure.

As a second example, consider a crystal structure in the form of a simple cubic Bravais lattice with a two-point monoatomic basis at atomic positions $(1/2, 0, 1/2)$ and $(0, 1/2, 1/2)$. This is equivalent to a fcc Bravais lattice. The structure factor

in this case is given by:

$$F(hkl) = \sum_j \exp[2\pi i(hx_j + ky_j + lz_j)] = \exp[\pi i(h+l)] + \exp[\pi i(k+l)] = (-1)^{h+l} + (-1)^{k+l} \quad (2.23)$$

Hence the systematic absences occur if h, k, l are not all even or all odd.

Finally, we consider the case of the monoatomic cubic diamond crystal structure, in which elements such as carbon, silicon, germanium or tin crystallise. This is not a Bravais lattice and must therefore be described as a lattice with a basis. The underlying Bravais lattice is fcc, and the two atom basis lies at atomic positions $0, 0, 0$ and $1/4, 1/4, 1/4$ relative to the conventional eight-atom unit cell. The structure factor consequently is:

$$F(hkl) = \sum_j \exp[2\pi i(hx_j + ky_j + lz_j)] = 1 + \exp\left[i\frac{\pi}{2}(h+k+l)\right] \quad (2.24)$$

This time we take the square of the above expression for convenience:

$$F^2(hkl) = F(hkl) F^*(hkl) = 2 \left[1 + \cos\left(\frac{\pi}{2}(h+k+l)\right) \right] \quad (2.25)$$

The above structure factor also applies to the β -Sn high pressure phase considered in this Thesis, since the relative atomic positions of the two-atom basis there are the same. The selection rules that determine the occurrence of systematic absences are therefore that $h+k+l$ be twice an odd number. To this selection rule, we must also superimpose that of the underlying fcc Bravais lattice that the structure factor falls to zero when h, k, l are mixed even or odd.

2.3 Theory of elasticity

2.3.1 Introduction

The theory describing the mechanical behaviour of elastic bodies under deformation is vast and mathematically very sophisticated. Only the essential elements will be introduced within the scope of this brief introduction. For a complete, rigorous and modern description of the subject, the reader is referred to the monograph by Lur'ie , which builds upon the classical work by Love, Timoshenko and Solkonikoff. Alternatively, there exists many textbooks which offer a more introductory treatment ?????.

Elasticity is an elegant theory that deals with the determination of the stress, strain and displacement distribution in an elastic solid under the influence of external forces. Following the usual assumptions of linear, small-deformation theory, the formulation establishes a mathematical model that allows solutions to problems that have applications in many engineering and scientific fields. These include Civil engineering applications, which normally consist in the analysis of stresses and deflections in structures such as rods, beams, plates and shells; applications in geomechanics, which involve the stresses in materials such as soil, rock, concrete and asphalt; Mechanical engineering, which uses elasticity in numerous problems in the analysis and design of machine elements involving general stress analysis, contact stresses, thermal stress analysis, fracture mechanics and fatigue; Material engineering, which uses elasticity to determine the stress fields in crystalline solids, around dislocations , and in materials with microstructure; and finally there are applications in aeronautical and aerospace engineering , which include stress, fracture and fatigue analysis in aereostrutures. The subject also

provides the foundations for more advanced work in inelastic material behaviour, including plasticity and viscoelasticity, and in the analysis of finite, non-linear deformations.

The concept of elastic force-deformation relation was first proposed by Robert Hooke in 1678. However, the major formulation of the mathematical theory of elasticity was not developed until the 19th century: in 1821 Navier presented his investigations on the general equations of equilibrium; he was quickly followed by Cauchy, who studied the basic elasticity equations and developed the notation of stress at a point. A long list of prominent scientists and mathematicians continued development of the theory, including Bernoulli, Lord Kelvin, Poisson, Lamé, Green, Airy, Rayleigh and Love.

2.3.2 Deformations and Strains

As a result of applied external loadings, elastic solids will change shape and thence deform, and these deformations can be quantified by knowing the displacements of material points in the body. The continuum hypothesis establishes a displacement field at all points within the elastic solid. Using appropriate geometry, particular measures of deformation can be constructed leading to the development of the strain tensor. As appropriate for linear elasticity, the relation between displacement and strain will be developed in this section under the conditions of small deformations. For finite or large deformation theory, the undeformed and deformed configurations can be significantly different leading to the so-called Lagrangian and Eulerian descriptions, which however lie beyond the scope of this introduction.

Typically when elastic solids deform under the effect of some external load, the deformation varies from point to point and is thus said to be non-homogeneous. In particular, elements within the material undergo extensional and shearing deformation. An elastic solid is thus said to be deformed or strained when the relative displacements between points in the body are changed. This contrasts with rigid-body motion where the distance between the points remains the same. Let us define the displacement vectors of two pairs of points between the undeformed and deformed configurations of a generic solid as \vec{u}_o and \vec{u} respectively. Since the points in each of the deformed and undeformed configurations are neighbouring and are therefore separated by a small distance \vec{r} and \vec{r} respectively, we can use a Taylor expansion to express the components of \vec{u} as:

$$\begin{aligned} u &= u_o + \frac{\partial u}{\partial x} r_x + \frac{\partial u}{\partial y} r_y + \frac{\partial u}{\partial z} r_z \\ v &= v_o + \frac{\partial v}{\partial x} r_x + \frac{\partial v}{\partial y} r_y + \frac{\partial v}{\partial z} r_z \\ w &= w_o + \frac{\partial w}{\partial x} r_x + \frac{\partial w}{\partial y} r_y + \frac{\partial w}{\partial z} r_z \end{aligned} \quad (2.26)$$

The change in the relative position vector \vec{r} can be written as:

$$\Delta r = \vec{r} - \vec{r} = \vec{u} - \vec{u}_o \quad (2.27)$$

and using Eq. (2.26) gives:

$$\begin{aligned} \Delta r_x &= \frac{\partial u}{\partial x} r_x + \frac{\partial u}{\partial y} r_y + \frac{\partial u}{\partial z} r_z \\ \Delta r_y &= \frac{\partial v}{\partial x} r_x + \frac{\partial v}{\partial y} r_y + \frac{\partial v}{\partial z} r_z \\ \Delta r_z &= \frac{\partial w}{\partial x} r_x + \frac{\partial w}{\partial y} r_y + \frac{\partial w}{\partial z} r_z \end{aligned} \quad (2.28)$$

or in index notation:

$$\Delta r_i = u_{i,j} r_j \quad (2.29)$$

The tensor $u_{i,j}$ is called the displacement gradient tensor, and may be expressed as:

$$u_{i,j} = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{pmatrix} \quad (2.30)$$

This tensor can be decomposed into symmetric and antisymmetric parts as:

$$u_{i,j} = e_{ij} + \omega_{ij} \quad (2.31)$$

where

$$\begin{aligned} e_{ij} &= \frac{1}{2} (u_{i,j} + u_{j,i}) \\ \omega_{ij} &= \frac{1}{2} (u_{i,j} - u_{j,i}) \end{aligned} \quad (2.32)$$

The tensor e_{ij} is called the strain tensor, while ω_{ij} is referred to as the rotation tensor. Hence for small deformation theory the change in the relative position vector between neighbouring points can be expressed in terms of a sum of strain and rotation components. This is because we are considering a general displacement field, so that the results include both strain deformation and rigid-body motion. We now wish to establish a more geometrical interpretation of these results. Consider the common deformational behaviour of a rectangular element as shown in Fig. 2.10. The usual types of motion include rigid-body rotation

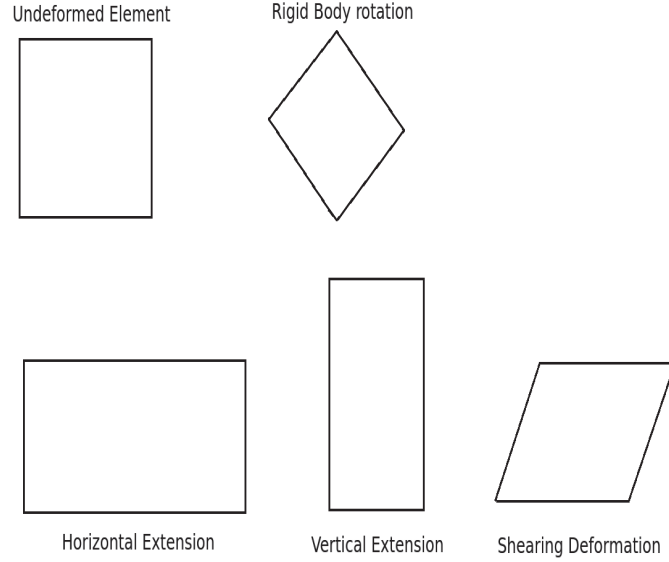


Figure 2.10: Possible types of deformation of a rectangular element.

and extensional and shearing deformation as illustrated. Rigid-body motion does not contribute to the strain field, and thus also does not affect the stresses. We therefore focus our study primarily on the extensional and shearing deformation.

Fig. 2.11 illustrates the two-dimensional deformation of a rectangular element with original dimensions dx by dy . After deformation, the element takes a rhombus form as shown in the dotted outline. The displacements of various corner reference points are indicated in the figure. Reference point A is taken at location (x, y) , and the displacement components of this point are thus $u(x, y)$ and $v(x, y)$, where u and v are the displacements fields along x and y respectively. The corresponding displacements of point B are $u(x + dx, y)$ and $v(x + dx, y)$, and the displacements of the other corner points are defined in an analogous manner. According to small deformation theory, $u(x + dx, y) \approx u(x, y) + (\partial u / \partial x) dx$, with similar expansions for all other terms. The normal or extensional strain compo-

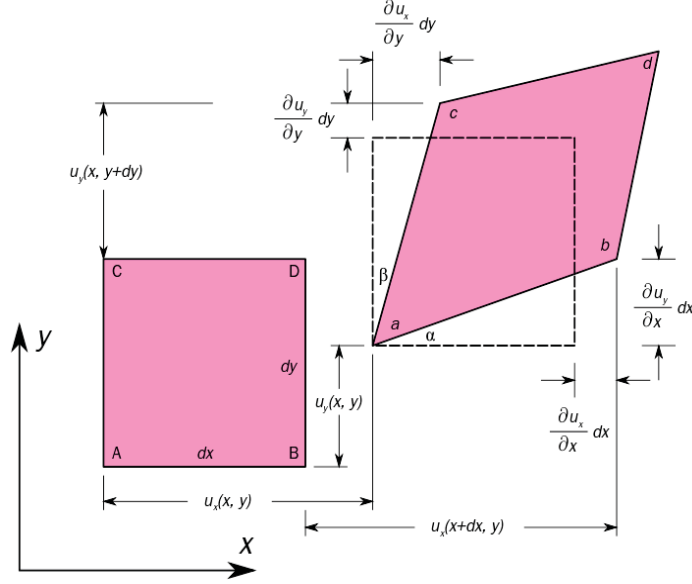


Figure 2.11: Two-dimensional geometric strain deformation.

ment in a direction n is defined as the change in length per unit length of fibres oriented in the n direction. Normal strain is positive if fibres increase in length and negative if the fibre is shortened. In Fig. 2.11, the normal strain in the x direction can thus be defined by:

$$\varepsilon_x = \frac{A'B' - AB}{AB} \quad (2.33)$$

From the geometry of Fig. 2.11:

$$A'B' = \sqrt{\left(dx + \frac{\partial u}{\partial x} dx\right)^2 + \left(\frac{\partial v}{\partial x} dx\right)^2} = \sqrt{1 + 2\frac{\partial u}{\partial x} + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2} dx \approx \left(1 + \frac{\partial u}{\partial x}\right) dx \quad (2.34)$$

Using these results and the fact that $AB = dx$, the normal strain in the x -direction reduces to:

$$\varepsilon_x = \frac{\partial u}{\partial x} \quad (2.35)$$

In a similar fashion, the normal strain in the y -direction becomes:

$$\varepsilon_y = \frac{\partial v}{\partial y} \quad (2.36)$$

A second type of strain is shearing deformation, which involves angle changes. Shear strain is defined as the change in angle between two originally orthogonal directions in the continuum material. Measured in radians, shear strain is positive if the right angle between the positive directions of the two axes decreases. Thus the sign of the shear strain depends on the coordinate system. In Fig. 2.11, the shear strain with respect to the x and y directions can be defined as:

$$\gamma_{xy} = \alpha + \beta \quad (2.37)$$

For small deformations, $\alpha \approx \tan\alpha$ and $\beta \approx \tan\beta$, and the shear strain can then be expressed as:

$$\gamma_{xy} = \frac{\frac{\partial v}{\partial x} dx}{dx + \frac{\partial u}{\partial x} dx} + \frac{\frac{\partial u}{\partial y} dy}{dy + \frac{\partial v}{\partial y} dy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \quad (2.38)$$

By simple interchange of x and y and u and v , it is apparent that $\gamma_{xy} = \gamma_{yx}$. By considering similar behaviours in the $y - z$ and $x - z$ planes, these results can be easily extended to the general three-dimensional case, giving the results:

$$\begin{aligned} \varepsilon_x &= \frac{\partial u}{\partial x}, & \varepsilon_y &= \frac{\partial v}{\partial y}, & \varepsilon_z &= \frac{\partial w}{\partial z} \\ \gamma_{xy} &= \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}, & \gamma_{yz} &= \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}, & \gamma_{zx} &= \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \end{aligned} \quad (2.39)$$

Thus, we define three normal and three shearing strain components leading to a total of six independent components that completely describe small deformation theory. This set of equations is normally referred to as the strain-displacement relations. However, tensorial elasticity theory prefers to use the formulation of the strain tensor e_{ij} given by Eq. (2.32). This represents only a minor change because the normal strains are identical and shearing strains differ by a factor of one-half. Therefore, using the strain tensor e_{ij} , the strain-displacement relations can be expressed in component form as:

$$\begin{aligned} e_x &= \frac{\partial u}{\partial x}, & e_y &= \frac{\partial v}{\partial y}, & e_z &= \frac{\partial w}{\partial z} \\ e_{xy} &= \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right), & e_{yz} &= \frac{1}{2} \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right), & e_{zx} &= \frac{1}{2} \left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \right) \end{aligned} \quad (2.40)$$

Using the more compact tensor notation, these relations are written as:

$$e_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i}) \quad (2.41)$$

while in direct vector/matrix notation the form reads:

$$\mathbf{e} = \frac{1}{2} \left[\nabla \mathbf{u} + (\nabla \mathbf{u})^T \right] \quad (2.42)$$

where \mathbf{e} is the strain matrix and $\nabla \mathbf{u}$ is the displacement gradient matrix. The strain is a symmetric second-order tensor ($e_{ij} = e_{ji}$) and is commonly written in matrix format:

$$\mathbf{e} = \begin{pmatrix} e_x & e_{xy} & e_{xz} \\ e_{xy} & e_y & e_{yz} \\ e_{xz} & e_{yz} & e_z \end{pmatrix} \quad (2.43)$$

It follows that because the strain is a symmetric second-order tensor, we can identify and determine its principal axes and values (the eigenvectors and eigenvalues). According to this theory, for any given strain tensor we can establish the principal value problem and solve the characteristic equation to explicitly determine the principal values and directions. The general characteristic equation for the strain tensor can be written as:

$$\det [e_{ij} - e\delta_{ij}] = -e^3 + \vartheta_1 e^2 - \vartheta_2 e + \vartheta_3 = 0 \quad (2.44)$$

where e is the principal strain and the fundamental invariants of the strain tensor can be expressed in terms of the three principal strains e_1, e_2, e_3 (the eigenvalues of the strain tensor) as:

$$\begin{aligned} \vartheta_1 &= e_1 + e_2 + e_3 \\ \vartheta_2 &= e_1 e_2 + e_2 e_3 + e_3 e_1 \\ \vartheta_3 &= e_1 e_2 e_3 \end{aligned} \quad (2.45)$$

The first invariant $\vartheta_1 = \vartheta$ is normally called the cubical dilatation, because it is related to the change in volume of material elements. The strain matrix in the principal coordinate system takes the special diagonal form:

$$e_{ij} = \begin{pmatrix} e_1 & 0 & 0 \\ 0 & e_2 & 0 \\ 0 & 0 & e_3 \end{pmatrix} \quad (2.46)$$

Notice that for this principal coordinate system, the deformation does not produce any shearing and thus is only extensional. Therefore, a rectangular element oriented along principal axes of strain will retain its orthogonal shape and undergo only extensional deformation of its sides.

In particular applications it is convenient to decompose the strain tensor into two parts called spherical and deviatoric strain tensors. The spherical strain is defined by

$$\widetilde{\mathbf{e}}_{ij} = \frac{1}{3}e_{kk}\delta_{ij} = \frac{1}{3}\vartheta\delta_{ij} \quad (2.47)$$

while the deviatoric strain is specified as:

$$\widehat{\mathbf{e}}_{ij} = e_{ij} - \frac{1}{3}e_{kk}\delta_{ij} \quad (2.48)$$

Note that the total strain is then simply the sum

$$e_{ij} = \widetilde{\mathbf{e}}_{ij} + \widehat{\mathbf{e}}_{ij} \quad (2.49)$$

The spherical strain represents only volumetric deformation and is an isotropic tensor, being the same in all coordinate systems. The deviatoric strain tensor then accounts for changes in shape of material elements. It can be shown that the principal directions of the deviatoric strain are the same as those of the strain

tensor.

2.3.3 Stress and equilibrium

When a structure is subjected to applied external loadings, internal forces are induced inside the body. Following the philosophy of continuum mechanics, these internal forces are distributed continuously within the solid. In order to study such forces, it is convenient to categorize them into two major groups, commonly referred to as body forces and surface forces.

Body forces are proportional to the body's mass and are reacted with an agent outside of the body. By using continuum mechanics principles, a body force density (force per unit volume) $F(x)$ can be defined such that the total resultant body force of an entire solid can be written as a volume integral over the body:

$$F_R = \iiint_V F(x) dV \quad (2.50)$$

Surface forces always act on a surface and result from physical contact with another body. Again, the resultant surface force over the entire surface S can be expressed as the integral of a surface force density function $T^n(x)$:

$$F_S = \iint_S T^n(x) dS \quad (2.51)$$

The surface force density is normally referred to as the traction vector. In order to quantify the nature of the internal distribution of forces within a continuum solid, consider a general body subject to arbitrary (concentrated and distributed) external loadings. To investigate the internal forces, a section is made through

the body. On this section consider a small area ΔA with unit normal vector \vec{n} . The resultant surface force acting on ΔA is defined by $\Delta \mathbf{F}$. The stress or traction vector is defined by:

$$\vec{T}^n(\vec{x}, \vec{n}) = \lim_{\Delta A \rightarrow 0} \frac{\Delta \vec{F}}{\Delta A} \quad (2.52)$$

Notice that the traction vector depends on both the spatial location and the unit normal vector to the surface under study. Thus, even though we may be investigating the same point, the traction vector still varies as a function of the orientation of the surface normal. Because the traction is defined as force per unit area, the total surface force is determined through integration according to Eq. (2.51). Consider now the special case in which ΔA coincides with each of the three coordinate planes with the unit normal vectors pointing along the positive coordinate axes. This concept is shown in Fig. 3.4, where the three coordinate surfaces for ΔA partition off a cube of material. For this case, the traction vector on each face can be written as

$$\begin{aligned} \vec{T}^n(\vec{x}, \vec{n} = \vec{e}_1) &= \sigma_x \vec{e}_1 + \tau_{xy} \vec{e}_2 + \tau_{xz} \vec{e}_3 \\ \vec{T}^n(\vec{x}, \vec{n} = \vec{e}_2) &= \tau_{yx} \vec{e}_1 + \sigma_y \vec{e}_2 + \tau_{yz} \vec{e}_3 \\ \vec{T}^n(\vec{x}, \vec{n} = \vec{e}_3) &= \tau_{zx} \vec{e}_1 + \tau_{zy} \vec{e}_2 + \sigma_z \vec{e}_3 \end{aligned} \quad (2.53)$$

where $\vec{e}_1, \vec{e}_2, \vec{e}_3$ are the unit vectors along each coordinate direction, and the nine quantities $\{\sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{yx}, \tau_{yz}, \tau_{zy}, \tau_{zx}, \tau_{xz}\}$ are the components of the traction vector on each of three coordinate planes as illustrated. These nine components are called the stress components, with $\sigma_x, \sigma_y, \sigma_z$ referred to as normal stresses and $\tau_{xy}, \tau_{yx}, \tau_{yz}, \tau_{zy}, \tau_{zx}, \tau_{xz}$ called the shearing stresses. The components

of stress σ_{ij} are commonly written in matrix format:

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{pmatrix} \quad (2.54)$$

and it can be formally shown that the stress is a second-order tensor. The positive directions of each stress component are illustrated in Fig. ?? . Regardless of the coordinate system, positive normal stress always acts in tension out of the face, and only one subscript is necessary because it always acts normal to the surface. The shear stress, however, requires two subscripts, the first representing the plane of action and the second designating the direction of the stress. Similar to shear strain, the sign of the shear stress depends on coordinate system orientation. For example, on a plane with a normal in the positive x direction, positive τ_{xy} acts in the positive y direction. Similar definitions follow for the other shear stress components.

Consider next the traction vector on an oblique plane with arbitrary orientation. The unit normal to the surface can be expressed by:

$$\vec{n} = n_x \vec{e}_1 + n_y \vec{e}_2 + n_z \vec{e}_3 \quad (2.55)$$

where n_x, n_y, n_z are the direction cosines of the unit vector \vec{n} relative to the given coordinate system. We now consider the equilibrium of the pyramidal element interior to the oblique and coordinate planes. Invoking the force balance between tractions on the oblique and coordinate faces gives:

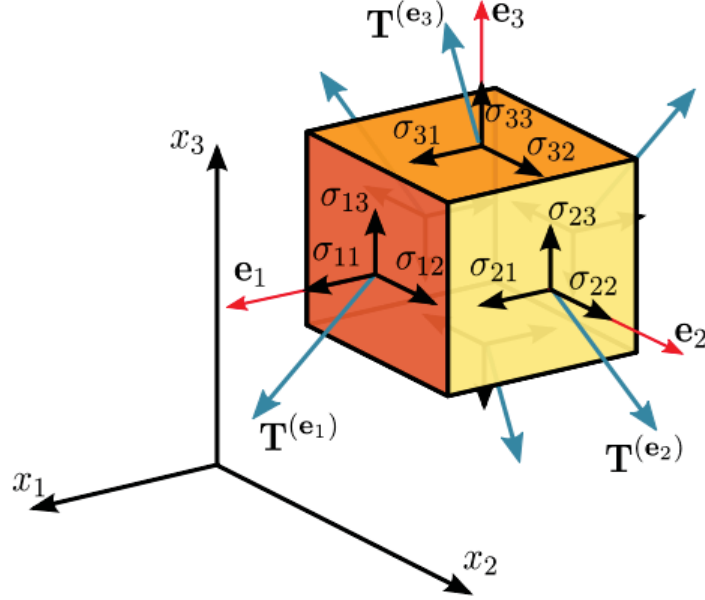


Figure 2.12: Components of the stress tensor and of the traction vectors.

$$\vec{T}^h = n_x \vec{T}^h(\vec{n} = \vec{e}_1) + n_y \vec{T}^h(\vec{n} = \vec{e}_2) + n_z \vec{T}^h(\vec{n} = \vec{e}_3) \quad (2.56)$$

and by using Eq. (2.53), this can be written as:

$$\vec{T}^h = (\sigma_x n_x + \tau_{yx} n_y + \tau_{zx} n_z) \vec{e}_1 + (\tau_{xy} n_x + \sigma_y n_y + \tau_{zy} n_z) \vec{e}_2 + (\tau_{xz} n_x + \tau_{yz} n_y + \sigma_z n_z) \vec{e}_3 \quad (2.57)$$

or in index notation

$$T_i^n = \sigma_{ji} n_j \quad (2.58)$$

Based on these previous definitions, the distinction between the traction vector and stress tensor should be carefully understood. Although each quantity has the same units of force per unit area, they are fundamentally different since the

traction is a vector while the stress is a second-order tensor (matrix). Components of traction can be defined on any surface, but particular stress components only exist on coordinate surfaces, as shown in Fig. 3.4 for the Cartesian case. Clearly, Eq. 2.58 establishes the relation between the two variables, thereby indicating that each traction component can be expressed as a linear combination of particular stress components.

It can be shown that the stress is a symmetric tensor. Using this fact, we can identify and determine principal axes and values for the stress (eigenvectors and eigenvalues of the stress tensor). For any given stress tensor we can establish the principal value problem and solve the characteristic equation to explicitly determine the principal values and directions. The general characteristic equation for the stress tensor becomes:

$$\det [\sigma_{ij} - \sigma \delta_{ij}] = -\sigma^3 + I_1 \sigma^2 - I_2 \sigma + I_3 = 0 \quad (2.59)$$

where σ are the principal stresses and the fundamental invariants of the stress tensor can be expressed in terms of the three principal stresses $\sigma_1, \sigma_2, \sigma_3$ as:

$$\begin{aligned} I_1 &= \sigma_1 + \sigma_2 + \sigma_3 \\ I_2 &= \sigma_1 \sigma_2 + \sigma_2 \sigma_3 + \sigma_3 \sigma_1 \\ I_3 &= \sigma_1 \sigma_2 \sigma_3 \end{aligned} \quad (2.60)$$

In the principal coordinate system, the stress matrix takes the special diagonal form:

$$\sigma_{ij} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \quad (2.61)$$

Notice that for the principal coordinate system, all shearing stresses vanish and thus the state includes only normal stresses. These issues should be compared to the equivalent comments made for the strain tensor at the end of the previous section.

As mentioned in our previous discussion on strain, it is often convenient to decompose the stress into two parts called the spherical and deviatoric stress tensors. Analogous to relations (2.47) and (2.48), the spherical stress is defined by:

$$\widetilde{\sigma}_{ij} = \frac{1}{3}\sigma_{kk}\delta_{ij} \quad (2.62)$$

while the deviatoric stress becomes

$$\widehat{\sigma}_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{kk}\delta_{ij} \quad (2.63)$$

Note that the total stress is then simply the sum

$$\sigma_{ij} = \widetilde{\sigma}_{ij} + \widehat{\sigma}_{ij} \quad (2.64)$$

The spherical stress is an isotropic tensor, being the same in all coordinate systems. It can also be shown that the principal directions of the deviatoric stress are the same as those of the stress tensor itself. We next briefly explore a

couple of particular stress components or combinations that have been defined in the literature and are commonly used in formulating failure theories related to inelastic deformation. It has been found that ductile materials normally exhibit inelastic yielding failures that can be characterized by these particular stresses. The octahedral normal and shear stresses are defined in terms of the principal values of stress as:

$$\begin{aligned}\sigma_{oct} &= \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3) = \frac{1}{3}\sigma_{kk} = \frac{1}{3}I_1 \\ \tau_{oct} &= \frac{1}{3}[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]^{1/2} = \frac{1}{3}(2I_1^2 - 6I_2)^{1/2}\end{aligned}\quad (2.65)$$

Another specially defined stress is known as the effective or von Mises stress and is given by the expression:

$$\sigma_{vonMises} = \frac{1}{\sqrt{2}}[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]^{1/2} \quad (2.66)$$

If at some point in the structure, the von Mises stress equals the yield stress, then the material is considered to be at the failure condition

2.3.4 Linear elastic solids

We now wish to complete our general formulation of the theory of elasticity by specializing to a particular material model that provides reasonable characterization of materials under small deformations. The model is that of a linear elastic material. This section presents the basics of the linear elastic model that applies to isotropic materials only. Related theory for anisotropic media is developed in the next section.

Relations that characterize the physical properties of materials are called constitutive equations. Our interest here is limited to a special class of solid materials with loadings resulting from mechanical or thermal effects. The mechanical behaviour of solids is normally defined by constitutive stress-strain relations. Commonly, these relations express the stress as a function of the strain, strain rate, strain history, temperature, and material properties. We choose a rather simple material model called the elastic solid that does not include rate or history effects. Furthermore, we restrict the constitutive stress-strain law to be linear, thus leading to a linear elastic solid. Although these assumptions greatly simplify the model, linear elasticity predictions have shown good agreement with experimental data and have provided useful methods to conduct stress analysis. Many structural materials including metals, plastics, ceramics, wood, rock, concrete, and so forth exhibit linear elastic behaviour under small deformations. Experimental testing is commonly employed in order to characterize the mechanical behaviour of real materials. It is observed experimentally that materials exhibit an initial stress-strain response for small deformation that is approximately linear. This is followed by a change to nonlinear behaviour that can lead to large deformation, finally ending with sample failure. For each material the initial linear response ends at a point normally referred to as the proportional limit. Another observation in this initial region is that if the loading is removed, the sample returns to its original shape and the strain disappears. This characteristic is the primary descriptor of elastic behavior. However, at some point on the stress-strain curve unloading does not bring the sample back to zero strain and some permanent plastic deformation results. The point at which this nonelastic behavior begins is called the elastic limit. Although some materials exhibit

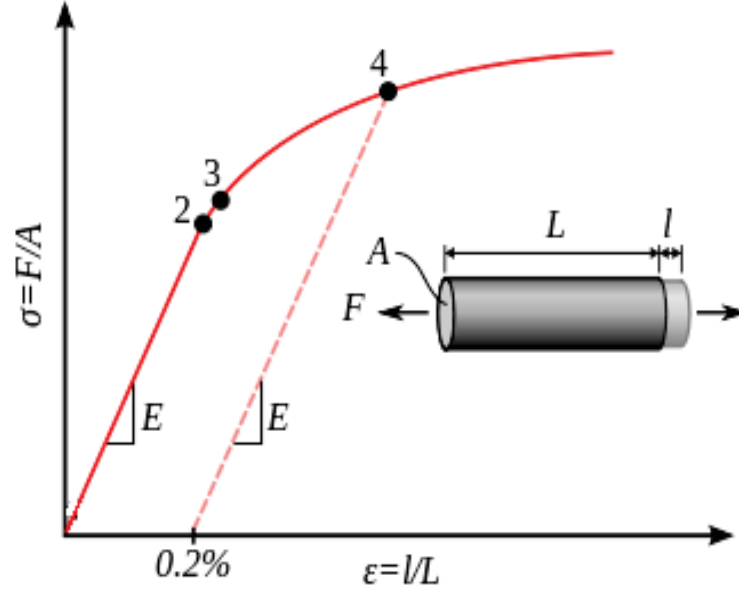


Figure 2.13: Stress-strain curve showing typical yield behaviour. Stress (σ) is shown as a function of strain (ϵ). The positions of the Proportionality limit (2), Elastic limit (3) and yield point (4) are indicated.

different elastic and proportional limits, many times these values are taken to be approximately the same. Another demarcation on the stress-strain curve is referred to as the yield point, defined by the location where large plastic deformation begins. The appearance of a typical stress-strain curve summarising the positions of the various limiting behaviours is shown in Fig. 2.13

It is therefore concluded that a large variety of real materials exhibit linear elastic behavior under small deformations. This would lead to a linear constitutive model for the one-dimensional axial loading case given by the relation $\sigma = E\epsilon$, where E is the slope of the uniaxial stress-strain curve. To construct a general three-dimensional constitutive law for linear elastic materials, we assume that each stress component is linearly related to each strain component. This relation can be cast into a matrix format as

$$\begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \\ \tau_{xy} \\ \tau_{yz} \\ \tau_{zx} \end{pmatrix} = \begin{pmatrix} C_{11} & \cdots & C_{16} \\ \vdots & \ddots & \vdots \\ C_{61} & \cdots & C_{66} \end{pmatrix} \begin{pmatrix} e_x \\ e_y \\ e_z \\ 2e_{xy} \\ 2e_{yz} \\ 2e_{zx} \end{pmatrix} \quad (2.67)$$

where the coefficients C_{ij} are material parameters and the factors of 2 arise because of the symmetry of the strain. Note that this relation could also be expressed by writing the strains as a linear function of the stress components. Relations (2.67) can also be expressed in standard tensor notation by writing:

$$\sigma_{ij} = C_{ijkl}e_{kl} \quad (2.68)$$

where C_{ijkl} is a fourth-order elasticity tensor whose components include all the material parameters necessary to characterize the material. Based on the symmetry of the stress and strain tensors, the elasticity tensor must have the following properties:

$$\begin{aligned} C_{ijkl} &= C_{jikl} \\ C_{ijkl} &= C_{ijlk} \end{aligned} \quad (2.69)$$

In general, the fourth-order tensor C_{ijkl} has 81 components. However, relations (2.69) reduce the number of independent components to 36, and this provides the required match with form (2.67). It can also be shown that the components of C_{ijkl} satisfy the relation $C_{ijkl} = C_{klij}$ or equivalently $C_{ij} = C_{ji}$, which

provides further reduction to 21 independent elastic components. The components of C_{ijkl} or equivalently C_{ij} are called elastic moduli and have units of stress (force/area).

If the material is homogeneous, the elastic behaviour does not vary spatially, and thus all elastic moduli are constant. For this case, the elasticity formulation is straightforward, leading to the development of many analytical solutions to problems of engineering interest. A homogeneous assumption is an appropriate model for most structural applications, and thus we primarily choose this particular case for subsequent analysis. Similar to homogeneity, another fundamental material property is isotropy. This property has to do with differences in material moduli with respect to orientation. For example, many materials including crystalline minerals, wood, and fiber-reinforced composites have different elastic moduli in different directions. Materials such as these are said to be anisotropic. Note that for most real anisotropic materials there exist particular directions where the properties are the same. These directions indicate material symmetries. However, for many engineering materials (most structural metals and many plastics), the orientation of crystalline and grain microstructure is distributed randomly so that macroscopic elastic properties are found to be essentially the same in all directions. Such materials with complete symmetry are called isotropic. As expected, an anisotropic model complicates the formulation and solution of problems. We therefore postpone development of such solutions until the next section and continue our current development under the assumption of isotropic material behaviour.

The tensorial form (2.68) provides a convenient way to establish the desired isotropic stress-strain relations. If we assume isotropic behavior, the elasticity

tensor must be the same under all rotations of the coordinate system. Using the basic transformation properties for fourth-rank tensors, it can be shown that the most general form that satisfies this isotropy condition is given by:

$$C_{ijkl} = \alpha \delta_{ij} \delta_{kl} + \beta \delta_{ik} \delta_{jl} + \gamma \delta_{il} \delta_{jk} \quad (2.70)$$

where α , β , and γ are arbitrary constants. Using the general form (2.70) in the stress-strain relation (2.68) gives:

$$\sigma_{ij} = \lambda e_{kk} \delta_{ij} + 2\mu e_{ij} \quad (2.71)$$

where we have relabelled particular constants using λ and ν . The elastic constant λ is called Lamé's constant, and ν (or alternatively G) is referred to as the shear modulus or modulus of rigidity. Eq. (2.71) can be written out in individual scalar equations as:

$$\begin{aligned} \sigma_x &= \lambda (e_x + e_y + e_z) + 2\mu e_x \\ \sigma_y &= \lambda (e_x + e_y + e_z) + 2\mu e_y \\ \sigma_z &= \lambda (e_x + e_y + e_z) + 2\mu e_z \\ \tau_{xy} &= 2\mu e_{xy} \\ \tau_{yz} &= 2\mu e_{yz} \\ \tau_{zx} &= 2\mu e_{zx} \end{aligned} \quad (2.72)$$

Relations (2.71) or 2.72 are called the generalized Hooke's law for linear isotropic elastic solids. It should be noted that only two independent elastic constants are needed to describe the behavior of isotropic materials. As shown in the next

section, additional numbers of elastic moduli are needed in the corresponding relations for anisotropic materials. The Stress-strain relations of Eq. (2.71) or 2.72 may be inverted to express the strain in terms of the stress. In order to do this it is convenient to use the index notation form (2.71) and set the two free indices the same (contraction process) to get:

$$\sigma_{kk} = (3\lambda + 2\mu) e_{kk} \quad (2.73)$$

This relation can be solved for e_{kk} and substituted back into Eq. (2.71) to get:

$$e_{ij} = \frac{1}{2\mu} \left(\sigma_{ij} - \frac{\lambda}{3\lambda + 2\mu} \sigma_{kk} \delta_{ij} \right) \quad (2.74)$$

which is more commonly written as:

$$e_{ij} = \frac{1 + \nu}{E} \sigma_{ij} - \frac{\nu}{E} \sigma_{kk} \delta_{ij} \quad (2.75)$$

where $E = \mu(3\lambda + 2\mu)/(\lambda + \mu)$ and is called the modulus of elasticity or Youngs modulus, and $\nu = \lambda/[2(\lambda + \mu)]$ is referred to as Poissons ratio. The index notation relation (2.75) may be written out in component (scalar) form, giving the six equations:

$$\begin{aligned}
e_x &= \frac{1}{E} [\sigma_x - \nu (\sigma_y + \sigma_z)] \\
e_y &= \frac{1}{E} [\sigma_y - \nu (\sigma_z + \sigma_x)] \\
e_z &= \frac{1}{E} [\sigma_z - \nu (\sigma_x + \sigma_y)] \\
e_{xy} &= \frac{1 + \nu}{E} \tau_{xy} = \frac{1}{2\mu} \tau_{xy} \\
e_{yz} &= \frac{1 + \nu}{E} \tau_{yz} = \frac{1}{2\mu} \tau_{yz} \\
e_{zx} &= \frac{1 + \nu}{E} \tau_{zx} = \frac{1}{2\mu} \tau_{zx}
\end{aligned} \tag{2.76}$$

Constitutive form (2.75) or (2.76) again illustrates that only two elastic constants are needed to formulate Hookes law for isotropic materials. By using any of the isotropic forms of Hookes law, it can be shown that the principal axes of stress coincide with the principal axes of strain. This result also holds for some but not all anisotropic materials.

For the isotropic case, the previously defined elastic moduli have simple physical meaning. These can be determined through investigation of particular states of stress commonly realized in laboratory materials testing. Consider first the simple tension test with a sample subjected to tension in the x direction. The state of stress is closely represented by the one-dimensional field:

$$\sigma_{ij} = \begin{pmatrix} \sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.77}$$

Using this in relation (2.75) gives a corresponding strain field:

$$e_{ij} = \begin{pmatrix} \frac{\sigma}{E} & 0 & 0 \\ 0 & -\frac{\nu}{E}\sigma & 0 \\ 0 & 0 & -\frac{\nu}{E}\sigma \end{pmatrix} \quad (2.78)$$

Therefore, $E = \sigma/e_x$ and is simply the slope of the stress-strain curve, while $\nu = -e_y/e_x = -e_z/e_x$ is the ratio of the transverse strain to the axial strain.

Secondly, we consider a thin-walled cylinder subjected to torsional loading, in which the state of stress on the surface of the cylindrical sample is given by:

$$\sigma_{ij} = \begin{pmatrix} 0 & \tau & 0 \\ \tau & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.79)$$

Again, by using Hookes law, the corresponding strain field becomes:

$$e_{ij} = \begin{pmatrix} 0 & \tau/2\mu & 0 \\ \tau/2\mu & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.80)$$

and thus the shear modulus is given by $\mu = \tau/2e_{xy} = \tau/\gamma_{xy}$, and this modulus is simply the slope of the shear stress-shear strain curve.

The final example is associated with the hydrostatic compression (or tension) loading of a cubical specimen. The state of stress for this case is given by:

$$\sigma_{ij} = \begin{pmatrix} -p & 0 & 0 \\ 0 & -p & 0 \\ 0 & 0 & -p \end{pmatrix} = -p\delta_{ij} \quad (2.81)$$

This is an isotropic state of stress and the strains follow from Hookes law:

$$e_{ij} = \begin{pmatrix} -\frac{1-2\nu}{E}p & 0 & 0 \\ 0 & -\frac{1-2\nu}{E}p & 0 \\ 0 & 0 & -\frac{1-2\nu}{E}p \end{pmatrix} \quad (2.82)$$

The dilatation that represents the change in material volume is thus given by $\vartheta = e_{kk} = -3(1 - 2\nu)p/E$, which can be written as:

$$p = -k\vartheta \quad (2.83)$$

where $k = E/[3(1 - 2\nu)]$ (alternatively labelled as K) is called the bulk modulus of elasticity. This additional elastic constant represents the ratio of pressure to the dilatation, which could be referred to as the volumetric stiffness of the material. Notice that as Poissons ratio approaches 0.5, the bulk modulus becomes unbounded and the material does not undergo any volumetric deformation and hence is incompressible. This represents the upper limit value of ν , the lower limit being given by -1 . E , μ , k and λ are all subject to the same bounds that they must be greater than zero.

Our discussion of elastic moduli for isotropic materials has led to the definition of five constants λ, μ, E, ν , and k . However, only two of these are needed to characterize the material. Although we have developed a few relationships between various moduli, many other such relations can also be found. In fact, it can be shown that all five elastic constants are interrelated, and if any two are given, the remaining three can be determined by using simple formulae. Results of these relations are conveniently summarized in Fig. 2.14.

	(K, E)	(K, λ)	(K, G)	(K, ν)	(E, G)	(E, ν)	(λ, G)	(λ, ν)	(G, ν)
$K =$	K	K	K	K	$\frac{EG}{3(3G-E)}$	$\frac{E}{3(1-2\nu)}$	$\lambda + \frac{2G}{3}$	$\frac{\lambda(1+\nu)}{3\nu}$	$\frac{2G(1+\nu)}{3(1-2\nu)}$
$E =$	E	$\frac{9K(K-\lambda)}{3K-\lambda}$	$\frac{9KG}{3K+G}$	$3K(1-2\nu)$	E	E	$\frac{G(3\lambda+2G)}{\lambda+G}$	$\frac{\lambda(1+\nu)(1-2\nu)}{\nu}$	$2G(1+\nu)$
$\lambda =$	$\frac{3K(3K-E)}{9K-E}$	λ	$K - \frac{2G}{3}$	$\frac{3K\nu}{1+\nu}$	$\frac{G(E-2G)}{3G-E}$	$\frac{E\nu}{(1+\nu)(1-2\nu)}$	λ	λ	$\frac{2G\nu}{1-2\nu}$
$G =$	$\frac{3KE}{9K-E}$	$\frac{3(K-\lambda)}{2}$	G	$\frac{3K(1-2\nu)}{2(1+\nu)}$	G	$\frac{E}{2(1+\nu)}$	G	$\frac{\lambda(1-2\nu)}{2\nu}$	G
$\nu =$	$\frac{3K-E}{6K}$	$\frac{\lambda}{3K-\lambda}$	$\frac{3K-2G}{2(3K+G)}$	ν	$\frac{E}{2G} - 1$	ν	$\frac{\lambda}{2(\lambda+G)}$	ν	ν

Figure 2.14: Relations between the five elastic constants $\lambda, \mu = G, E, \nu$, and $k = K$ in a isotropic linear elastic material. This class of materials has its elastic properties uniquely determined by any two moduli among these, thus given any two, any other of the elastic moduli can be calculated according to these formulae.

2.3.5 Anisotropic elasticity

It has long been recognized that deformation behaviour of many materials depends upon orientation, that is, the stress-strain response of a sample taken from the material in one direction will be different if the sample were taken in a different direction. The term anisotropic is generally used to describe such behaviours. Our previous development of the linear elastic stress-strain relations in the previous section began with the general case of inhomogeneous and anisotropic behaviour. However, this generality was quickly eliminated, and only the homogeneous isotropic case was subsequently developed in detail. We now wish to go back and further investigate the anisotropic homogeneous case.

The directional-dependent behaviours found in anisotropic solids normally

result from particular microstructural features within the material. Our previous isotropic model neglected these effects, thus resulting in a material that behaved the same in all directions. Micro features commonly arise in natural and synthetic materials in such a way as to produce a stress-strain response with particular symmetries. This concept is based on the Neumann principle that symmetry in material microgeometry corresponds to identical symmetry in the constitutive response. These symmetries generally lead to a reduction in the complexity of the stress-strain constitutive relation.

From the previous section, the general form of Hookes law was given by:

$$\sigma_{ij} = C_{ijkl}e_{kl} \quad (2.84)$$

The fourth-order elasticity tensor C_{ijkl} contains all of the elastic stiffness moduli, and we have previously established the following symmetry properties:

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij} \quad (2.85)$$

Relations (2.85) reduce the original 81 independent elastic constants within C_{ijkl} to a set of 21 elastic moduli for the general case. We shall assume that the material is homogeneous and thus the moduli are independent of spatial position. On occasion we may wish to invert Eq. (2.84) and write strain in terms of stress:

$$e_{ij} = S_{ijkl}\sigma_{kl} \quad (2.86)$$

where S_{ijkl} is the elastic compliance tensor, which has identical symmetry properties as those in relations (2.85). Because of the various preexisting symme-

tries, stress-strain relations (2.84) and (2.86) contain many superfluous terms and equations. To avoid these, a convenient contracted notation has been developed, sometimes referred to as Voigt matrix notation:

$$\begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \\ \tau_{yz} \\ \tau_{zx} \\ \tau_{xy} \end{pmatrix} = \begin{pmatrix} C_{11} & \cdots & C_{16} \\ \vdots & \ddots & \vdots \\ C_{61} & \cdots & C_{66} \end{pmatrix} \begin{pmatrix} e_x \\ e_y \\ e_z \\ 2e_{yz} \\ 2e_{zx} \\ 2e_{xy} \end{pmatrix} \quad (2.87)$$

or in compact notation

$$\sigma_i = C_{ij}e_j \quad (2.88)$$

Note that the 6×6 \mathbf{C} matrix is symmetric; that is, $C_{ij} = C_{ji}$, and thus only 21 independent elastic constants exist. A similar scheme can be established for Eq. (2.86), and a compliance matrix S_{ij} can be defined by:

$$e_i = S_{ij}\sigma_j \quad (2.89)$$

We thus determined that for the general anisotropic case (sometimes referred to as triclinic material), 21 independent elastic constants are needed to characterize the material response. However, most real materials have some types of symmetry, which further reduces the required number of independent elastic moduli. Orientations for which an anisotropic material has the same stress-strain

response can be determined by coordinate transformation (rotation) theory. In order to determine various material symmetries, it is more convenient to work in the noncontracted form. Thus, applying this theory, Hookes law (2.84) can be expressed in a new coordinate system as:

$$\sigma'_{ij} = C'_{ijkl} e'_{kl} \quad (2.90)$$

Now because the stress and strain must transform as second-order tensors:

$$\begin{aligned} \sigma'_{ij} &= Q_{ik} Q_{jl} \sigma_{kl}, & \sigma_{ij} &= Q_{ki} Q_{lj} \sigma'_{kl} \\ e'_{ij} &= Q_{ik} Q_{jl} e_{kl}, & e_{ij} &= Q_{ki} Q_{lj} e'_{kl} \end{aligned} \quad (2.91)$$

Combining equations (2.90) and (2.91) and using the orthogonality conditions between the transformation tensors \mathbf{Q} yields the transformation law for the elasticity tensor:

$$C'_{ijkl} = Q_{im} Q_{jn} Q_{kp} Q_{lq} C_{mnpq} \quad (2.92)$$

If under a specific transformation \mathbf{Q} the material response is to be the same, relation (2.92) reduces to:

$$C_{ijkl} = Q_{im} Q_{jn} Q_{kp} Q_{lq} C_{mnpq} \quad (2.93)$$

This material symmetry relation will provide a system of equations that allows reduction in the number of independent elastic moduli. We now consider some specific cases of practical interest. We first investigate the case of a material with a plane of symmetry. Such a medium is commonly referred to as a monoclinic

material. We consider the case of symmetry with respect to the x, y -plane. For this particular symmetry, the required transformation is simply a mirror reflection about the x, y -plane and is given by:

$$Q_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (2.94)$$

Using this specific transformation in relation (2.93) gives $C_{ijkl} = -C_{ijkl}$ if the index 3 appears an odd number of times, and thus these particular moduli would have to vanish. Thus, the elasticity matrix takes the form:

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & C_{16} \\ . & C_{22} & C_{23} & 0 & 0 & C_{26} \\ . & . & C_{33} & 0 & 0 & C_{36} \\ . & . & . & C_{44} & C_{45} & 0 \\ . & . & . & . & C_{55} & 0 \\ . & . & . & . & . & C_{66} \end{pmatrix} \quad (2.95)$$

It is therefore observed that 13 independent elastic moduli are needed to characterize monoclinic materials.

A material with three mutually perpendicular planes of symmetry is called orthotropic. Common examples of such materials include wood and fiber-reinforced composites. To investigate the material symmetries for this case, it is convenient to let the symmetry planes correspond to $x - y - z$ coordinate planes. The symmetry relations can be determined by using 180° rotations about each of the coordinate axes. Thus, the elasticity matrix for the orthotropic case reduces to

having only nine independent stiffnesses given by:

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ . & C_{22} & C_{23} & 0 & 0 & 0 \\ . & . & C_{33} & 0 & 0 & 0 \\ . & . & . & C_{44} & 0 & 0 \\ . & . & . & . & C_{55} & 0 \\ . & . & . & . & . & C_{66} \end{pmatrix} \quad (2.96)$$

For orthotropic materials, the compliance matrix has similar form but is commonly written using notation related to isotropic theory:

$$S_{ij} = \begin{pmatrix} \frac{1}{E_1} & -\frac{\nu_{21}}{E_2} & -\frac{\nu_{31}}{E_3} & 0 & 0 & 0 \\ -\frac{\nu_{12}}{E_1} & \frac{1}{E_2} & -\frac{\nu_{32}}{E_3} & 0 & 0 & 0 \\ -\frac{\nu_{13}}{E_1} & -\frac{\nu_{23}}{E_2} & \frac{1}{E_3} & 0 & 0 & 0 \\ . & . & . & \frac{1}{\mu_{23}} & 0 & 0 \\ . & . & . & . & \frac{1}{\mu_{31}} & 0 \\ . & . & . & . & . & \frac{1}{\mu_{12}} \end{pmatrix} \quad (2.97)$$

where E_i are Youngs moduli in the three directions of material symmetry, ν_{ij} are the Poissons ratios defined by $-e_j/e_i$ for a stress in the i direction, and μ_{ij} are the shear moduli in the i, j -planes. Symmetry of this matrix requires that $\nu_{ij}/E_i = \nu_{ji}/E_j$.

Another common form of material symmetry is with respect to rotations about an axis. This concept can be specified by stating that a material possess an axis of symmetry of order n when the elastic moduli remain unchanged for rotations of $2\pi/n$ radians about the axis. The transformation for arbitrary rotations θ about

the z -axis is given by:

$$Q_{ij} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.98)$$

Using this transformation and invoking symmetry for arbitrary rotations corresponds to the case of $n \rightarrow \infty$, and such materials are called transversely isotropic. The elasticity stiffness matrix for this case reduces to:

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ . & C_{11} & C_{13} & 0 & 0 & 0 \\ . & . & C_{33} & 0 & 0 & 0 \\ . & . & . & C_{44} & 0 & 0 \\ . & . & . & . & C_{44} & 0 \\ . & . & . & . & . & (C_{11} - C_{12})/2 \end{pmatrix} \quad (2.99)$$

Thus, for transversely isotropic materials, only five independent elastic constants exist.

For the case of complete symmetry, the material is referred to as isotropic, and the fourth-order elasticity tensor has been previously given by:

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \quad (2.100)$$

This form can be determined by invoking symmetry with respect to two orthogonal axes, which implies symmetry about the remaining axis. In contracted matrix form, this result would be expressed as:

$$C_{ij} = \begin{pmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ . & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ . & . & \lambda + 2\mu & 0 & 0 & 0 \\ . & . & . & \mu & 0 & 0 \\ . & . & . & . & \mu & 0 \\ . & . & . & . & . & \mu \end{pmatrix} \quad (2.101)$$

Thus, as shown previously, only two independent elastic constants exist for isotropic materials. For each case presented, a similar compliance elasticity matrix could be developed.

2.4 The physics of shock-waves

2.4.1 Introduction

Planar shock waves in solid crystalline materials are defined as propagating disturbances capable of giving rise to a discontinuity in the material's pressure, temperature (or internal energy) and density, thus resulting in an intense uniaxial dynamic compression of the internal crystal structure of the material along the shock propagation direction. A "Shock Wave" is then the term which is used to denote the interface of this discontinuity. The circumstances that surround the propagation of shock waves in condensed matter are made unique by the extreme physical conditions that can be attained, together with their exceptionally high rates of variation. These are typically characterized by pressures ranging from ~ 100 MPa up to several hundreds of GPa (and up to ten TPa in some unique experiments), and temperatures up to tens of thousands of degrees Kelvin, cou-

pled with durations of the load application varying from 10^{-9} to 10^{-5} s. Only the surface of the vast and active area of research that is the study of shock-induced compression in solid materials will be scratched within the scope of this short introduction. For a more comprehensive overview, the reader is referred to the numerous introductory review textbooks on the subject already available in the literature, such as those provided by Zel'dovich and Raizer [?], Isbell [,] Bushman ^{and} Meyers [?]. In addition, Springer offers a comprehensive collection of textbooks on shock-compression of condensed matter as part of its "Shock Wave and High Pressure Phenomena" series ^{??????}. Further comprehensive coverage of the field is given in the "Handbook of shock-waves" series of three volumes. Finally, a more concise introduction to the subject is offered in the review paper by Duvall [.]

2.4.2 Rankine-Hugoniot jump conditions

The concept of shock-wave propagation and the equations of conservation of mass, momentum and energy can be very easily understood and derived by means of a simplified conceptual framework involving a cylinder of unit cross-sectional area onto which a piston penetrates. This is shown in Fig. 2.15. Initially the piston is at rest. It then is pushed into the compressible material, initially, at a pressure P_0 and having density ρ_0 at a velocity U_p . After a time t_1 , the highly compressed region ahead of the piston has moved forward by a distance equal to $U_S t_1$, where U_S is the velocity of propagation of the disturbance ahead of the piston. During this time interval, the piston has moved by a distance equal to $U_p t_1$. The compressed region has a pressure P and a density ρ , and the particles

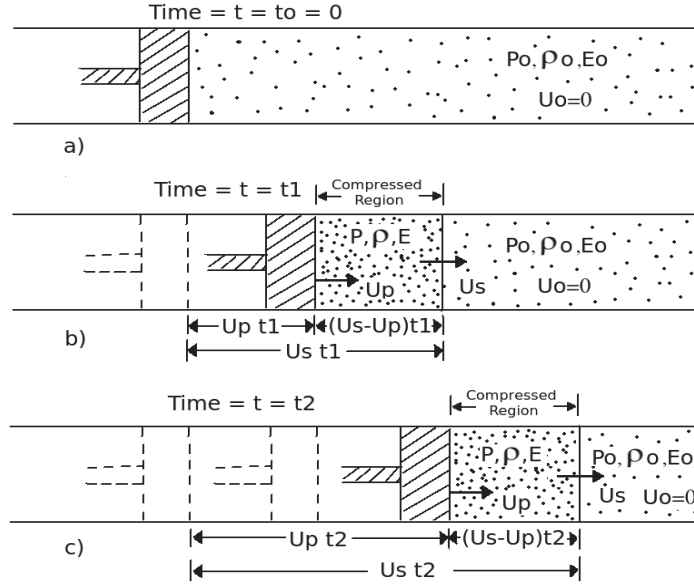


Figure 2.15: Successive positions of an idealised piston moving into a cylinder with a compressible fluid inside.

within it move at a velocity U_p equal to that of the piston. A shock front can be visualised as a plane discontinuity separating moving from stationary fluid in a cylinder with a moving piston. The velocity U_S of the shock-wave is greater than the piston velocity U_p , such that when the piston has traveled a distance $U_p t_1$ in time t_1 , the shock-wave has travelled a distance $U_S t_1$.

We can now look at the conservation of mass, momentum and energy across the shock-front. Since the internal cross-sectional area of the cylinder is unity, one has, for the compression length $(U_S - U_p) t_1$:

$$U_S t_1 \rho_0 = \rho (U_S - U_p) t_1 \quad (2.102)$$

The term $U_S t_1 \rho_0$ is the mass of the initial uncompressed material. This equation applies to any time t , and therefore we have for conservation of mass:

$$U_S \rho_0 = \rho (U_S - U_p) \quad (2.103)$$

The momentum is defined as the product of mass and velocity. The change in momentum of a system is equal to the impulse given to the system. Initially, the momentum is zero. The momentum at t_1 is the product of the mass $(U_S - U_p) \rho t$ and the velocity of the material particles in the compressed volume, U_p . Thus we have for conservation of momentum:

$$\rho (U_S - U_p) U_p t - 0 = (P - P_0) t \quad (2.104)$$

or

$$\rho (U_S - U_p) U_p = (P - P_0) \quad (2.105)$$

The equation of conservation of energy simply states that the work by the external forces is equal to the change in internal energy plus change of kinetic energy. In the compression region, at time t , the change in internal energy is:

$$E_1 [\rho (U_S - U_p) t] - E_0 [\rho_0 U_S t] = (E_1 - E_0) \rho_0 U_S t \quad (2.106)$$

where E is the internal energy density. The change in kinetic energy is given by $1/2mv^2$:

$$\frac{1}{2} \rho (U_S - U_p) U_p^2 t - 0 = \frac{1}{2} \rho_0 U_S U_p^2 t \quad (2.107)$$

For a stationary shock-wave, the change in kinetic energy is equal to the

change in internal energy:

$$E_1 - E_0 = \frac{1}{2}U_p^2 \quad (2.108)$$

Equations (2.103), (2.117) and (2.108) are, in essence, the Rankine-Hugoniot relationships for a material in which a pressure discontinuity propagates. These equations apply to a piston moving into a compressible medium (gas) and can be extended to a shock-wave propagating into gas, liquid, or solid or to a detonation wave. This treatment can be regarded as an analogue to the treatment for developing Rankine-Hugoniot conservation equations for shock-waves, which will be illustrated next.

Shock-waves are characterised by a steep front in the form of a discontinuous surface with no apparent thickness and require in a first instance a state of uniaxial strain along the shock-propagation direction (no lateral flow of materials) which allows the build-up of the shear component of stress to high levels. The requirement for the formation of shock-waves can be understood by considering the case of an ideal gas being compressed isentropically by a high-pressure disturbance travelling through it. The corresponding equation of state for the process is given by:

$$PV^\gamma = K = \text{constant} \quad (2.109)$$

By differentiating the above we obtain:

$$\gamma PV^{\gamma-1}dV + V^\gamma dP = 0 \rightarrow \frac{dP}{dV} = -\gamma \frac{P}{V} \quad (2.110)$$

Furthermore, it can be shown that the velocity of the disturbance in the gas,

assuming a one-dimensional configuration, is given by:

$$u = \left(\frac{d\sigma/d\varepsilon}{\rho} \right)^{1/2} = \left(\frac{dP/dV}{\rho} \right)^{1/2} = |K|^{1/2} \quad (2.111)$$

where K is the bulk modulus of the gas. Hence the velocity of the disturbance is proportional to $(dP/dV)^{1/2}$. Since the value of P/V increases with pressure, it can be concluded from Eq. (2.110) that high-pressure amplitude isentropic disturbances travel faster than low-amplitude ones in gases. Hence a disturbance front will "steepen up" as it travels through the material because the higher amplitude regions of the front travel faster than the low amplitude regions. This is the fundamental requirement for the establishment of a shock-wave, that the velocity of the pulse, U , increases with pressure, that is:

$$\left(\frac{\partial^2 P}{\partial U^2} \right) > 0 \quad (2.112)$$

as P and U increase. Fig. 2.16 demonstrates the steepening of a ramp as it travels into a sample. The higher pressure components advance faster and catch up to the low pressure leading edge to form an increasingly steep front with time, which eventually results in the formation of a shock-wave. The opposite case on the other hand, where the velocity of the disturbances is inversely proportional to the pressure amplitude, leads to the dispersion of the wave, which makes it impossible for the shock-wave to be steady.

The mathematical theory of shock-waves was originally developed by Rankine and Hugoniot for fluids. The development of the Rankine-Hugoniot equations relies on several assumptions, most notably that the shear modulus of the material

is zero, such that it responds to the wave as a fluid, that body forces (such as gravitational forces) and heat conduction at the shock front are negligible, and finally that there is no elastoplastic behaviour or phase transitions. The Rankine-Hugoniot equations can be easily developed by considering regions immediately ahead of and behind the shock front. Fig. 2.17 illustrates a shock-front. Ahead of the front, the pressure is P_0 , the density is ρ_0 , and the temperature is T_0 ; behind it they are P , ρ and T respectively. The velocity of the front is U_S , and the particles are stationary ahead of the front. At the front and behind it, they are moving at a velocity U_p . This displacement of particles is responsible for the pressure build-up. If one considers the centre of reference as the shock front and moves with it at a velocity U_S into a region of particle velocity $U_p = U_0$ and density ρ_0 , then the apparent velocity of the fluid moving towards the centre of reference is $U_S - U_0$. At the same time, the material leaving the front or receding from the centre of reference is moving at a velocity $U_S - U_p$. With this in mind, we will set up the equations for the conservation of mass, momentum and energy.

The equation of conservation of mass stems from the fact that the mass moving toward the front in time dt , given by $A\rho_0(U_S - U_0)dt$ must be equal to the mass moving away from the front, given by $A\rho(U_S - U_p)dt$. Hence, if $U_0 = 0$, we have:

$$\rho_0 U_S = \rho (U_S - U_p) \quad (2.113)$$

This is the equation for the conservation of mass.

The conservation of momentum requires that the difference in momentum, given by the product of mass and velocity, be equal to the impulse Fdt acting in time dt per unit cross sectional area. The difference in momentum across the

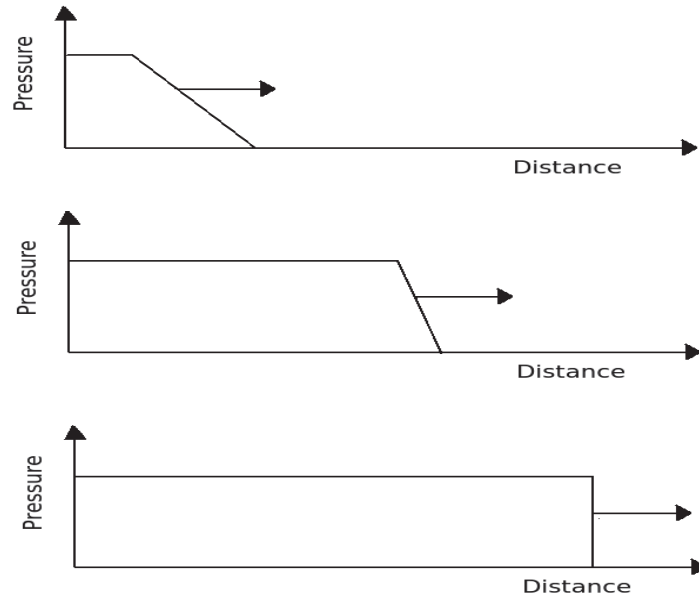


Figure 2.16: A schematic of the steepening of a ramped compression launched into a material sample.

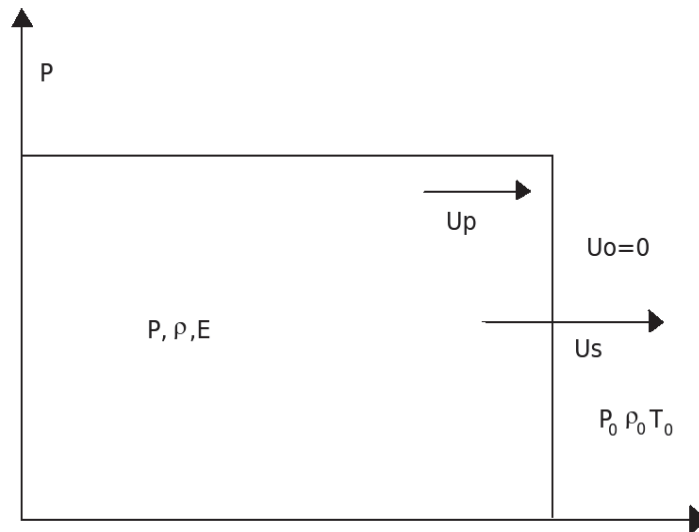


Figure 2.17: Schematic of a shock front moving at velocity U_S , with the definition of the material variables ahead of it and behind it.

shock front is therefore given by:

$$\underbrace{\rho A (U_S - U_p) dt}_{\text{mass}} \underbrace{U_p}_{\text{velocity}} - \underbrace{\rho_0 A (U_S - U_0) dt}_{\text{mass}} \underbrace{U_0}_{\text{velocity}} \quad (2.114)$$

The impulse on the other hand is given by:

$$F dt = (P A - P_0 A) dt \quad (2.115)$$

Then, equating change in momentum with impulse, we obtain:

$$\rho_0 (U_S - U_0) (U_p - U_0) = P - P_0 \quad (2.116)$$

If $U_0 = 0$:

$$(P - P_0) = \rho_0 U_S U_p \quad (2.117)$$

This is the equation for the conservation of momentum, and the quantity $\rho_0 U_S$ is often called the shock impedance.

The conservation of energy is obtained by setting up an equation in which the work done by P minus the work done by P_0 is equal to the difference in the total energy (kinetic plus internal) between the two sides of the front. The difference of work done by P and P_0 is given by:

$$\Delta W = \underbrace{(P A)}_{\text{force}} \underbrace{(U_p dt)}_{\text{distance}} - \underbrace{(P_0 A)}_{\text{force}} \underbrace{(U_0 dt)}_{\text{distance}} \quad (2.118)$$

The difference in total energy (kinetic plus internal) per unit mass is equal to the final minus the initial energy:

$$\Delta E = \frac{1}{2} [\rho A (U_S - U_p) dt] U_p^2 + E A \rho (U_S - U_p) dt - \left\{ \frac{1}{2} [\rho_0 A (U_S - U_0) dt] U_0^2 + E_0 A \rho_0 (U_S - U_0) dt \right\} \quad (2.119)$$

So equating ΔW to ΔE and taking $U_0 = 0$ yields:

$$P U_p = \frac{1}{2} \rho (U_S - U_p) U_p^2 - E_0 \rho_0 U_S + E \rho (U_S - U_p) \quad (2.120)$$

But from conservation of mass, $\rho (U_S - U_p) = \rho_0 U_S$. Substituting into the above we get:

$$P U_p = \frac{1}{2} \rho_0 U_S U_p^2 + \rho_0 U_S (E - E_0) \quad (2.121)$$

This equation can be simplified to obtain a more common form:

$$E - E_0 = \frac{P U_p}{\rho_0 U_S} - \frac{1}{2} \rho_0 \frac{U_S U_p^2}{\rho_0 U_S} \quad (2.122)$$

But, from conservation of momentum, $U_p = (P - P_0)/\rho_0 U_S$. Substituting for U_p in the above, we obtain:

$$E - E_0 = \frac{P (P - P_0)}{\rho_0^2 U_S^2} - \frac{1}{2} \frac{(P - P_0)^2}{\rho_0^2 U_S^2} \quad (2.123)$$

Again, from conservation of mass, $\rho_0 U_S = \rho (U_S - U_p)$, and using conservation of momentum:

$$\rho_0 U_S^2 = -\rho (P - P_0) \frac{1}{(\rho_0 - \rho)} \quad (2.124)$$

If $1/\rho = V$ then, simplifying, we get:

$$\rho_0^2 U_S^2 = \frac{P - P_0}{V_0 - V} \quad (2.125)$$

Substituting this back into Eq. (2.123), we get:

$$E - E_0 = \frac{1}{2} (P + P_0) (V_0 - V) \quad (2.126)$$

Eq. (2.126) is the more common form of the conservation of energy. In the three conservation equations of mass, momentum and energy, Equations (2.113), (2.117), (2.126), there are five variables: pressure P , particle velocity U_p , shock velocity U_S , specific volume V or density ρ , and energy E . Hence, an additional equation is needed if one wants to determine all parameters as a function of one of them. This fourth equation, which can be conveniently expressed as the relationship between shock and particle velocities, has to be experimentally determined. A polynomial equation with parameters $C_0, S_1, S_2, S_3, \dots$ empirically describes the relationship between U_S and U_p :

$$U_S = C_0 + S_1 U_p + S_2 U_p^2 + \dots \quad (2.127)$$

Eq. (5.38) is often known as the equation of state (EoS) of a material. Here, S_1 and S_2 are empirical parameters and C_0 is the sound velocity in the material at zero pressure. For most metals, $S_2 = 0$, and Eq. (5.38) thus reduces to a linear relationship:

$$U_S = C_0 + S_1 U_p \quad (2.128)$$

This linear relationship between U_S and U_p describes fairly well the shock

response of materials not undergoing phase transitions and with no porosity. Values of C_0 and S_1 are often tabulated in the literature; hence knowing these values and applying the conservation equations and the EoS, one can calculate the $P - U_S$, $P - U_p$, $P - \rho$, $P - V/V_0$, $E - U_S$, and other relationships. It is important to note that one needs only two shock parameters to determine the remaining ones once the constants in the EoS are known. There are in total five shock parameters with which we deal: P , E , ρ (or V), U_S and U_p . With the use of EoS, these relationships can be separated into 10 pairs. These 10 pairs provide, in turn, 20 equations, which are summarised in page 116 of Ref. ?.

2.4.3 The Hugoniot curve

Eq. (2.126), derived from the conservation of energy, establishes a relation between P and ρ immediately behind the shock-front. This pressure-density relationship is usually known as the Rankine-Hugoniot equation, or simply the "Hugoniot", and in graphic form is shown in Fig. 3.20 as a $P - V$ curve. The Hugoniot is therefore defined as the locus of all shocked states in a material and essentially describes the material properties. The straight line joining (P_0, V_0) and (P_1, V_1) is known as the Rayleigh line and refers to the shock state at P_1 .

It is very important to realise that when pressure is increased in a shock front, it does not follow the Hugoniot curve. Rather, it changes discontinuously from its initial value P_0 to its value P_1 as defined by the Rankien Hugoniot relationships. This discontinuity is explained by the slope of the Rayleigh line that is proportional to the square of the velocity U_S of the shock-wave. One can see this from the following derivation. Starting from the equation of conservation

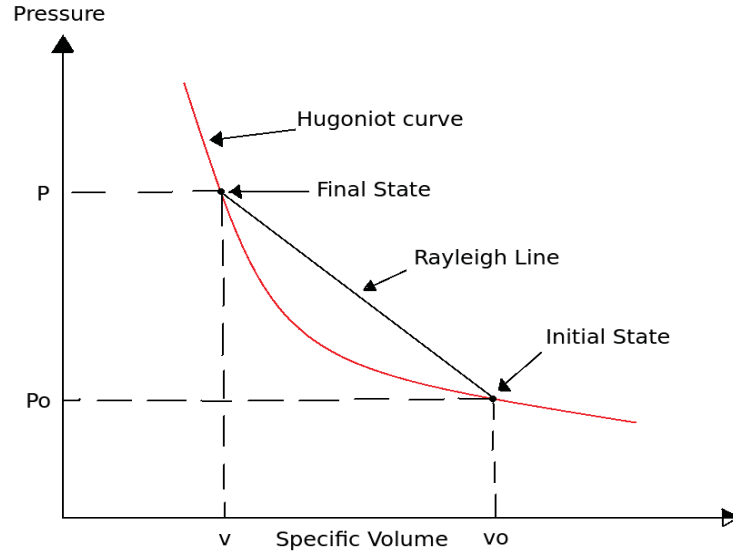


Figure 2.18: Hugoniot ($P - V$) curve denoting all the shocked states that can be attained in a material. The Rayleigh line connecting the initial to the final states is also shown.

of momentum:

$$P - P_0 = \rho_0 U_S U_p \quad (2.129)$$

or

$$\frac{P - P_0}{U_p} = \rho_0 U_S \quad (2.130)$$

But from conservation of mass we have:

$$\rho_0 U_S = \rho (U_S - U_p) \rightarrow U_p = U_S \frac{\rho - \rho_0}{\rho} \quad (2.131)$$

Thus:

$$\frac{P - P_0}{1 - \rho_0/\rho} = \rho_0 U_S^2 \quad (2.132)$$

But:

$$\frac{\rho_0}{\rho} = \frac{V}{V_0} \quad (2.133)$$

So:

$$\frac{P - P_0}{1 - V/V_0} = \rho_0 U_S^2 \quad (2.134)$$

This is equivalent to:

$$\frac{P - P_0}{V - V_0} = -\rho_0^2 U_S^2 \quad (2.135)$$

The left-hand side of the above equation is the slope of the Rayleigh line. One sees clearly that the higher the pressure, the higher the magnitude of the slope and the higher the velocity of the wave.

2.4.4 Real shock-wave profiles and the Hugoniot Elastic Limit

While an ideal shock-wave profile would predict a discontinuity at the front, a plateau at the top, and a gradual return to zero pressure, real shock-waves exhibit a number of peculiarities that are material and pressure dependent. The profiles of idealised and realistic shock-waves are shown in Fig. 2.19. The most important difference is that the pressure-volume curve for a real material is not the same as

the ideal Hugoniot curve because of the deviatoric component of stress present in the former. The difference between the two curves is shown in Fig. 2.20. Firstly, the rate of rise of stress with volume of the real material's curve is much higher in the elastic range. Secondly, when the elastic limit under the imposed stress and strain-rate conditions is reached above some limit of the value of the deviatoric stress, a dynamic yield point known as the Hugoniot Elastic Limit (HEL), the pressure-volume curve of the real material exhibits a change in slope and curves downward. The value of stress at the HEL is a material-specific parameter that specifies the elastic-to-plastic transition. The three arrows in Fig. 2.20 indicate the flow strength of the material at the imposed strains, defined as the difference between the ideal Hugoniot curve and the actual pressure-volume curve of the material. The three cases shown in Fig. 2.20 represent in turn:

- Curve 1: Flow strength of material is independent of pressure.
- Curve 2: Flow strength of material decreases with pressure, leading to softening.
- Curve 3: Flow strength of material increases with pressure, leading to hardening.

The HEL of metals is fairly low, and these effects are relatively unimportant. However, in ceramics this is not the case. For example, the HEL of sapphire is close to 20 GPa, whereas that of alumina is around 6-8 GPa. The position of the HEL separating the elastic portion of the wave from the plastic portion is shown in the realistic time profile of Fig. 2.19. This elastic portion, below the HEL, normally travels at a velocity higher than the plastic wave. The propagation

velocity of the elastic wave is equal to the longitudinal sound velocity, which is always larger than the plastic-wave velocity. For the elastic shock-wave, stress and strain at the shock-front can be given by:

$$\begin{aligned}\sigma_x &= (\lambda + 2\mu) \varepsilon_x \\ \sigma_y &= \sigma_z = \lambda \varepsilon_x\end{aligned}\tag{2.136}$$

for the case of shock-propagation along the x direction, where σ_x , σ_y and σ_z denote the components of the stress tensor, ε_x denotes the one-dimensional strain component, and λ , μ denote the Lamé' constants. If yielding occurs behind the elastic precursor wave, the shock yield stress Y can be given by:

$$Y = 2\mu\varepsilon_x\tag{2.137}$$

This result is obtained by assuming either the maximum shear stress or von Mises criterion for yielding due to the condition of one-dimensional strain.

At this point we are in a position to explain all the phenomena observed in Fig. 2.19(b). The leading wave is an elastic shock limited in amplitude by the shear strength of the material. In an ideal elastic-plastic solid, after the initial steep rise in pressure (or particle velocity), this elastic wave attains the HEL of the material. Permanent plastic deformation occurs through a splitting of the leading elastic precursor wave into two waves, an elastic wave and a slower plastic wave. The latter wave is what gives rise to the plastic behaviour. The ensuing rate of rise of the pressure is dictated by the constitutive behaviour of the material. In general, the rise becomes steeper as the shock-induced pressure increases. If there is a phase transition following the regime of plastic deformation, it is accompanied by

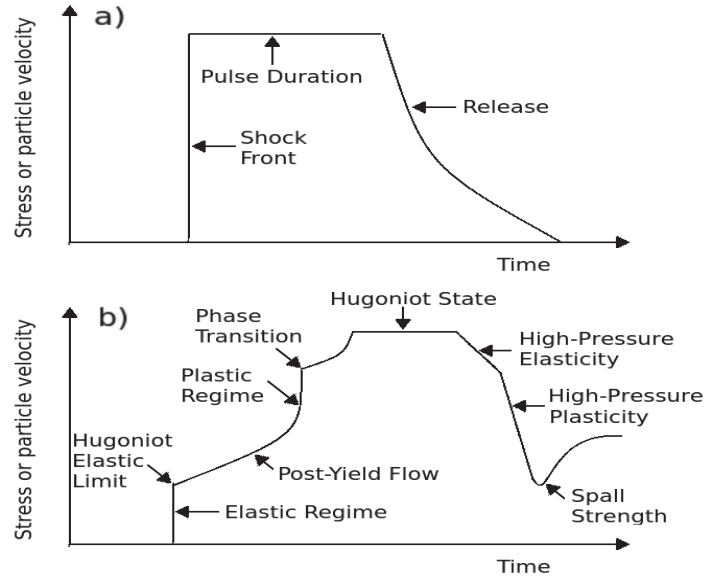


Figure 2.19: Idealised (a) and "generic" realistic (b) shock-wave stress or particle velocity profiles. In the latter case, the positions of the features in the profile characteristic of the onset of various types of material behaviour are labelled. This variety of material behaviour is explained in the text of this subsection.

a clear signal in the wave profile, which normally consists in the further splitting into two of the wave. The phase transition occurs via the arrival of the phase transition wave resulting from this splitting, and goes to completion if the driving pressure is large enough. This phase transition wave travels even more slowly than the plastic wave, which in turn travels more slowly than the elastic precursor. At the top of the plot of Fig. 2.19(b) we have the pulse duration plateau labelled as the Hugoniot State, since this represents the endstate which is attained on the characteristic pressure-volume curve of the material. When unloading starts following this plateau, it occurs initially elastically and then plastically. This elastoplastic transition in unloading leaves a signal analogous to the signal left by the HEL on loading. This signal is normally referred to as the Release or

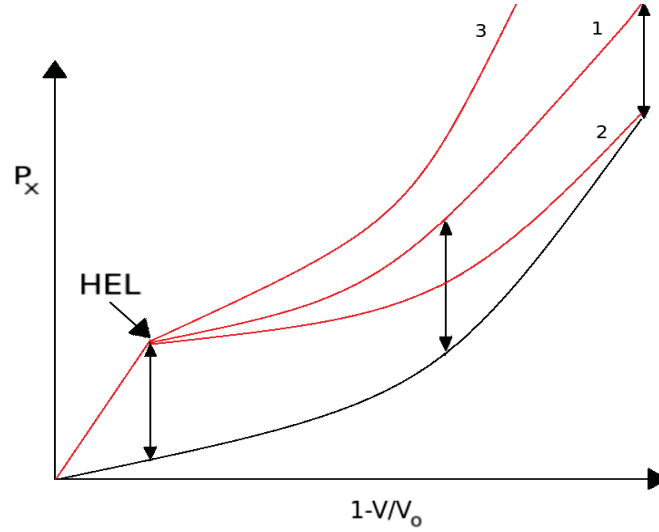


Figure 2.20: Representation of the pressure-volume curves for a real material (red) compared to that of an idealised material (black). The position of the HEL in the former is labelled. Three different red curves labelled 1,2 and 3 are shown, corresponding to the behaviour of three different materials.

Rarefaction wave, and returns the material to something approaching its original state. At the end of the unloading process, the wave reflects from the back surface of the sample and this phenomenon, known as spalling, can result in the fracturing of the material.

At high-pressures, structural phase transitions will take place in various materials. During these transformations, the lattice structure changes to a more stable configuration for the pressure and temperature realised by strong shock-compression. In such cases, the shock Hugoniot compression curve has the appearance as in Fig. 2.21. According to the thermodynamics of the problem, which will be explored in more detail in Chapter 4, the isotherm around first-order phase transitions in the pressure-volume plane has a horizontal line segment

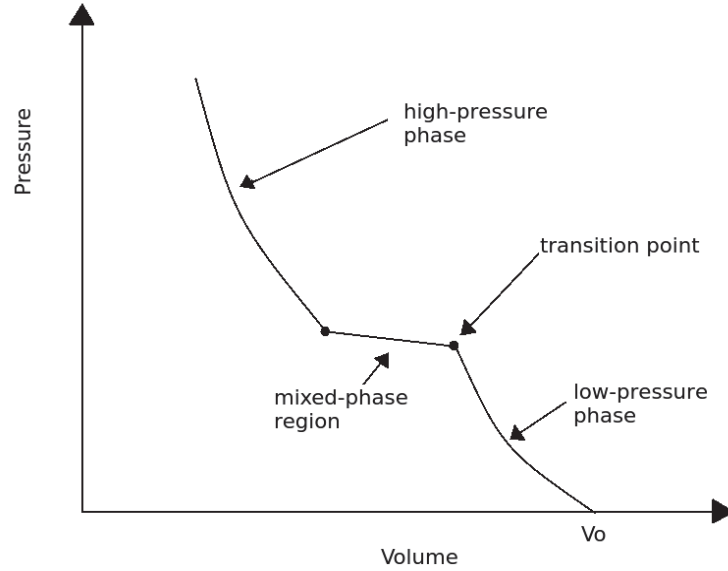


Figure 2.21: Hugoniot curve for a material undergoing a high-pressure polymorphic phase transition.

at the mixed-phase region. The shock-hugoniot in this region on the other hand should have a negative slope due to the temperature and entropy increase caused by the decreasing volume.

Finally, we analyse the response of the shock-loaded sample at the microstructural level of the underlying crystal lattice. It is widely accepted that the material morphology and timescales of atomistic phenomena have a profound impact on bulk properties, such as plasticity, phase transformations, and damage. Fig 2.22 shows the various stages of development of the shock-induced deformation of the crystal structure put in direct correlation with the various features of the real-material wave profile of Fig. 2.19. In the case of a planar shock considered here, the initial response of the crystalline lattice to shock-loading is a uniaxial, one-dimensional compression of the crystal along the direction of shock propagation. This uniaxial response can remain elastic, that is once the disturbance

is removed, the lattice will relax back to its original configuration. Furthermore, this uniaxial compression leads to the accumulation of strong levels of shear stress in the lattice. Under high-stress conditions and given sufficient time, the lattice will undergo an irreversible (permanent) deformation. Various morphologies of permanent deformation can occur, consisting in either plasticity, melting, resolidification, or solid-solid structural transformations. These atomistic changes can have dramatic and important consequences to most macroscopically observable behavior such as the materials thermodynamic state (pressure, temperature, density), strength, and failure. If the shock-wave propagating through the lattice exceeds the HEL of the material, the shear stress accumulated during the elastic phase of the compression leads to nucleation and propagation of dislocations, allowing the lattice to plastically relax to a quasi-hydrostatic, three-dimensional state of compression on some characteristic timescale. A fully 3D relaxed state is ideally defined as having zero shear stress and hydrostatic compression. The rate of relaxation to hydrostatic conditions is controlled by the nucleation rates and mobility of dislocations, which in turn depend upon microstructural parameters such as grain boundaries, barriers, etc. Experimentally, this relaxation of the lattice behind a shock has been studied with time-resolved X-ray diffraction in FCC metals. As an example, such kinetics have been studied in single crystal copper shocked to 18 GPa. The degree of lattice relaxation was determined through diffraction recorded from planes both parallel and perpendicular to the shock loading direction. These samples were found to have approached a fully 3D state in less than 400 ps after the passage of the shock front. Under shock compression, solids can also exhibit a rich variety of solid-to-solid phase transformations, some of which are metastable and, due to the uniaxial and ultrafast nature of

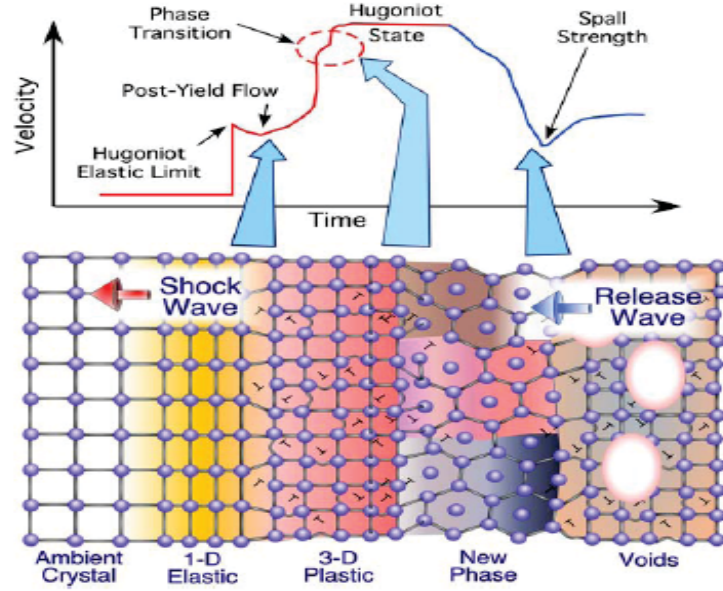


Figure 2.22: A schematic diagram of a shocked crystalline solid, linking microstructural processes to the bulk behavior measured by inspection of the pressure or particle velocity wave profile. The rearrangement of the atoms as the lattice relaxes plastically or undergoes a phase-transition alters the response of the material, causing the corresponding change in the wave profile.

the compression, may differ from those encountered under longer timescale and more hydrostatic conditions such as during static DAC-based experiments. Phase transitions in general provide an alternative microscopic mechanism for the relief of shear stress from plastic deformation. Carbon, for example, demonstrates a wide range of structural phases along the Hugoniot, including cubic, hexagonal, nanocrystalline, and amorphous diamond. Solid-to-melt transformations can also occur under sufficiently strong shocks.

Historically, two approaches have been employed to probe the dynamic response properties of materials to shock-compression: sample recovery and *in situ* bulk property measurements. The first approach affords for microstructural analysis but does not allow direct probing of the samples under dynamic loading. The

examination of shock recovered specimens in fact provides only end-state information that can only be useful for *inferring* dynamic behaviour. In the second approach, real-time measurements of the bulk response are recorded with fast diagnostics such as surface velocimetry (VISAR) and internal stress gauges. While such approaches have proven themselves to be valuable and are largely responsible for our current understanding, these continuum level methodologies are limited in the insight that they can provide about atomistic lattice-level processes under dynamic conditions. It is clear that a fundamental understanding requires the direct probing and study of relevant, transient physical processes at their characteristic lattice length (of order Å) and time (of order picosecond) scales. The technologies of ultrabright, pulsed X-ray sources for *in situ* time-resolved X-ray diffraction experiments, and massively parallel computation employing such techniques as classical Molecular Dynamics described in the next section, have emerged as particularly promising for investigating the shock-response of solids at the atomistic level. In particular, there exist two main classes of experimental approaches: the first one is based on the use of high-power lasers for both driving the shock-wave into the sample and generating X-rays by plasma-line emission from a backlighting metallic foil. The second approach is accelerator-based such as realised in synchrotrons or at 4th generation X-ray Free Electron Laser sources such as LCLS. The latter class of facilities in particular possesses unparalleled X-ray beam quality, brightness, pulse duration shortness and repetition rate for performing time-resolved measurements. These approaches are complementary in their capacity to probe from the nano-to-macroscale in the spatial and picosecond-to-nanosecond in the temporal regimes.

2.5 Classical Molecular Dynamics

2.5.1 Introduction

A general definition of classical molecular dynamics (MD) simulation is a computation technique by which the atomic trajectories of a system of N particles are generated by numerical integration of Newton's equation of motion (within a given statistical ensemble). This integration is performed in conjunction with the operation of a specific interatomic potential to calculate the forces on each atom, and with certain initial conditions (IC) and boundary conditions (BC). Let us consider for example a system with N atoms in a volume Ω . The internal energy of the system E can be defined as the sum of the kinetic energy K and potential energy U , $E = K + U$. The kinetic energy is given by:

$$K = \sum_{i=1}^N \frac{1}{2} m_i |\dot{\mathbf{x}}_i(t)|^2 \quad (2.138)$$

whereas the potential energy can be expressed as:

$$U = U(\mathbf{x}^{3N}(t)) \quad (2.139)$$

where $\mathbf{x}^{3N}(t)$ denotes the group of three-dimensional coordinates $\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_N(t)$.

If the system is isolated, then E is a conserved quantity and therefore constant with time. MD simulations can be structured into four main steps, very much like experiments: system setup, equilibration, simulation run and output. In the equilibration step, we fine-tune the system until it reaches the desired conditions, such as for example temperature T and pressure P , and then in the simulation

step we perform property averages, for example calculating the radial distribution function $g(r)$ or thermal conductivity.

2.5.2 Elements of Molecular Dynamics simulations

There are five key ingredients to a MD simulation, which are boundary conditions, initial conditions, force calculation, integrator/ensemble, and property calculation. A brief overview of each one of these elements will now be given.

There are two major types of boundary conditions: isolated boundary condition (IBC) and periodic boundary condition (PBC). IBC is best suited for studying clusters and molecules, while PBC is suited for studying bulk liquids and solids. There can also be mixed boundary conditions such as slab or wire configurations for which the system is assumed to be periodic in some directions but not in the others. In IBC, the N -particle system is surrounded by vacuum. These particles are still able to interact among themselves, but they are so far away from everything else in the universe that no interactions with the outside occur except perhaps responding to some well-defined external forcing. In PBC on the other hand, the motion of the N particles in the supercell are followed, but the supercell is surrounded by infinitely replicated, periodic images of itself. Therefore a particle may interact not only with particles in the same supercell, but also with particles in adjacent image supercells. The simplest and most often used space filling unit which can serve as the PBC supercell is a parallelepiped, specified by its three edge vectors \mathbf{h}_1 , \mathbf{h}_2 and \mathbf{h}_3 . IBC can in fact often be well reproduced by a large enough PBC supercell such that the images do not interact among themselves.

Since Newton's equations of motion are second-order ordinary differential equations (ODE), specifying the initial conditions basically means specifying $\mathbf{x}^{3N}(t=0)$ and $\dot{\mathbf{x}}^{3N}(t=0)$, the initial particle positions and velocities. For crystalline solids, generating the IC is usually straightforward, but for liquids and amorphous solids less so. Liquids and amorphous solids can be generated by melting and quenching of crystalline solids during a MD run. Let us focus on IC for crystalline solids only. $\mathbf{x}^{3N}(t=0)$ can for example be a perfect fcc crystal, or an interface between two crystalline phases. According to the equipartition theorem, each independent degree of freedom should possess $k_B T/2$ kinetic energy. So, one should draw each component of the $3N$ -dimensional $\dot{\mathbf{x}}^{3N}(t=0)$ vector from a Gaussian-Maxwell normal distribution. After that, the centre of mass velocity should be eliminated, and for clusters, the net angular momentum as well.

Two approximations are employed in the classical equation of motion to describe the motion of the atoms:

$$m_i \frac{d^2 \mathbf{x}_i(t)}{dt^2} = \mathbf{f}_i = -\frac{\partial U}{\partial \mathbf{x}_i}, i = 1, \dots, N \quad (2.140)$$

The first is the Born-Oppenheimer approximation, which assumes that the electronic state couples adiabatically to nuclei motion. The second is that the nucleus motion is far removed from the Heisenberg uncertainty lower bound: $\Delta E \Delta t \gg \hbar/2$. If we insert $\Delta E = k_B T/2$, the kinetic energy, and $\Delta t = 1/\omega$, where ω is a characteristic vibrational frequency, we obtain $k_B T/\hbar\omega \gg 1$. In solids, this means that the temperature should be significantly greater than the Debye temperature. The evaluation of the right-hand side of Eq. (2.140) is

usually the most computationally-expensive step in a MD simulation, so its efficiency is of crucial importance. There exist special algorithms to break up long range Coulomb interactions into two contributions, a short-range interaction and a smooth field-like interaction, both of which can be computed efficiently in separate ways.

Eq. (2.140) is a set of second-order ordinary differential equations, which can be strongly non-linear. By converting them to first-order Ordinary differential equations in the $6N$ -dimensional space of $\{\mathbf{x}_N, \dot{\mathbf{x}}_N\}$, general numerical algorithms for solving ordinary differential equations such as the Runge-Kutta method can be applied. However, these general methods are rarely used in MD, because the existence of a Hamiltonian allows for more accurate integration algorithms, including the family of predictor-corrector integrators and the family of symplectic integrators. Integrators will be discussed in more detail in the next section. Ensembles such as the micro-canonical, canonical and grand-canonical are concepts in statistical physics that refer to the distribution of initial conditions. However, ensemble and integrator are often grouped together because it is possible to generate the desired ensemble distribution via time integration.

One of the great powers of classical MD is that it provides a nearly complete amount of information at the atomistic level. All properties that are well defined in classical mechanics and statistical mechanics can in principle be computed. The only limiting factors are the accuracy of the calculations, which is limited by the choice of the interatomic potential, and the computational efficiency. The material properties that can be calculated by MD can be roughly grouped into four categories. Firstly, we have structural characterisation, which includes for example the calculation of radial distribution functions and dynamic structure factors.

Secondly, we have equation of state calculations, which includes the calculation of free energy functions, phase diagrams, and static response functions like the thermal expansion coefficient. A third type of calculations are those which pertain to the transport properties of the material, which includes for example viscosity, thermal conductivity, correlation functions and diffusivity. Finally, MD is capable of calculating the non-equilibrium response of the sample, which includes such phenomena as plastic deformation and pattern formation.

2.5.3 Integrators

An integrator is defined as the algorithm which advances the trajectory of the atoms in the MD simulation over small time increments Δt :

$$\mathbf{x}^{3N}(t_0) \rightarrow \mathbf{x}^{3N}(t_0 + \Delta t) \rightarrow \mathbf{x}^{3N}(t_0 + 2\Delta t) \rightarrow \dots \rightarrow \mathbf{x}^{3N}(t_0 + L\Delta t) \quad (2.141)$$

where L usually ranges from 10^4 to 10^7 . We will now give a brief overview of some popular integrator algorithms: central difference (Verlet, leap-frog, velocity Verlet and Beeman algorithm), predictor-corrector, and symplectic integrators.

In the Verlet Algorithm, assuming the $\mathbf{x}^{3N}(t)$ trajectory is smooth, we perform the Taylor expansion:

$$\mathbf{x}_i(t_0 + \Delta t) + \mathbf{x}_i(t_0 - \Delta t) = 2\mathbf{x}_i(t_0) + \ddot{\mathbf{x}}_i(t_0)(\Delta t)^2 + \mathcal{O}((\Delta t)^4) \quad (2.142)$$

Since $\ddot{\mathbf{x}}_i(t_0) = \mathbf{f}_i(t_0)/m_i$ can be evaluated given the atomic positions $\mathbf{x}^{3N}(t_0)$ at

$t = t_0$, $\mathbf{x}^{3N}(t_0 + \Delta t)$ in turn may be approximated by:

$$\mathbf{x}_i(t_0 + \Delta t) = -\mathbf{x}_i(t_0 - \Delta t) + 2\mathbf{x}_i(t_0) + \left(\frac{\mathbf{f}_i(t_0)}{m_i}\right)(\Delta t)^2 + \mathcal{O}((\Delta t)^4) \quad (2.143)$$

By discarding the $\mathcal{O}((\Delta t)^4)$ term, we obtain a recursion formula to compute $\mathbf{x}^{3N}(t_0 + \Delta t), \mathbf{x}^{3N}(t_0 + 2\Delta t), \dots$ successively, which is the Verlet algorithm. The velocities do not appear in the recursion but are needed for property calculations. They are given approximately by:

$$\mathbf{v}_i(t_0) \equiv \dot{\mathbf{x}}_i(t_0) = \frac{1}{2\Delta t} [\mathbf{x}_i(t_0 + \Delta t) - \mathbf{x}_i(t_0 - \Delta t)] + \mathcal{O}((\Delta t)^2) \quad (2.144)$$

We now investigate to what degree the above recursion reproduces the real trajectory $\mathbf{x}^{3N}(t)$. If the $\mathbf{x}_i(t_0)$ and $\mathbf{x}_i(t_0 - \Delta t)$ terms in Eq. (2.143) are exact, and assuming the computer employed is perfect and has no machine errors in storing the numbers or carrying out floating-point operations, the computed $\mathbf{x}_i(t_0 + \Delta t)$ would still be different from the real $\mathbf{x}_i(t_0 + \Delta t)$ by $\mathcal{O}((\Delta t)^4)$, which is defined as the local truncation error (LTE). LTE is an intrinsic error of the Verlet algorithm. In addition to LTE, there is a round-off error due to the computer's finite precision.

In the so-called Leap-frog Algorithm, we start out with $\mathbf{v}^{3N}(t_0 - \Delta t/2)$ and $\mathbf{x}^{3N}(t_0)$. The iterative scheme to advance by one step is given by:

$$\mathbf{v}_i\left(t_0 + \frac{\Delta t}{2}\right) = \mathbf{v}_i\left(t_0 - \frac{\Delta t}{2}\right) + \left(\frac{\mathbf{f}_i(t_0)}{m_i}\right) \Delta t \quad (2.145)$$

followed by:

$$\mathbf{x}_i(t_0 + \Delta t) = \mathbf{x}_i(t_0) + \mathbf{v}_i\left(t_0 + \frac{\Delta t}{2}\right) \Delta t \quad (2.146)$$

The velocity at time t_0 can be approximated by:

$$\mathbf{v}_i(t_0) = \frac{1}{2} \left[\mathbf{v}_i \left(t_0 - \frac{\Delta t}{2} \right) + \mathbf{v}_i \left(t_0 + \frac{\Delta t}{2} \right) \right] + \mathcal{O}((\Delta t)^2) \quad (2.147)$$

Thirdly, in the velocity Verlet algorithm, we start with $\mathbf{x}^{3N}(t_0)$ and $\mathbf{v}^{3N}(t_0)$, and then we have:

$$\mathbf{x}_i(t_0 + \Delta t) = \mathbf{x}_i(t_0) + \mathbf{v}_i(t_0) \Delta t + \frac{1}{2} \left(\frac{\mathbf{f}_i(t_0)}{m_i} \right) (\Delta t)^2 \quad (2.148)$$

We next evaluate $\mathbf{f}^{3N}(t_0 + \Delta t)$, and then:

$$\mathbf{v}_i(t_0 + \Delta t) = \mathbf{v}_i(t_0) + \frac{1}{2} \left(\frac{\mathbf{f}_i(t_0)}{m_i} + \frac{\mathbf{f}_i(t_0 + \Delta t)}{m_i} \right) \Delta t \quad (2.149)$$

and we have advanced by one step. Since this algorithm can yield simultaneously $\mathbf{x}^{3N}(t_0)$ and $\mathbf{v}^{3N}(t_0)$, it is commonly used in MD codes.

Fourthly, we consider the so-called Beeman algorithm, which is similar to the velocity Verlet algorithm. We start with $\mathbf{x}^{3N}(t_0)$, $\mathbf{f}^{3N}(t_0 - \Delta t)$, $\mathbf{f}^{3N}(t_0)$ and $\mathbf{v}^{3N}(t_0)$, and then we have for the iterative scheme to advance by one step:

$$\mathbf{x}_i(t_0 + \Delta t) = \mathbf{x}_i(t_0) + \mathbf{v}_i(t_0) \Delta t + \left[\frac{4\mathbf{f}_i(t_0) - \mathbf{f}_i(t_0 - \Delta t)}{m_i} \right] \frac{(\Delta t)^2}{6} \quad (2.150)$$

We next evaluate $\mathbf{f}^{3N}(t_0 + \Delta t)$, and finally we obtain for the velocity:

$$\mathbf{v}_i(t_0 + \Delta t) = \mathbf{v}_i(t_0) + \left[\frac{2\mathbf{f}_i(t_0 + \Delta t) + 5\mathbf{f}_i(t_0) - \mathbf{f}_i(t_0 - \Delta t)}{m_i} \right] \frac{\Delta t}{6} \quad (2.151)$$

Just like the leap-frog and the velocity-Verlet algorithms, the Beeman algorithm

gives identical trajectory as the Verlet algorithm in the absence of machine error, with a fourth-order LTE in position. However, it gives better velocity estimate (third-order LTE) than the leap-frog or velocity Verlet (second-order LTE). The best velocity estimate (fourth-order LTE) can be achieved by the so-called velocity-corrected Verlet algorithm, but it requires knowledge of the next two steps' positions.

We now describe the predictor-corrector algorithm. Let us take the commonly used 6-value predictor-corrector algorithm as an example. We start with the $6 \times 3N$ storage composed of $\mathbf{x}^{3N(0)}(t_0)$, $\mathbf{x}^{3N(1)}(t_0)$, $\mathbf{x}^{3N(2)}(t_0)$, ..., $\mathbf{x}^{3N(5)}(t_0)$, where $\mathbf{x}^{3N(k)}(t)$ is defined by:

$$\mathbf{x}_i^{(k)}(t) \equiv \left(\frac{d^k \mathbf{x}_i^{(t)}}{dt^k} \right) \left(\frac{(\Delta t)^k}{k!} \right) \quad (2.152)$$

The iteration consists of the prediction, evaluation and correction steps. The prediction step is based on the following general formula:

$$\mathbf{x}_i^{(k)} = \sum_{k'=k}^{M-1} \left[\frac{k'!}{(k'-k)!k!} \right] \mathbf{x}_i^{(k')}, \quad k = 0, \dots, M-2 \quad (2.153)$$

with $M = 6$ here. The evaluation must proceed from 0 to $M - 2$ sequentially. In the subsequent evaluation step, we evaluate the force \mathbf{f}^{3N} using the newly obtained $\mathbf{x}^{3N(0)}$. Finally, in the correction step, we define the error \mathbf{e}^{3N} as:

$$\mathbf{e}_i \equiv \mathbf{x}_i^{(2)} - \left(\frac{\mathbf{f}_i}{m_i} \right) \left(\frac{(\Delta t)^2}{2!} \right) \quad (2.154)$$

Table 2.5: Gear predictor-correctors coefficients.

C_{Mk}	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
$M = 4$	1/6	5/6	1	1/3				
$M = 5$	19/120	3/4	1	1/2	1/12			
$M = 6$	3/20	251/360	1	11/18	1/6	1/60		
$M = 7$	863/6048	665/1008	1	25/36	35/144	1/24	1/360	
$M = 8$	1925/14112	19087/30240	1	137/180	5/16	17/240	1/120	1/2520

Then we apply corrections:

$$\mathbf{x}_i^{(k)} = \mathbf{x}_i^{(k)} - C_{Mk} \mathbf{e}_i, \quad k = 0, \dots, M - 1 \quad (2.155)$$

where C_{Mk} are constants listed in Table 2.5. The LTE for \mathbf{x}^{3N} is $\mathcal{O}\left((\Delta t)^M\right)$ after the prediction step. But one can show that the LTE is reduced to $\mathcal{O}\left((\Delta t)^{M+1}\right)$ after the correction step if \mathbf{f}^{3N} depends on \mathbf{x}^{3N} only, and not on the velocity.

Finally, we consider the case of Symplectic Integrators, which have the unique quality of being able to rigorously maintain the phase-space volume conservation property of Hamiltonian dynamics (Liouville's theorem), in the absence of round-off errors. This property severely limits the possibilities of mapping from initial to final states, and for this reason symplectic integrators tend to have much better total energy conservation in the long run. The velocity Verlet algorithm is in fact symplectic, followed by higher order extensions.

2.5.4 The Tersoff Potential

The semi-empirical interatomic potentials employed in MD simulations permit the calculation of structural properties and energetics of complex systems. In particular interatomic potentials give the total energy E of a set of particles

as an explicit mathematical function of the set $\{\mathbf{r}\}$ of particle coordinates. If this function is sufficiently easy to calculate, and if it gives a sufficiently accurate description of the real physical system of interest, then one can perform realistic calculations of the properties of relatively large systems. Of course the accuracy delivered by such potential-based calculations will always inevitably be worse than that characteristic of truly quantum-mechanical ab-initio methods such as Density Functional Theory. Most MD potentials are created by fitting the adjustable parameters to bulk properties of the material under investigation. Common material properties used for fitting empirical potentials include lattice parameter, bulk modulus and elastic tensor components.

Pair potentials, like the notable Lennard-Jones potential and the exponential Morse potential, have long been used to describe rare-gas atoms, simple metals, and highly ionic systems. These potentials describe the interaction of atoms as a function of their separation, including only two-body interactions. Such potentials can be directly applied to a completely arbitrary configuration of atoms but do not accurately describe any but the simplest closed-shell systems. In particular, pair-potentials are completely inapplicable to strongly covalent systems such as semiconductors. This pair potential approach corresponds to the leading term in a mathematical expansion of the energy, viewed as a function of the atomic positions. The energy of N interacting particles may be written as:

$$E = \sum_i V_i(\mathbf{r}_i) + \sum_{i<j} V_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i<j<k} V_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \quad (2.156)$$

where \mathbf{r}_n is the position of the n -th particle and the function V_m is called an m -body potential. The first (one-body) term corresponds to an external potential.

The first term which describes the interactions between the particles is the second (two-body) term, which when taken alone constitutes a pair potential. Thus, in this expansion, the pair potential is the simplest possible model for the interaction of a set of particles. A general feature of atomic pair potentials is that they favour the formation of closed-packed structures, and therefore are unsuitable for describing covalent systems, which assume more open structures.

The more recent Tersoff scheme offers a new approach for constructing potentials affording for a improved description of covalent materials such as semiconductors like silicon and other interesting materials such as ceramics, polymers and refractory metals. The central idea behind Tersoff is that, in real systems, the bond order (that is, the strength of each bond) depends upon the local environment. In particular, an atom with many neighbours forms weaker bonds than an atom with few neighbours. For covalent materials, this dependence is sufficient to stabilise structures with low atomic coordination number. Including this dependence explicitly in the Tersoff scheme appears to solve some of the most serious problems of describing covalent systems without introducing any major increase in computational complexity. Tersoff appears to be particularly suited for modelling silicon, which represents a particular challenge because it has many polymorphs with qualitatively different bonding which, nevertheless, have similar cohesive energies. For this reason, Tersoff is by far the most widely employed potential for modelling silicon.

For describing covalent systems, a natural first step was to include the next term in the expansion of Eq. (2.156), and therefore obtain a three-body potential. This additional term could stabilise more open structures. However, it was found that a three-body potential was inadequate for accurately describing

the cohesive energy of silicon over a wide range of bonding geometry and coordination. However, at the same time, a general form for a four- or a five-body potential would probably prove intractable, and would contain far too many free parameters. As an alternative, the Tersoff approach was introduced going beyond the three-body potential. In attempting to construct an accurate and tractable potential, it abandons the use of a general N -body form and attempts to identify the relevant physics and to build it directly into the form of the potential.

It is a well-known fact that the more neighbours an atom has, the weaker the bond to each neighbour will be. The bond strength, or bond order, in general depends in a complicated way on the geometry. However the most important single variable appears to be the coordination number, that is the number of neighbours close enough to form bonds. In particular, the bond order appears to be a monotonically decreasing function of coordination, with the trade-off between bond-order and number of bonds determining the equilibrium coordination. Silicon in particular is notable for assuming structures with a large range of coordination over modest changes of pressure. The difference in cohesive energy among these structures is remarkably small. This is because the decrease in bond strength with increasing coordination number very nearly cancels the increase in the number of bonds, over a large range of coordination. Silicon therefore provides a particularly stringent test of the ability of the Tersoff potential to accurately describe the dependence of bonding upon coordination, and therefore the dependence of cohesion upon structure.

Because of the crucial role of bond order and its dependence upon local geometry, Tersoff includes an environment-dependent bond order explicitly into the

potential in the following way:

$$\begin{aligned}
 E &= \sum_i E_i = \frac{1}{2} \sum_{i \neq j} V_{ij} \\
 V_{ij} &= f_C(r_{ij}) [a_{ij} f_R(r_{ij}) + b_{ij} f_A(r_{ij})]
 \end{aligned}
 \tag{2.157}$$

Here E is the total energy of the system, which is decomposed into a site energy E_i and a bond energy V_{ij} . The indices i and j run over the atoms of the system, and r_{ij} is the distance from atom i to atom j . The function f_R represents a repulsive pair potential and f_A represents an attractive pair potential associated with bonding. The extra term f_C is merely a smooth cutoff function, to limit the range of the potential, since for many applications short-ranged functions permit a significant reduction in computational expense. The function b_{ij} is the only novel feature of the Tersoff potential. It represents a measure of the bond-order, and is assumed to be a monotonically decreasing function of the coordination of atoms i and j . Determining a satisfactory form for b_{ij} is by far the most difficult part of applying the Tersoff approach.

The functions f_R , f_A , f_C , a_{ij} and especially b_{ij} still need to be determined. The choice of exponential functions for f_R and f_A , as in a Morse potential, proves to be a convenient one:

$$\begin{aligned}
 f_R(r) &= A \exp(-\lambda_1 r) \\
 f_A(r) &= -B \exp(-\lambda_2 r)
 \end{aligned}
 \tag{2.158}$$

The cutoff function is simply taken as:

$$f_C(r) = \begin{cases} 1, & r < R - D \\ \frac{1}{2} - \frac{1}{2} \sin \left[\frac{\pi}{2}(r - R)/D \right], & R - D < r < R + D \\ 0, & r > R + D \end{cases} \quad (2.159)$$

which has continuous value and derivative for all r , and goes from 1 to 0 in a small range around R . R is chosen to include only the first-neighbour shell for most structures of interest. The parameters R and D specify the position and the width of the cutoff region. The short range of the potential is numerically advantageous in many applications. Furthermore, b_{ij} is taken to have the following form:

$$\begin{aligned} b_{ij} &= (1 + \beta^n \xi_{ij}^n)^{-1/2n} \\ \xi_{ij} &= \sum_{k \neq i, j} f_C(r_{ik}) g(\theta_{ijk}) \exp [\lambda_3^3 (r_{ij} - r_{ik})^3] \\ g(\theta) &= 1 + c^2/d^2 - c^2/[d^2 + (h - \cos \theta)^2] \end{aligned} \quad (2.160)$$

where θ_{ijk} is the bond angle between bonds ij and ik . Finally, the form proposed by Tersoff for a_{ij} is:

$$\begin{aligned} a_{ij} &= (1 + \alpha^n \eta_{ij}^n)^{-1/2n} \\ \eta_{ij} &= \sum_{k \neq i, j} f_C(r_{ik}) \exp [\lambda_3^3 (r_{ij} - r_{ik})^3] \end{aligned} \quad (2.161)$$

with α taken sufficiently small that $a_{ij} \simeq 1$ unless η_{ij} is exponentially large, which will only occur for atoms outside the first-neighbour shell. The terms a_{ij} and b_{ij} contain the three-body interactions and modify the strengths of the repulsive and attractive terms according to the bonds local environment. The resulting Tersoff potential has performed quite well in describing many properties of silicon.

The Tersoff potential was further refined by Erhart and Albe for the cases of silicon, carbon and silicon carbide with the addition of some new free parameters. The cohesive energy is now written as a sum over individual bond energies:

$$E = \sum_{i>j} f_C(r_{ij}) \left[V_R(r_{ij}) - \underbrace{\frac{b_{ij} + b_{ji}}{2}}_{b_{ij}} V_A(r_{ij}) \right] \quad (2.162)$$

with the pairwise attractive and repulsive contributions given by:

$$V_R(r) = \frac{D_0}{S-1} \exp \left[-\beta \sqrt{2S} (r - r_0) \right] \quad (2.163)$$

and

$$V_A(r) = \frac{SD_0}{S-1} \exp \left[-\beta \sqrt{2/S} (r - r_0) \right] \quad (2.164)$$

where D_0 and r_0 are the dimer energy and bond length. β and S are adjustable free parameters of the potential. The cutoff function is the same as that of the general Tersoff potential. The bond-order is given by:

$$b_{ij} = (1 + \chi_{ij})^{-1/2} \quad (2.165)$$

with

$$\chi_{ij} = \sum_{k \neq i,j} f_C(r_{ik}) \exp [2\mu (r_{ij} - r_{ik})] g(\theta_{ijk}) \quad (2.166)$$

and the angular function:

$$g(\theta) = \gamma \left(1 + \frac{c^2}{d^2} - \frac{c^2}{d^2 + [h + \cos \theta]^2} \right) \quad (2.167)$$

The three-body interactions are determined by the parameters 2μ , γ , c , d , and h , which leads in total to up to nine adjustable parameters, all of them depending on the type of atoms i and j . This new bond-order potential for silicon, carbon and silicon carbide reproduces well the elastic, thermal and point defect properties. The description of higher coordinated structures for silicon and carbon is comparable to the Tersoff potential, while elastic, defect and thermal properties are superior.

2.6 Density Functional Theory

2.6.1 The Kohn-Sham formulation of Density Functional Theory

The *raison d'être* of Density Functional Theory (DFT) is the solution of the many-body Hamiltonian for a system of N electrons and M positively-charged nuclei (or ions) in a crystalline solid, the starting point for nearly all problems in solid-state physics. Neglecting all relativistic and magnetic effects, this Hamiltonian can be written as:

$$\begin{aligned} \widehat{H}_N = & -\frac{\hbar^2}{2m_e} \sum_{i=1}^N \nabla_i^2 - \sum_{i,I=1}^{N,M} \frac{Z_I e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i,j}^{i \neq j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \\ & - \sum_{I=1}^M \frac{\hbar^2}{2M_I} \nabla_I^2 + \frac{1}{2} \sum_{I,J}^{I \neq J} \frac{Z_I Z_J e^2}{4\pi\epsilon_0 |\mathbf{R}_I - \mathbf{R}_J|} \end{aligned} \quad (2.168)$$

where the electrons are denoted by lower case subscripts and the nuclei, with charge Z_I and mass M_I , by upper case subscripts. The terms in this Hamiltonian represent, from left to right, the kinetic energy of the electrons, the attractive

potential acting on the electrons due to the nuclei, the repulsive electron-electron interaction, the kinetic energy of the nuclei and finally the repulsive internuclear electrostatic interaction. In practice, such an elementary equation as Eq. 2.168 is impossible to solve analytically for all but the most trivial systems because of the enormously complicating effects of the interactions between electrons, which leads to coupling between the electronic coordinates in the system. This coupling manifests itself under the well-known phenomena of inter-electron exchange and correlation interactions in Quantum Mechanics . The issue central to the theory of electronic structure is therefore the development of approximate methods using simplifying physical ideas to treat electronic interactions and correlations with sufficient accuracy such that, starting from Eq. 2.168, one can still gain insight into the diverse array of electronic properties and phenomena exhibited by real solid matter.

One initial simple approximation involves setting the mass of the ions M_I to infinity, which is tantamount to saying that the positions of the ions are assumed fixed to their corresponding lattice points within the crystal structure. This so-called Born-Oppenheimer, or Adiabatic, approximation in electronic structure calculations allows the ionic kinetic energy term in Eq. 2.168 to be ignored, which is an excellent approximation for most intent and purposes. Neglecting also the final inter-ionic electrostatic interaction term, which is essential in total energy calculations but reduces to a simple classical additive term within the Born-Oppenheimer approximation, the initial problem of solving the electronic structure for the entire crystalline solid is consequently simplified to the treatment of the crystal electron cloud exclusively. The corresponding N -electron

Schroedinger equation can be rewritten as:

$$\left[\sum_{i=1}^N \left(-\frac{\hbar}{2m_e} \nabla_i^2 + V_{ion}(\mathbf{r}_i) \right) + \frac{1}{2} \sum_{i,j}^{i \neq j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \right] \Psi_N(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2; \dots, \mathbf{r}_N, \sigma_N) = E_N \Psi_N(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2; \dots, \mathbf{r}_N, \sigma_N) \quad (2.169)$$

where $V_{ion}(\mathbf{r}_i) = -\sum_{I=1}^M \frac{Z_I e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{R}_I|}$ represents the combined ionic potential acting on each electron, E_N is the total electronic energy and $\Psi_N(r_1, \sigma_1; r_2, \sigma_2; \dots, r_N, \sigma_N)$ is the N -electron wavefunction expressed in terms of the spatial (\mathbf{r}_i) and spin (σ_i) quantum numbers for the i^{th} electron in the material.

Historically, an important milestone for going beyond the independent-electron approximation, embodied by the Sommerfeld free-electron model, and incorporating approximately the effect of inter-electron interactions in solids was the development of the Hartree-Fock equations¹. Even though they were the first technique which allowed for a neat separation between electron exchange and correlation effects, in practice their treatment remains difficult and requires careful further approximations. The need for more accurate, widely-applicable and easy-to-implement methods for calculating the ground state properties of many-body molecules and solid-state systems from first-principles, or “ab-initio”, while still accounting for the effects of inter-electron exchange and correlation interactions, provided an incentive for P. Hohenberg and W. Kohn to conceive the current modern formulation of DFT in 1964². In this pioneering work, the authors demonstrated that the density $n(\mathbf{r})$ of particles in the ground state of a quantum many-body system can be assigned a special role and considered as a

¹For an introduction to the mathematical formalism of Hartree-Fock theory, see for example page 332 in Ref. ².

“fundamental” variable, and that all properties of the system under investigation, including the effects of interactions and correlations among the particles, can be considered to be unique *functionals*¹ of this fundamental quantity. It follows from this crucial conjecture that the ground state electron density can be used as an effective replacement for the ground state many-body wavefunction $\Psi_G(r_1, \sigma_1; r_2, \sigma_2; \dots r_N, \sigma_N)$ appearing in Eq. 2.169, thus significantly simplifying the scale and complexity of the problem from the full $3N$ degrees of freedom for N electrons embodied by Ψ_G . This change of variables can be expressed conceptually as follows:

$$\Psi_G = \Psi(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_i, \sigma_i; \dots; \mathbf{r}_N, \sigma_N) \xrightarrow{DFT} n(\mathbf{r}) = \left\langle \Psi_G \left| \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \Psi_G \right. \right\rangle \quad (2.170)$$

This reformulation of the many-body problem is encapsulated by two cardinal theorems, known as Hohenberg-Kohn (HK) theorems, which together establish the theoretical foundations of all modern formulations of DFT. Apart from encompassing the important role assigned to the ground state density in DFT mentioned previously, the HK theorems also provide insight in the relation between the ground state density of a system of mutually-interacting electrons and any general external potential $V_{ext}(\mathbf{r})$ acting upon it (for example, the Coulomb potential due to the nuclei in the periodic crystal structure of a material). The HK theorems, which will be stated here without proof², can be expressed as follows:

1. **Theorem 1:** For any system of interacting particles bathed in an external

¹Familiarity with the mathematical concept of Functional is assumed in this report. For an introduction to the theory of Functionals and their applications in Calculus of Variations, the reader is referred to Chapt. 22 in Ref ?.

²For a formal proof of the HK theorems, consult for example section 6.2 in Ref. ?.

potential $V_{ext}(\mathbf{r})$, the potential $V_{ext}(\mathbf{r})$ is determined uniquely, except for a constant, by the ground state particle density $n_o(\mathbf{r})$.

Corollary 1: Since the hamiltonian of the system is thus fully determined, except for a constant shift of the energy, it follows that the many-body wavefunctions for all states (ground and excited), and consequently all internal properties of the system, are completely determined from the solution of the corresponding Schroedinger Equation given only the ground state density.

2. **Theorem 2:** Any internal property of the system, including its total energy, can be expressed as a unique functional of the density $n(\mathbf{r})$, valid for any external potential $V_{ext}(\mathbf{r})$. For any particular $V_{ext}(\mathbf{r})$ the exact ground state energy of the system is the global minimum value of the functional of the total energy $E[n]$, and the density $n(\mathbf{r})$ that minimizes the functional is the exact ground state density $n_o(\mathbf{r})$.

Corollary 2: The functional $E[n]$ alone is sufficient to determine only the exact ground state energy and density. In general, excited states of the electrons must be determined by other means.

The logic behind the HK theorems is summarized schematically in Fig 2.23. The challenge posed by the Hohenberg-Kohn theorems is how to make use of the reformulation of many-body theory in terms of functionals of the density. Such a functional for the total ground-state energy of the system can be written in general as:

$$E_N[n(\mathbf{r})] = T_S[n(\mathbf{r})] + E_{Hartree}[n(\mathbf{r})] + E_{XC}[n(\mathbf{r})] + \int d^3\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r}) + E_{II} \quad (2.171)$$

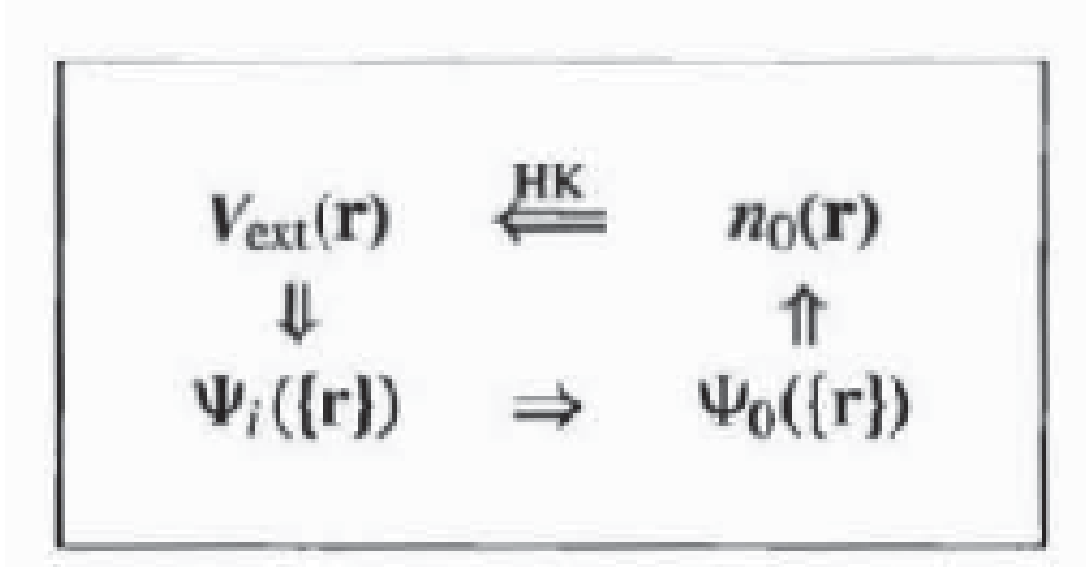


Figure 2.23: Flow diagram illustrating the one-to-one correspondence between the external potential $V_{\text{ext}}(\mathbf{r})$ acting on the system of interacting electrons, its ground state density $n_0(\mathbf{r})$ and its set of eigenstates $\Psi_i(\{\mathbf{r}\})$, as demonstrated by the first HK theorem. (Figure reproduced from Ref. ?)

where E_{II} is the interaction energy of the nuclei among themselves, and $T[n]$, $E_{\text{Hartree}}[n]$ and $E_{XC}[n]$ represent respectively the kinetic, classical Coulomb interaction (Hartree) and Exchange-Correlation energies of the electron cloud, all of which are internal properties of the electron system and are therefore expressible as functionals of the density according to the second HK theorem. At this stage however we are still left with the problem of finding explicit forms for the constituent functionals in Eq. 2.171. While on one hand there exists a simple expression linking the Hartree potential energy to the density:

$$E_{\text{Hartree}}[n] = \frac{1}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} \quad (2.172)$$

this task is more problematic for the Kinetic energy term $T[n]$, for which there

is no known direct connection to the density, and for the Exchange-Correlation energy functional $E_{XC}[n] = E_X[n] + E_C[n]$ incorporating all the complex many-body exchange ($E_X[n]$) and correlation ($E_C[n]$) effects, which cannot be modeled analytically.

A significant step forward from this impasse was accomplished by the seminal work performed in 1965 by W. Kohn and L.J. Sham [?], which lead to the now well-established Kohn-Sham (KS) approach to DFT . The defining *Ansatz* of the KS formulation of DFT was to map the difficult interacting many-body system obeying the Hamiltonian of Eq. 2.169 to a different fictitious auxiliary system of non-interacting particles that can be solved more easily, with the condition that its ground state density be equal to that of the original interacting system and consequently that its internal physical properties be equivalent. This leads to the relation between the actual and auxiliary systems depicted in Fig. 2.24. Even though it has never been proven formally that this auxiliary system is capable of reproducing precisely all the physical properties of the original interacting problem, this basic assumption yields excellent approximations for the problems of greatest practical interest, and in particular for ground-state calculations which are by far the most widespread applications of DFT. Solution of the KS auxiliary system for the ground state can be viewed as the problem of minimization of the total ground state energy functional of Eq. 2.171 with respect to small variations in the density $\delta(n(\mathbf{r}, \sigma))$. One of the major qualities of the KS approach is that it allows the electronic density of the indepent-particle auxiliary system to be expressed as a simple sum of the probability densities associated with each single-

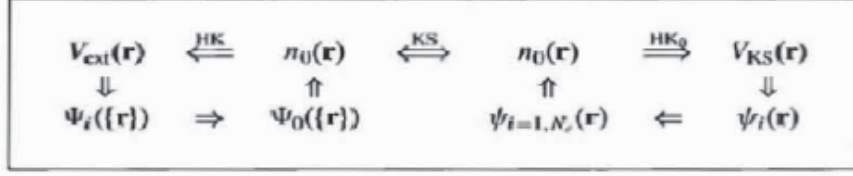


Figure 2.24: Schematic representation of the KS ansatz. The label “ HK_o ” denotes the HK theorem applied to the non-interacting auxiliary system of electrons. The connection in both directions between the two systems is established by the arrow labeled “ KS ”, showing that in principle solution of the independent-particle KS problem determines all properties of the original interacting system. (Figure reproduced from Ref. ?)

particle wavefunction $\psi_i^\sigma(\mathbf{r})$:

$$n(\mathbf{r}) = \sum_{\sigma} n(\mathbf{r}, \sigma) = \sum_{\sigma} \sum_{i=1}^{N^\sigma} |\psi_i^\sigma(\mathbf{r})|^2 \quad (2.173)$$

The establishment of such a relation between density and wavefunctions allows in turn the functional for the kinetic energy in Eq. 2.171 to be kept expressed in terms of the single-particle orbitals, without therefore the need to find an explicit dependence on the density:

$$T_s = -\frac{\hbar^2}{2m_e} \sum_{\sigma} \sum_{i=1}^{N^\sigma} \langle \psi_i^\sigma | \nabla^2 \psi_i^\sigma \rangle = -\frac{\hbar^2}{2m_e} \sum_{\sigma} \sum_{i=1}^{N^\sigma} |\nabla \psi_i^\sigma|^2 \quad (2.174)$$

Given that all other functionals of the density in Eq. 2.171 can now also be expressed in terms of the independent-particle orbitals $\psi_i^\sigma(\mathbf{r})$ via Eq. 2.173, the problem of minimizing Eq. 2.171 to find the ground-state energy of the system can be reformulated in terms of variations of the orbitals via the chain rule, which

yields the following variational equation:

$$\frac{\delta E_N}{\delta \psi_i^{\sigma*}(\mathbf{r})} = \frac{\delta T_S}{\delta \psi_i^{\sigma*}(\mathbf{r})} + \left[\frac{\delta E_{ext}}{\delta n(\mathbf{r}, \sigma)} + \frac{\delta E_{Hartree}}{\delta n(\mathbf{r}, \sigma)} + \frac{\delta E_{XC}}{\delta n(\mathbf{r}, \sigma)} \right] \frac{\delta n(\mathbf{r}, \sigma)}{\delta \psi_i^{\sigma*}(\mathbf{r})} = 0 \quad (2.175)$$

subject to the orthonormalization constraint on each of the single-particle orbitals:

$$\langle \psi_i^\sigma | \psi_j^{\sigma'} \rangle = \delta_{i,j} \delta_{\sigma,\sigma'} \quad (2.176)$$

Using expressions 2.174 and 2.173 for calculating the derivatives involving the Kinetic Energy and the density:

$$\frac{\delta T_S}{\delta \psi_i^{\sigma*}(\mathbf{r})} = -\frac{1}{2} \nabla^2 \psi_i^\sigma(\mathbf{r}); \quad \frac{\delta n^\sigma(\mathbf{r})}{\delta \psi_i^{\sigma*}(\mathbf{r})} = \psi_i^\sigma(\mathbf{r}) \quad (2.177)$$

and the Lagrange multiplier method for constrained minimization¹, we obtain the following set of Schroedinger-like equations, one for each of the N single-particle orbitals in the auxiliary system:

$$(H_{KS}^\sigma - \epsilon_i^\sigma) \psi_i^\sigma(\mathbf{r}) = 0 \quad (2.178)$$

$$\{\psi_i^\sigma : i = 1, N\}$$

where the Lagrange parameters ϵ_i represent the possible eigenvalues. It is important to stress that these eigenvalues do not correspond to the single-particle energies as in a normal Schroedinger Equation, and in fact have no obvious physical meaning except for the highest one, which corresponds to the ionization energy

¹An introduction to this mathematical technique is provided in section 5.9 in Ref. ?

of the system. H_{KS} on the other hand represents the effective KS Hamiltonian:

$$H_{KS}^\sigma(\mathbf{r}) = -\frac{\hbar^2}{2m_e}\nabla^2 + V_{KS}^\sigma(\mathbf{r})$$

where $V_{KS}^\sigma(\mathbf{r}) = \frac{\delta E_{ext}}{\delta n(\mathbf{r}, \sigma)} + \frac{\delta E_{Hartree}}{\delta n(\mathbf{r}, \sigma)} + \frac{\delta E_{XC}}{\delta n(\mathbf{r}, \sigma)} = V_{ext}^\sigma(\mathbf{r}) + V_{Hartree}^\sigma(\mathbf{r}) + V_{XC}^\sigma(\mathbf{r})$

(2.179)

Equations 2.178 and 2.179, collectively referred to as the *Kohn-Sham equations*, represent the culminating point of the entire KS formulation of DFT. By diagonalizing and solving the equations self-consistently (in practice by numerical means) one can first compute all the single-particle KS eigenstates $\psi_i^\sigma(\mathbf{r})$, and from there the ground-state density and total ground-state energy of the original interacting system from Equations 2.173 and 2.171 respectively, with an accuracy limited only by the approximations in the exchange-correlation functional (given that all other functionals in Eq. 2.171 can be computed exactly within the KS approach) . In fact, the development and availability today of highly accurate exchange-correlation functionals with a broad range of applications has been the determining factor for elevating DFT to its current position as the most accurate, computationally-efficient and widely used technique for ab-initio electronic structure calculations in a wide range of atoms, molecules and condensed matter systems. The next section in this document is devoted to the description of the approximations for this functional most entrenched in the Condensed-Matter Physics community.

2.6.2 The Local Density and Generalized Gradient approximations for the exchange-correlation energy functional

Already in their seminal paper ?, Kohn and Sham pointed out that solids can often be considered as close to the limit of the homogeneous electron gas. In this limit, it is known that the effects of exchange and correlation are rather local in character, and they proposed making the local density approximation (LDA) (or more generally the local spin density approximation (LSDA), which accounts also for the spins of the electrons in a spin-polarized system) in which the exchange-correlation energy is simply an integral over all space, with the exchange-correlation energy density $\epsilon_{XC}([n], \mathbf{r})$ at each point \mathbf{r} assumed to be the same as in a homogeneous electron gas of interacting electrons with the same local density:

$$E_{XC}^{LDA}[n] = \int d^3\mathbf{r} n(\mathbf{r}) \epsilon_{XC}^{\text{hom}}(n(\mathbf{r})) = \int d^3\mathbf{r} n(\mathbf{r}) [\epsilon_X^{\text{hom}}(n(\mathbf{r})) + \epsilon_C^{\text{hom}}(n(\mathbf{r}))] \quad (2.180)$$

The corresponding exchange-correlation potential appearing in the KS equations is consequently given by:

$$V_{XC}^{\sigma}(\mathbf{r}) = \frac{\delta E_{XC}^{LDA}[n]}{\delta n(\mathbf{r}, \sigma)} = \left[\epsilon_{XC}^{\text{hom}} + n \frac{\partial \epsilon_{XC}^{\text{hom}}}{\partial n^{\sigma}} \right]_{\mathbf{r}, \sigma} \quad (2.181)$$

The only information needed within the LDA approximation is therefore the exchange-correlation energy of the homogeneous gas as a function of density. Since the exchange energy of an homogeneous electron gas can be calculated

analytically within Hartree-Fock theory, and is given by ¹:

$$\epsilon_X^{\text{hom}} = -\frac{3}{4} \frac{e^2 k_F}{\pi} = -\frac{3e^2}{4\pi a_o} (k_F a_o) \quad (2.182)$$

where k_F is the Fermi wavevector and a_o the Bohr Radius, the LDA simply boils down to fitting numerical correlation energies for the homogeneous gas, for example by Monte Carlo total energy calculations ^{??}. A variety of LDA expressions for the correlation energy have been proposed with time, and the most celebrated are summarized in Appendix B of Ref. [?]. The LDA is expected to give the best results for solids close to a homogeneous electron gas in which the charge density is slowly varying (like a nearly-free-electron metal) and worst for very inhomogeneous cases, like for the case of atoms where the density must go continuously to zero outside the atom, or indeed in any general covalently-bound solid. However, experience has proved the LDA to be a surprisingly good approximation for a wide variety of solid-state systems.

The remarkable success of the LDA in most applications has stimulated ideas for the development of various Generalized Gradient Approximations (GGAs) ^{???}, with marked improvement over LDA in many cases. The first step beyond the local approximation was the introduction of a functional of the magnitude of the density gradient $|\nabla n^\sigma(\mathbf{r})|$ in addition to the value of $n(\mathbf{r})$ at each point \mathbf{r} , which lead to the so-called Gradient Expansion Approximation (GEA). However, the GEA was found to often yield worse results than the LDA due to the large density gradients found in real materials, which cause the expansion to break down. The GGA provided the solution by introducing functions that modify the

¹A derivation of this expression can be found in Chapt. 17 of Ref. [?].

behavior at large gradients in such a way as to preserve the desired properties of the material under investigation. Within the the context of the GGA, it is convenient to generalize the expression for the Exchange-correlation functional given by Eq. 2.180 to include explicitly the dependence on the various orders of the density gradient $|\nabla^m n^\sigma(\mathbf{r})|$:

$$\begin{aligned} E_{XC}^{GGA}[n] &= \int d^3r n(r) \varepsilon_{XC}(n, |\nabla n|, |\nabla^2 n|, \dots, |\nabla^m n|) \\ &\equiv \int d^3r n(r) \varepsilon_X^{\text{hom}}(n) F_{XC}(n, |\nabla n|, |\nabla^2 n|, \dots, |\nabla^m n|) \end{aligned} \quad (2.183)$$

where F_{XC} is a dimensionless function, known as the enhancement factor of the exchange-correlation functional, and $\epsilon_x^{\text{hom}}(n)$ is the exchange energy density of the homogeneous electron gas given by Eq. 2.182. An expression for the exchange-correlation potential appearing in the KS equations can also be found by calculating the change $\delta E_{XC}[n]$ to linear order in δn and $\delta \nabla n = \nabla \delta n$:

$$V_{XC}^\sigma(\mathbf{r}) = \frac{\delta E_{XC}^{GGA}[n]}{\delta n(\mathbf{r}, \sigma)} = \left[\varepsilon_{XC} + n \frac{\partial \varepsilon_{XC}}{\partial n^\sigma} - \nabla \left(n \frac{\partial \varepsilon_{XC}}{\partial \nabla n^\sigma} \right) \right]_{\mathbf{r}, \sigma} \quad (2.184)$$

It is natural to work in terms of dimensionless reduced density gradients of m^{th} order that can be defined by:

$$s_m = \frac{|\nabla^m n|}{(2k_F)^m n} \quad (2.185)$$

With this notation, the lowest order terms in the expansion of the exchange part F_X and correlation part F_C of F_{XC} have been calculated analytically:

$$F_X = 1 + \frac{10}{81} s_1^2 + \frac{146}{2025} s_2^2 + \dots \quad (2.186)$$

and

$$F_C = \frac{\varepsilon_C^{LDA}(n)}{\varepsilon_X^{LDA}(n)} (1 - 0.219s_1^2 + \dots) \quad (2.187)$$

where F_C is expressed as a correction to the corresponding LDA expressions for the exchange and correlation energy densities. One of the defining characteristics of GGAs is that they lead to an exchange energy lower than their LDA counterparts, since $F_x \geq 1$. The resulting reduction of the cohesive energy improves markedly the agreement with experiment for atoms, molecules and solids, and therefore constitutes a significant step forward over the LDA overbinding ???. Numerous different forms have been proposed for the higher-order terms in the expansions for F_C and F_X by choosing different physical conditions for $s \rightarrow \infty$, leading to very different behaviors in the region of large density gradients. Hence, even if one form of GGA somehow gives the correct result for a certain physical property while others fail, the same form is not necessarily superior for other properties in which different physical conditions prevail.

2.6.3 Computational self-consistent solution of the Kohn-Sham equations: the plane-wave pseudopotential method

In order to solve in practice the eigenvalue problem presented by the KS equations (2.178), the single-particle eigenstates $\psi_i^\sigma(\mathbf{r})$ must be expanded in some orthogonal basis set. The simplest and most popular choice for this basis is to use orthogonal plane-wave functions, under the so-called orthogonalized planewave

(OPW) method ??? . These plane-waves can be expressed as ¹:

$$\psi_{\mathbf{k}}^i(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{k}}^n(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \quad (2.188)$$

where \mathbf{k} is the crystal momentum and the sum is over all Bravais lattice vectors \mathbf{G} of the reciprocal lattice of the system under investigation. This discrete basis set can in theory be made infinite in size by application of periodic boundary conditions to the problem, but in practice a finite number of plane-wave basis states is normally sufficient to converge systematically most physical quantities of interest in typical computational problems. As explained in Fig. 2.25, in order to truncate the basis set, the sum term in Eq. (2.188) is limited to a set of reciprocal lattice vectors encompassed by a sphere with radius defined by the cutoff kinetic energy, E_{cut} :

$$\frac{\hbar^2 |\mathbf{k}^2 + \mathbf{G}^2|}{2m} \leq E_{cut} \quad (2.189)$$

so that the entire set of plane-waves can be defined by this maximum kinetic energy component. More recently, the OPW method has been reformulated and adapted to modern techniques for calculation of total energy, forces and stress under the so-called Projector-Augmented Wave (PAW) method ?, which introduces projectors and auxiliary localized functions for a more efficient solution of the KS eigenvalue problem.

The choice of plane waves as basis set however requires to be operated in conjunction with the implementation of pseudopotentials ??? in order to limit the

¹This form for the plane-wave functions can be shown to be the most general solution satisfying Bloch's Theorem for periodic boundary conditions in a crystal lattice.

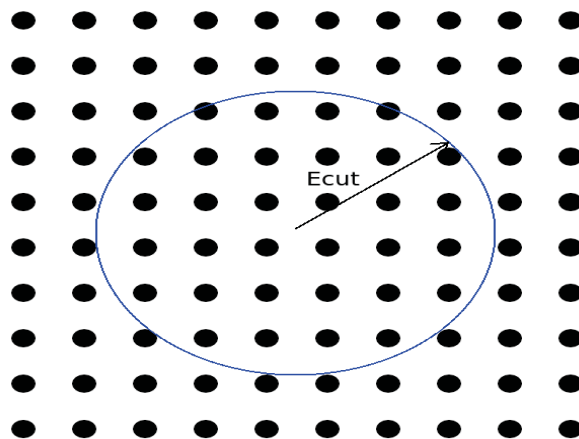


Figure 2.25: Truncation of the plane wave basis states expansion at a sphere of radius E_{cut} in reciprocal space. The blue dots represent the discrete reciprocal lattice vectors available for the expansion. (Figure reproduced from Ref. ?.)

size of this expansion to computationally-feasible proportions for the solution of realistic systems. Pseudopotentials allow for a neat separation between the treatment of core electrons in an atom, that is those electrons which are tightly-bound to the nucleus and are therefore not involved in the interactions between atoms, and the valence electrons, which on the other hand are more loosely bound and consequently dictate the chemical properties of the material. As a direct consequence of the orthogonality requirement on the single-electron wavefunctions, the superposition of the valence and core electron wavefunctions results in a many-electron wavefunction exhibiting strong-oscillations in the region near the nucleus, as illustrated in Fig. 2.26. These oscillations become inevitably quite difficult to model using a plane-wave basis set, requiring many plane-waves for an accurate description. It therefore becomes convenient to ignore altogether the region in the immediate surroundings of the nucleus, and this is indeed the approach adopted by the pseudopotential approximation, where only the valence

electrons of atoms are explicitly considered and the screening effects of the inner core electrons contained within the cutoff nuclear radius r_c indicated in Fig. 2.26 are integrated within a new effective ionic potential (the pseudopotential). The remaining smoother variation of the all-electron valence wavefunction in the outskirts of the nuclear region beyond r_c can be modeled accurately with a much more restricted size for plane-wave basis set, with the consequent sharp gain in computational efficiency. The pseudopotential approach turns out to be an excellent approximation for practical calculations since, as mentioned previously, this region remains almost completely shielded from all neighboring interactions, and the oscillations are consequently of very little consequence for the electronic structure of the solid.

When constructing a new pseudopotential, the main goals that need to be attained are threefold: firstly, the pseudopotential should be as soft as possible, meaning that it should allow expansion of the valence pseudo-wavefunctions using as few plane-waves as possible. Secondly, it should be as transferrable as possible, in the sense that a pseudopotential generated for a given atomic configuration should accurately reproduce others. This helps to ensure that its application in solid-state systems, where the overall crystal potential is inevitably different from an atomic potential, be capable of reliable results. Finally, the pseudo-charge density, that is the charge density constructed using the pseudo-wavefunctions, should reproduce the valence charge density of the atom under consideration as accurately as possible. The concept of *norm-conservation* in pseudopotential theory ? provided the first systematic solution for reconciling these conflicting goals. Within norm-conserving pseudopotentials, the pseudo-wavefunctions are designed and constructed to be equivalent to the actual valence wavefunctions

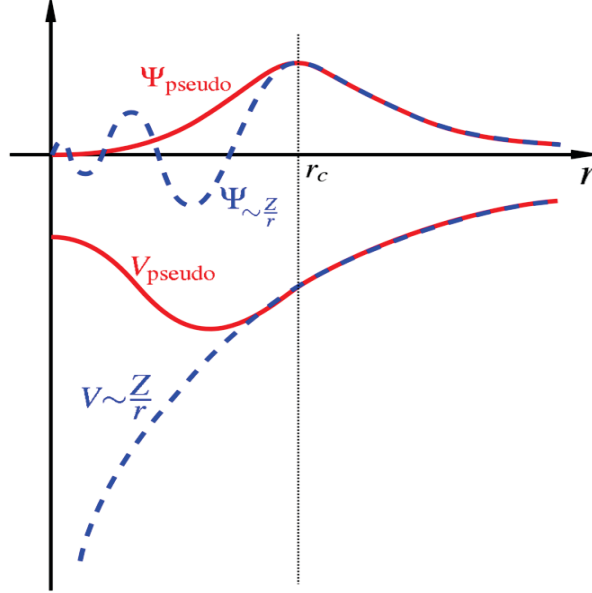


Figure 2.26: Schematic representation of the concept of pseudopotential. Ψ labels the true all-electron wavefunction, while Ψ_{pseudo} the all-electron pseudowavefunction which results from substituting the true bare Coulombic nuclear potential Z/r with the modified pseudopotential V_{pseudo} incorporating the screening effects of the inner core electrons. The smooth variation of Ψ beyond the core radius r_c implies that an accurate emulation of the atomic interaction with Ψ_{pseudo} requires only a limited plane-wave basis expansion, thus making the pseudopotential a very efficient approximation for most atomic systems. (Figure reproduced from Ref. ?.)

outside the core radius r_c . Inside r_c on the other hand, the pseudo-wavefunctions are allowed to differ from the true wavefunctions, but their norm is constrained to be the same.

A radical departure from the concept of norm-conservation in pseudopotentials was eventually proposed by Vanderbilt and co-workers [1]. In this approach, which became to be known as the *ultra-soft* pseudopotential method, the pseudo-wavefunctions are required to be equal to the all-electron wavefunctions outside r_c as with norm-conserving pseudopotentials, but inside r_c they are allowed to

be as soft as possible, which is accomplished by removing the norm-conservation constraint. Even though the use of non-normalized wavefunctions complicates the solution of the KS equations, this approach offers the major advantage of greatly reducing the required plane-wave cutoff without sacrificing the accuracy of the calculation, which explains its popularity for use in large-scale calculations. The computational work presented in this report was no exception.

One final aspect of the computational implementation of DFT worth mentioning is the techniques used for determining the charge density and other quantities which can be extracted from DFT calculations such as total energy and atomic forces. The evaluation of all these quantities require summations to be performed over the occupied states, which for crystals translate into integrals over the Brillouin zone. Exploiting the symmetry properties of the crystal, this integration can then be limited to the symmetry-irreducible wedge of the zone (IBZ). In practice, these integrals are calculated numerically using wavefunctions and eigenvalues at a finite number of \mathbf{k} -points in the electronic bandstructure of the IBZ. This leads to the so-called *special points method* for Brillouin-zone integration ???, the most commonly used approach for this type of calculations, in which integrations are performed as weighted sums over a grid of representative, or special, \mathbf{k} -points chosen to yield optimum convergence for smoothly varying functions of \mathbf{k} . This application of the special points method however proves problematic for the case of metals, where electronic bands intersect the Fermi energy E_f . This leads to discontinuities in the electronic occupation numbers and therefore in the integrands at the Fermi surface, thus yielding slower convergence for a given number of \mathbf{k} -points. This difficulty can be overcome by the introduction of an artificial temperature-induced and convergence-accelerating broadening of the Fermi sur-

face corresponding to a smoother electron distribution function than the original step function, such as a finite-temperature Fermi distribution. Care must however be taken to ensure that the introduction of such broadening factor does not affect physical quantities of interest.

Once equipped with all the above-mentioned computational models and techniques, a solution to the set of KS equations minimizing the total energy functional of Eq. (2.171) can be found numerically according to the iterative scheme portrayed in Fig. 2.27. Any valid solution must be *self-consistent*, meaning that it must be subject to the constraint that the effective KS potential $V_{KS}^\sigma(\mathbf{r})$ and the resulting electron density $n(\mathbf{r}, \sigma)$ be consistent with one another, as highlighted in the figure. In practice this is done by successively looping over changes in $V_{KS}^\sigma(\mathbf{r})$ and $n(\mathbf{r}, \sigma)$ until the self-consistent agreement is reached.

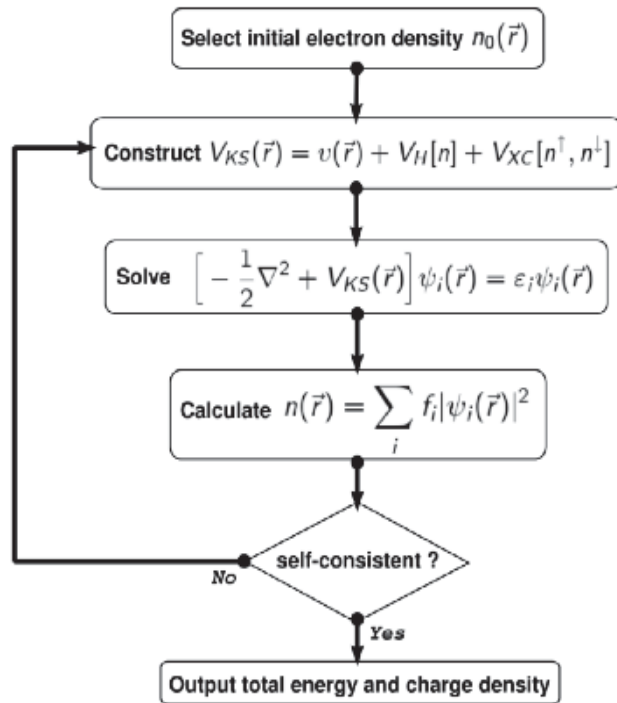


Figure 2.27: Flow chart illustrating the self-consistent iteration for solving the KS equations. (Figure reproduced from Ref. ?)