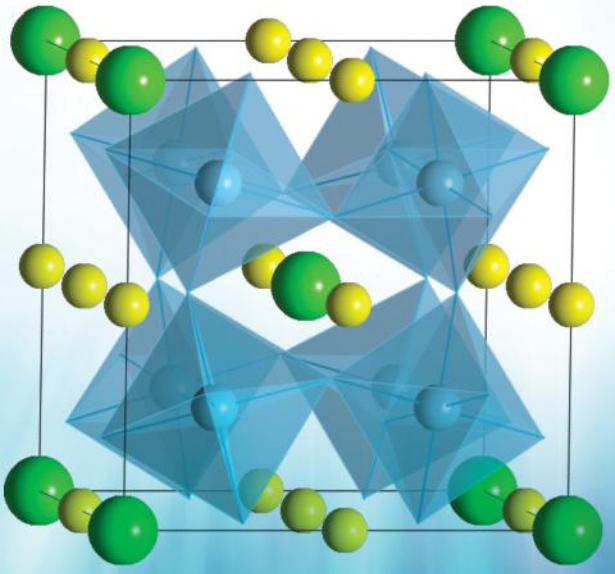
SOLID STATE CHEMISTRY

AND ITS APPLICATIONS

SECOND EDITION STUDENT EDITION



ANTHONY R. WEST

WILEY

IUPAC Periodic Table of the Elements

1																	18
1 H hydrogen [1.007; 1.009]	2		Key: atomic number **Symbol** name standard atomic weight									13	14	15	16	17	2 He helium 4.003
3 Li lithium [6.938; 6.997]	4 Be beryllium 9.012											5 B boron [10.80; 10.83]	6 C carbon [12.00; 12.02]	7 N nitrogen [14.00; 14.01]	8 O oxygen [15.99; 16.00]	9 F fluorine 19.00	10 Ne neon 20.18
11 Na sodium 22.99	12 Mg magnesium 24.31	3	4	5	6	7	8	9	10	11	12	13 Al aluminium 26.98	14 Si silicon [28.08; 28.09]	15 P phosphorus 30.97	16 S sulfur [32.05; 32.08]	17 Cl chlorine [35.44; 35.46]	18 Ar argon 39.95
19 K potassium 39.10	20 Ca calcium 40.08	21 Sc scandium 44.96	22 Ti titanium 47.87	23 V vanadium 50.94	24 Cr chromium 52.00	25 Mn manganese 54.94	26 Fe iron 55.85	27 Co cobalt 58.93	28 Ni nickel 58.69	29 Cu copper 63.55	30 Zn zinc 65.38(2)	31 Ga gallium 69.72	32 Ge germanium 72.63	33 As arsenic 74.92	34 Se selenium 78.96(3)	35 Br bromine 79.90	36 Kr krypton 83.80
37 Rb rubidium 85.47	38 Sr strontium 87.62	39 Y yttrium 88.91	40 Zr zirconium 91.22	41 Nb niobium 92.91	42 Mo molybdenum 95.96(2)	43 Tc technetium	44 Ru ruthenium 101.1	45 Rh rhodium 102.9	46 Pd palladium 106.4	47 Ag silver 107.9	48 Cd cadmium 112.4	49 In indium 114.8	50 Sn tin 118.7	51 Sb antimony 121.8	52 Te tellurium 127.6	53 I iodine 126.9	54 Xe xenon 131.3
55 Cs caesium 132.9	56 Ba barium 137.3	57-71 lanthanoids	72 Hf hafnium 178.5	73 Ta tantalum 180.9	74 W tungsten 183.8	75 Re rhenium 186.2	76 Os osmium 190.2	77 Ir iridium 192.2	78 Pt platinum 195.1	79 Au gold 197.0	80 Hg mercury 200.6	81 Tl thallium [204.3; 204.4]	82 Pb lead 207.2	83 Bi bismuth 209.0	84 Po polonium	85 At astatine	86 Rn radon
87 Fr francium	88 Ra radium	89-103 actinoids	104 Rf rutherfordium	105 Db dubnium	106 Sg seaborgium	107 Bh bohrium	108 Hs hassium	109 Mt meitnerium	110 Ds darmstadtium	111 Rg roentgenium	112 Cn copernicium		114 Fl flerovium		116 Lv livermorium		

57 La lanthanum 138.9	58 Ce cerium 140.1	59 Pr praseodymium 140.9	60 Nd neodymium 144.2	61 Pm promethium	62 Sm samarium 150.4	63 Eu europium 152.0	64 Gd gadolinium 157.3	65 Tb terbium 158.9	66 Dy dysprosium 162.5	67 Ho holmium 164.9	68 Er erbium 167.3	69 Tm thulium 168.9	70 Yb ytterbium 173.1	71 Lu lutetium 175.0
89 Ac actinium	90 Th thorium 232.0	91 Pa protactinium 231.0	92 U uranium 238.0	93 Np neptunium	94 Pu plutonium	95 Am americium	96 Cm curium	97 Bk berkelium	98 Cf californium	99 Es einsteinium	100 Fm fermium	101 Md mendelevium	102 No nobelium	103 Lr lawrencium

INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY

Notes

- IUPAC 2009 Standard atomic weights abridged to four significant digits (Table 4 published in *Pure Appl. Chem.* 83, 359-396 (2011); doi:10.1351/PAC-REP-10-09-14). The uncertainty in the last digit of the standard atomic weight value is listed in parentheses following the value. In the absence of parentheses, the uncertainty is one in that last digit. An interval in square brackets provides the lower and upper bounds of the standard atomic weight for that element. No values are listed for elements which lack isotopes with a characteristic isotopic abundance in natural terrestrial samples. See PAC for more details.

- "Aluminum" and "cesium"are commonly used alternative spellings for "aluminium"and "caesium."

- Claims for the discovery of all the remaining elements in the last row of the Table, namely elements with atomic numbers 113,115,117 and 118, and for which no assignments have yet been made, are being considered by a IUPAC and IUPAP Joint Working Party.

For updates to this table, see iupac.org/reports/periodic_table/. This version is dated 1 June 2012.

Solid State Chemistry and its Applications

Second Edition

Solid State Chemistry and its Applications

Second Edition

Student Edition

ANTHONY R. WEST
Department of Materials Science and Engineering,
University of Sheffield, UK

WILEY

This edition first published 2014

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data

West, Anthony R.
Solid state chemistry and its applications / Anthony R. West. – Second edition, student edition.
pages cm
Includes index.
ISBN 978-1-119-94294-8 (pbk.)
1. Solid state chemistry. I. Title.
QD478.W47 2014
541′.0421–dc23

2013029528

A catalogue record for this book is available from the British Library.

ISBN: 9781119942948

Cover images created using CrystalMaker® software. CrystalMaker Software Ltd, www.crystalmaker.com

Set in 10/12pt Times by Aptara Inc., New Delhi, India

1 2014

For Sheena, Isla, Graeme, Jenny and Susie

Contents

Preface

This book is a slimmed down, student edition of 'Solid State Chemistry and its Applications' whose second edition is scheduled for publication in 2015/6. It is modelled on the authors' 'Basic Solid State Chemistry', but has been completely rewritten with about 40% new material added and all the diagrams drawn professionally, in full colour. The nine chapters in 'Basic' have become ten in this new edition since 'Magnetic and Optical Properties' is split into separate chapters.

In the period since the second edition of 'Basic' was published in 1999, we have witnessed many major new discoveries and developments in the solid state chemistry of inorganic materials with topics such as colossal magnetoresistance, multiferroics, light emitting diodes and graphene. New materials synthesis techniques have evolved such as mechanosynthesis, microwave-hydrothermal synthesis and atomic layer deposition and of course, there have been many improvements in the techniques used to characterise solids including use of synchrotrons for diffraction and spectroscopy as well as high resolution scanning transmission electron microscopy permitting atomic-level identification and structural imaging. It was felt that an updated version of both 'Basic' and 'Solid State Chemistry and its Applications' was long overdue, therefore.

A major feature of this new edition is the extensive coverage of the crystal structures of important families of inorganic solids. Purchasers of the book will be able to download, free, a bespoke and easy-to-use CrystalMaker® viewer program. The CrystalViewer software is accompanied by more than 100 crystal structure models which users will be able to view on their computers with the facility to rotate the structures, view them from different orientations and either highlight or hide different structural features. CrystalViewer and the accompanying structure files can be downloaded from the companion website at http://www.wiley.com/go/west/solidstatechemistrystudent.

Many people have helped and encouraged me in preparing this new edition. Special thanks are due to: John McCallum who produced many of the crystal structure drawings and files, Frances Kirk who prepared the whole manuscript, in electronic format, and Wiley staff Sarah Hall and Sarah Tilley for their enthusiastic encouragement and involvement: in particular, Sarah Hall was instrumental in making the CrystalMaker® arrangements and Sarah Tilley oversaw all the artwork preparations.

Anthony R. West
Sheffield
July 2013

Chemistry – Solid State Chemistry – Materials Chemistry – Materials Science and Engineering

Chemistry is an evolving subject! Traditionally, there have been three branches of chemistry: organic, physical and inorganic, with some arguments in favour of including analytical as a fourth branch. An alternative, fairly new classification (favoured by the author!) divides chemistry into two broad areas: molecular (which includes liquids and gases) and non-molecular (or solid state). The ways in which we think about, make, analyse and use molecular and non-molecular substances are completely different, as shown by a comparison of one 'simple' substance in each category, toluene and aluminium oxide:

Comparison of the chemistries of molecular and non-molecular materials

Characteristic	Toluene	Aluminium oxide
Formula	Fixed, $C_6H_5CH_3$	Usually fixed, Al_2O_3, but for other oxides may be variable, e.g. $Fe_{1-x}O$
Are defects present?	Not allowed: missing or mis-placed atoms give rise to different molecules	Unavoidable: small concentration of vacancies, interstitials and dislocations are always present
Doping possibilities	Not possible without producing a different molecule	Doping or solid solution formation allows control and optimisation of properties, e.g. ruby is Cr-doped Al_2O_3
Structure and its determination	Molecular structure can be determined spectroscopically: NMR/Mass Spec/IR. Determine packing arrangement, bond lengths and angles, by single crystal X-ray diffraction. Usually, structural information is then complete.	Full characterisation of a solid requires structural and compositional information across the length scales from local, to unit cell, nano and microscales. Many diffraction, spectroscopic and microscopic techniques are needed for full characterisation.
Properties and applications	Controlled by molecular formula and configuration; cannot be modified by doping. Some properties (e.g. pharmaceutical activity) may depend on molecular packing arrangements in crystals.	Properties/applications depend on crystal structure, defects, dopants, surface structure, particle size and whether the material is a powder, single crystal, film, etc. Consider the diverse applications of Al_2O_3: films and ceramics used as insulators; powders used as abrasive; with Cr^{3+} dopants, ruby is used for lasers; porous solids used as catalyst supports.

Thus, for toluene, once its formula and molecular structure had been determined there were few remaining issues to be resolved other than, perhaps, the detailed packing arrangement of molecules in crystalline toluene at low temperatures or the possible discovery and evaluation, even today, of as-yet unknown chemical, biological or pharmaceutical properties of pure toluene.

Alumina, by contrast, is a highly complex material; its properties, and therefore potential applications, depend on different aspects of its structure (bulk, defect, surface, nano), the methods needed to fabricate it in different forms and shapes, the possibility of doping to modify its properties and the characterisation or determination of its structure (and its composition, whether homogeneous or heterogeneous, if doped) across all length scales. This is solid state chemistry!

The biggest contrast between molecular and non-molecular materials is that the latter can be doped, allowing modification and control of properties such as magnetism, superconductivity and colour/optical band gap. By contrast, attempts to dope molecules are inevitably frustrated since replacing one atom in the molecule by another, or creating defects such as missing atoms, lead to entirely different molecules.

In recent decades, *materials chemistry* has emerged as a distinct branch of chemistry which covers both non-molecular, solid state materials (oxides, halides, etc.) and many molecular materials (especially, functional polymers and organic solids with potentially useful physical properties). Materials chemistry cuts across the traditional disciplines of chemistry but also includes something extra which is an interest in the physical properties of compounds and materials. In the past, solid state physics and materials science have been the usual 'home' for physical properties; but now, they are an intrinsic part of solid state and materials chemistry.

The distinction between *materials chemistry* and *materials science* is often unclear but can be summarised broadly as follows:

Materials chemistry

Synthesis – structure determination – physical properties – new materials

Materials science

Processing and fabrication – characterisation – optimisation of properties and testing – improved/new materials for engineering applications in products or devices.

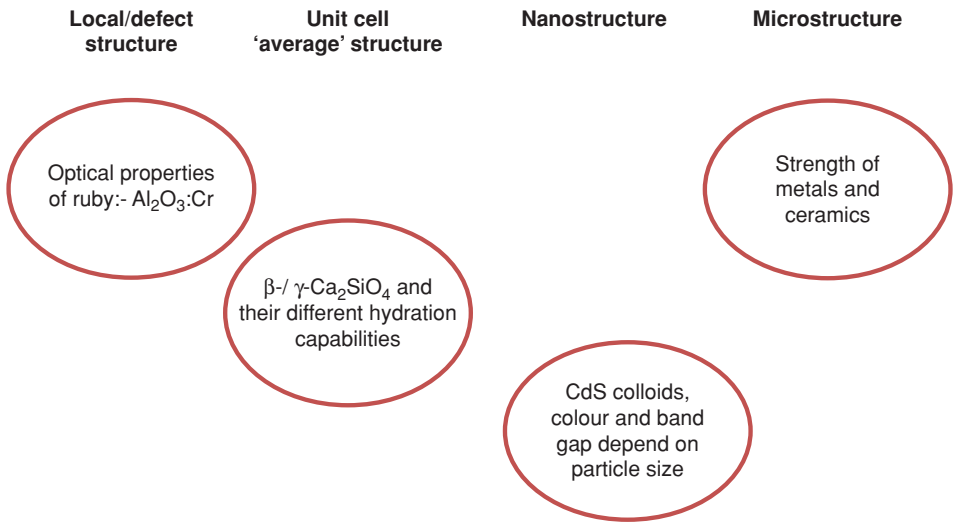
Materials science focuses on materials that are already known to be useful or have the potential to be developed for applications, either by compositional control to optimise properties or by fabrication into desired forms, shapes or products. Materials science therefore includes whatever aspects of chemistry, physics and engineering that are necessary to achieve the desired aims.

Materials chemistry is much more than just a subset of materials science, however, since it is freed from the constraint of a focus on specific applications; materials chemists love to synthesise new materials and measure their properties, some of which may turn out to be useful and contribute to the development of new industries, but they do this within an overarching interest in new chemistry, new structures and improved understanding of structure – composition – property relationships.

A curious fact is that, in the early days of chemistry, inorganic chemistry had as its main focus, the elements of the periodic table and their naturally occurring or easy-to-make compounds such as oxides and halides. Inorganic chemistry subsequently diversified to include organometallic chemistry and coordination chemistry but interestingly, many traditional inorganic materials have returned to centre-stage and are now at the heart of solid state materials science. Examples include: Cr-doped Al_2O_3 for lasers; doped Si semiconductors

for microelectronics; doped ZrO_2 as the solid electrolyte in solid oxide fuel cells; $BaTiO_3$ as the basis of the capacitor industry with a total annual production worldwide exceeding 10^{12} units; copper oxide-based materials for superconductor applications; and many, many more. The scope for developing new solid state materials/applications is infinite, judging by the 'simple' example of Al_2O_3 described above. Most such materials tend not to suffer from problems such as volatilisation, degradation and atmospheric attack, which are often a drawback of molecular materials, and can be used safely in the environment.

It is important to recognise also that physical properties of inorganic solids often depend on structure at different length scales, as shown by the following examples:



Thus in the case of ruby, which is a natural gemstone and was the first material in which LASER action – light amplification by stimulated emission of radiation – was demonstrated, two structural aspects are important. One is the host crystal structure of corundum, α-Al_2O_3 and the other is the Cr^{3+} dopant which substitutes at random for about 1% of the Al^{3+} ions in the corundum lattice: the Cr-O bond lengths and the octahedral site symmetry are controlled by the host structure; the two together combine to give the red ruby colour by means of d–d transitions within the Cr chromophore and the possibility of accessing the long-lived excited states that are necessary for LASER action.

A remarkable example of the effect of crystal structure details at the unit cell scale on properties is shown by dicalcium silicate, Ca_2SiO_4 which is readily prepared in two polymorphic forms at room temperature. One, the β-polymorph, reacts with water to give a semicrystalline calcium silicate hydrate which sets rock-solid and is a main constituent of concrete; the other polymorph, γ-Ca_2SiO_4, does not react with water. Just think, the entire construction industry rests on the detailed polymorphism of dicalcium silicate! It is not sufficient that one of the key components of cement has the right composition, Ca_2SiO_4; in addition, the precise manner in which ions are packed together in the solid state is critical to its hydration properties and whether or not it turns into concrete.

At the nanoscale, crystalline particles may contain many hundreds of unit cells but often their properties are different from powders, ceramics or single crystals of the same material with larger-sized grains simply because of the influence of surface energies. In small nanoparticles, surface free energies and structures

increasingly dominate the total free energy of a material, as shown by the colour, and associated band gap, of CdS nanoparticles (or colloids in older terminology) which can be fine-tuned by controlling the particle size.

Some properties are determined by structure at the micron ($1\ \mu m = 10^3\ nm = 10^4\ Å = 10^{-3}\ mm$) scale and this is the reason why 'microstructure' features strongly in the characterisation of metals and ceramics, primarily using optical and electron microscopy techniques. Frequently, impurities/dopants may precipitate at grain boundaries and surfaces and these can have a dramatic influence on for instance, the mechanical properties.

These examples illustrate the awesome challenges that must be met before an inorganic solid can be regarded as fully characterised across the length scales. This, coupled with the enormous number of inorganic crystal structures that are known, and the possibility to introduce dopants which modify properties, underlines why solid state chemistry is a central subject to many areas of physical science, engineering and technology.

This book concerns solid state chemistry and focuses on inorganic solids: their crystal structures, defect structures and bonding; the methods used to synthesise them and determine their structures; their physical properties and applications. Organic and other molecular materials are included in the coverage if their properties in the solid state complement, or relate to, those of inorganic solids. Physical properties are an intrinsic part of solid state chemistry since the whole area of structure–property relations requires the insights and input of chemistry to synthesise and characterise materials, as well as a good understanding of physical properties and the factors that control them.

Companion Website

This textbook is supported by a website which contains a variety of supplementary resources:

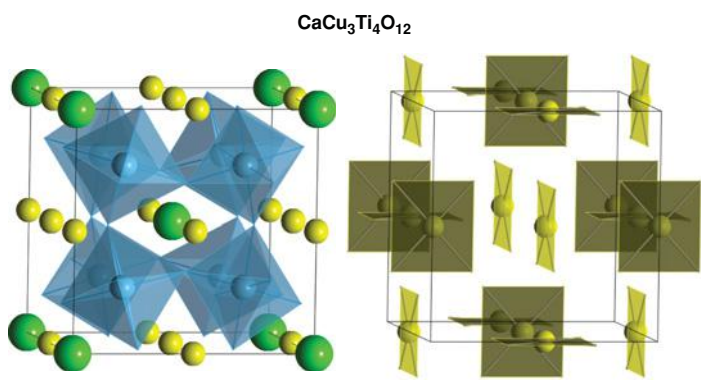
http://www.wiley.com/go/west/solidstatechemistrystudent

Online you will find PowerPoint slides of all figures from the book, as well as solutions to the set of questions. The website also gives you access to a CrystalMaker® viewer program. The CrystalViewer software is available for Windows and Mac, and is accompanied by a broad array of crystal structures for you to view and manipulate.

CrystalViewer

CrystalViewer is a visualisation program for displaying and manipulating crystal structures. The CrystalViewer software facilitates the exploration of crystal structures from the book in three dimensions, allowing users to view the structures in different orientations, and highlight/hide different structural features so as to aid the interpretation of complex crystal structures. The CrystalViewer program is accompanied by over 100 crystal structure files; many of these structures relate directly to illustrations from the book, identified by their figure numbers, and a variety of additional structures are provided to complement the concepts and applications discussed in the text.

The CrystalViewer software and accompanying structure files can be downloaded from the companion website at http://www.wiley.com/go/west/solidstatechemistrystudent



CrystalMaker®CrystalMaker Software Ltd, www.crystalmaker.com

An example of how a crystal structure can appear very different, depending on which aspects are emphasised, is shown here for $CaCu_3Ti_4O_{12}$, in which the two diagrams highlight either the TiO_6 octahedra or the CuO_4 square planar units.

Crystal Structure Library

A Crystal Structure Library is available on the companion website containing >100 structures which can be examined in detail using the CrystalViewer Software. The structures which correspond directly to figures in the book are listed below, with the relevant figure number noted in parentheses. Many more crystal structures are available online, including minerals and other inorganic structures. Further structures may be added from time to time.

Major Inorganic Structure Types (and relevant book diagrams)

β-alumina, $NaAl_{11}O_{17}$ (8.23 and 8.24)
$BaTiO_3$ (8.40)
bcc metal (2.12)
Brass, ZnCu (2.11)
Brownmillerite, $Ca_2(Fe,Al)_2O_5$ (1.42)
CaC_2 (1.10)
$CaCu_3Ti_4O_{12}$ (1.42)
$CdCl_2$ (1.40)
CdI_2 (1.39)
Chevrel Phase, $BaMo_6S_8$ (8.6)
Corundum, α-Al_2O_3 (1.46)
CsCl (1.36)
Diamond (1.33)
fcc metal (1.20)
Fluorite/antifluorite, CaF_2 (1.29, 1.30 and 1.34)
Garnet, $Y_3Fe_5O_{12}$ (1.49)
$GdFeO_3$ (1.41)
hcp metal (1.21)
Hollandite (8.27)
Ilmenite, $FeTiO_3$ (1.46)
K_2NiF_4 (1.50)
Layered double hydroxides (4.11)
Li_3N (8.32)
$LiCoO_2$/α-$NaFeO_2$ (8.35)
$LiNbO_3$ (1.46)
Magnetoplumbite (9.14)
MgB_2 (1.51)
Nasicon, $NaZr_2(PO_4)_3$ (8.27)
Nickel arsenide, NiAs (1.35)
Olivine, $LiFePO_4$ (1.45)
PbFCl, matlockite (8.6)
PbO (3.14)
Perovskite, $SrTiO_3$ (1.41)
Pyrochlore (1.48)
Rock salt, NaCl (1.2, 1.29 and 1.31)
Rutile, TiO_2 (1.37)
Spinel (1.44)
Tetragonal tungsten bronze (1.43)
Wurtzite, ZnS (1.35)
$YBa_2Cu_3O_6$ (8.8)
$YBa_2Cu_3O_7$ (8.8)
Zinc blende/sphalerite, ZnS (1.29 and 1.33)
ZrCuSiAs (8.6)

Biography

Tony West obtained his BSc degree in Chemistry at University College Swansea and his PhD at the University of Aberdeen, where he worked with Professor F. P. Glasser on silicate chemistry. He was appointed as a Lecturer in Aberdeen in 1971 and developed a lifetime interest in the then-emerging field of solid state chemistry with special interest in the synthesis of new oxide materials, their crystal structures and electrical properties. He was awarded a DSc from Aberdeen in 1984 and rose through the ranks to become Professor of Chemistry in 1989 before moving to the University of Sheffield, Department of Materials Science and Engineering, as Head of Department in 1999, a post he held until 2007.

Tony was founding editor of the *Journal of Materials Chemistry* and subsequently established the Materials Chemistry Forum, which has now become the Materials Chemistry Division of the Royal Society of Chemistry. He organised the First International Conference on Materials Chemistry, MCI, in Aberdeen, 1993, and co-organised the first Materials Discussion, MDI, in Bordeaux, 1998. He also served as President of the Inorganic Chemistry Division of IUPAC, 2004–2007.

Tony is a Fellow of the Royal Society of Chemistry, the Institute of Physics, the Institute of Materials, Minerals and Mining (IOM^3), and the Royal Society of Edinburgh. Over the years he has received several awards, including an Industrial Award in Solid State Chemistry from the RSC (1996), the Griffiths Medal and Prize from the IOM^3 (2008), the Epsilon de Oro Award from the Spanish Society of Glass and Ceramics (2007) and the Chemical Record Lectureship from the Chemical Societies of Japan (2007). He has been awarded the 2013 John B. Goodenough Award in Materials Chemistry by the RSC, a lifetime award which recognises exceptional and sustained contributions to the field of materials chemistry.

1

Crystal Structures and Crystal Chemistry

Solid state chemistry is concerned mainly with crystalline inorganic materials, their synthesis, structures, properties and applications. A good place to begin is with *crystal structures* and *crystal chemistry*. All necessary crystal structure information is contained in data on unit cells, their dimensions and the positions or atomic coordinates of atoms inside the unit cell. Crystal chemistry combines this basic structural information with information about the elements, their principal oxidation states, ionic radii, coordination requirements and preferences for ionic/covalent/metallic bonding. A working knowledge of the Periodic Table and the properties of elements is, of course, invaluable to be able appreciate crystal chemistry, but conversely, knowledge of crystal structures and especially crystal chemistry provides a very useful way to gain increased understanding of the elements and their compounds.

Many of the properties and applications of crystalline inorganic materials revolve around a surprisingly small number of structure types. In this chapter, the main families of inorganic structures are reviewed, especially those which have interesting properties; more details of the vast array of structures may be found in the encyclopaedic text by Wells and also in the Wyckoff *Crystal Structures* book series. First, however, we must consider some basic concepts of crystallography.

1.1 Unit Cells and Crystal Systems

Crystals are built up of regular arrangements of atoms in three dimensions; these arrangements can be represented by a repeat unit or motif called the *unit cell*. The unit cell is defined as *the smallest repeating unit which shows the full symmetry of the crystal structure*. Let us see exactly what this means, first in two dimensions. A section through the NaCl structure is shown in Fig. 1.1(a); possible repeat units are given in (b) to (e). In each, the repeat unit is a square and adjacent squares share edges and corners. Adjacent squares are identical, as they must be by definition; thus, all the squares in (b) have Cl^- ions at their corners and centres. The repeat units in (b), (c) and (d) are all of the same size and, in fact, differ only in their relative position. The choice of origin of the repeat unit is to some extent a matter of personal taste, even though its size, shape and orientation are fixed. The repeat unit of NaCl is usually chosen as (b) or (c) rather than (d) because it is easier to draw and visualise the structure as a whole if the repeat unit contains atoms or ions at special positions such as corners and edge centres. Another guideline is that usually the origin is chosen so that the symmetry of the structure is evident (next section).

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

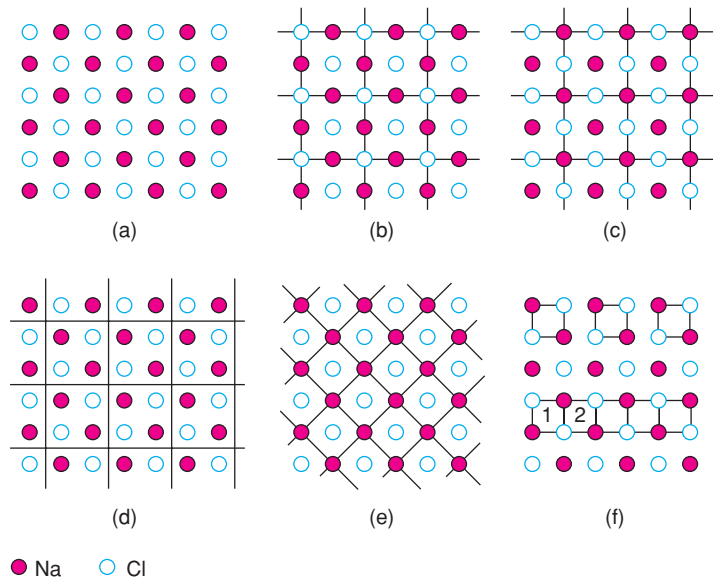


Figure 1.1 ***(a) Section through the NaCl structure, showing (b) to (e) possible repeat units and (f) incorrect units.***

In the hypothetical case that two-dimensional (2D) crystals of NaCl could form, the repeat unit shown in (e), or its equivalent with Cl at the corners and Na in the middle, would be the correct unit. Comparing (e) and, for example, (c), both repeat units are square and show the 2D symmetry of the structure; as the units in (e) are half the size of those in (c), (e) would be preferred according to the above definition of the unit cell. In three dimensions, however, the unit cell of NaCl is based on (b) or (c), rather than (e) because only they show the cubic symmetry of the structure (see later).

In (f) are shown two examples of what is *not* a repeat unit. The top part of the diagram contains isolated squares whose area is one-quarter of the squares in (c). It is true that each square in (f) is identical but it is not permissible to isolate unit cells or areas from each other, as happens here. The bottom part of the diagram contains units that are not identical; thus square 1 has Na in its top right corner whereas 2 has Cl in this position.

The unit cell of NaCl in three dimensions is shown in Fig. 1.2; it contains Na at the corner and face centre positions with Cl at the edge centres and body centre. Each face of the unit cell looks like the unit area shown

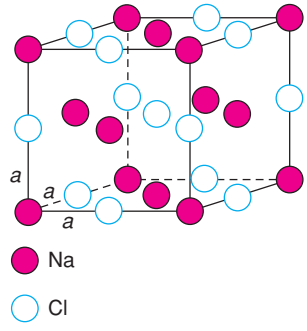


Figure 1.2 ***Cubic unit cell of NaCl,*** **a = b = c.**

Table 1.1 ***The seven crystal systems***

Crystal system	Unit cell shape[b]	Essential symmetry	Allowed lattices
Cubic	$a = b = c$, $\alpha = \beta = \gamma = 90^\circ$	Four threefold axes	P, F, I
Tetragonal	$a = b \neq c$, $\alpha = \beta = \gamma = 90^\circ$	One fourfold axis	P, I
Orthorhombic	$a \neq b \neq c$, $\alpha = \beta = \gamma = 90^\circ$	Three twofold axes or mirror planes	P, F, I, A (B or C)
Hexagonal	$a = b \neq c$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$	One sixfold axis	P
Trigonal (a)	$a = b \neq c$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$	One threefold axis	P
Trigonal (b)	$a = b = c$, $\alpha = \beta = \gamma \neq 90^\circ$	One threefold axis	R
Monoclinic[a]	$a \neq b \neq c$, $\alpha = \gamma = 90^\circ$, $\beta \neq 90^\circ$	One twofold axis or mirror plane	P, C
Triclinic	$a \neq b \neq c$, $\alpha \neq \beta \neq \gamma \neq 90^\circ$	None	P

[a]Two settings of the monoclinic cell are used in the literature, the most commonly used one given here, with b as the unique axis and the other with c defined as the unique axis: $a \neq b \neq c$, $\alpha = \beta = 90^\circ$, $\gamma \neq 90^\circ$.
[b]The symbol $\neq$ means 'not necessarily equal to'. Sometimes, crystals possess *pseudo-symmetry*. For example, a unit cell may be geometrically cubic but not possess the essential symmetry elements for cubic symmetry; the true symmetry is then lower, perhaps tetragonal.

in Fig. 1.1(c). As in the 2D case, the choice of origin is arbitrary; an equally valid unit cell could be chosen in which Na and Cl are interchanged. The unit cell of NaCl is *cubic*. The three edges: a, b and c are equal in length. The three angles: α (between b and c), β (between a and c) and γ (between a and b) are all 90°. A cubic unit cell also possesses certain symmetry elements, and these, together with the shape define the cubic unit cell.

The seven *crystal systems* listed in Table 1.1 and shown in Fig. 1.3 are the seven independent unit cell shapes that are possible in three-dimensional (3D) crystal structures. Six of these unit cell shapes are closely inter-related and are either cubic or can be derived by distorting a cube in various ways, as shown in Fig. 1.3(b).

Thus, if one axis, c, is of different length to the others, the shape is *tetragonal*; if all three axes are different, the shape is *orthorhombic*. If, now, one of the angles, β, is not 90°, the shape is *monoclinic*, whereas if all three angles differ from 90°, the shape is *triclinic*. Finally, if the cube is stretched, or compressed, along a body diagonal so that all three angles remain equal, but different from 90°, the shape is *trigonal*.

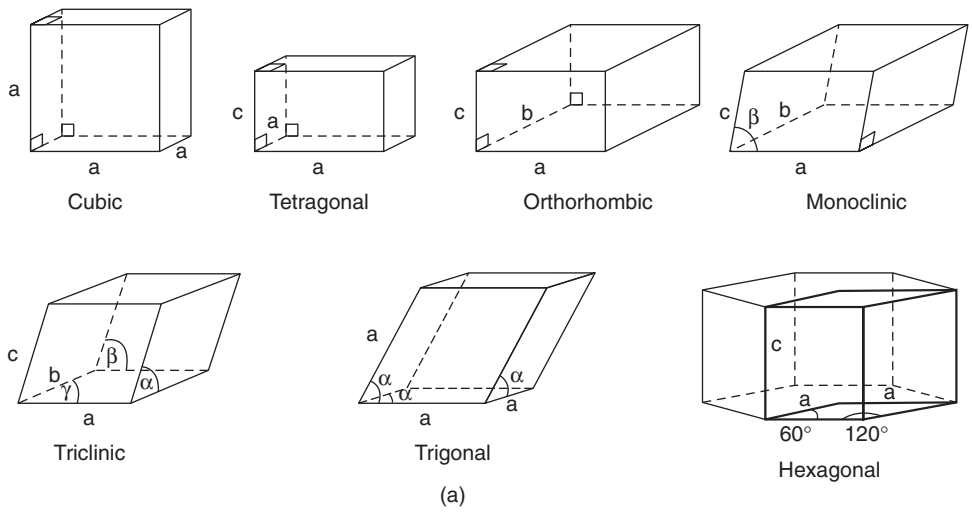
The remaining unit cell shape is *hexagonal*. A hexagonal-shaped box is shown in Fig. 1.3 and discussed later with reference to Fig. 1.21, but the true unit cell is only one-third of this size, as shown.

Although it is common practice to describe unit cells by their shapes, it is more correct to describe them by the presence or absence of symmetry. Thus, for example, if a unit cell has four intersecting threefold axes, it must be cubic in shape; the reverse does not necessarily apply and the unit cell could be fortuitously cubic but not have the threefold symmetries in the atomic arrangements. The essential symmetry for each crystal system is given in the third column of Table 1.1. Let us deal next with symmetry.

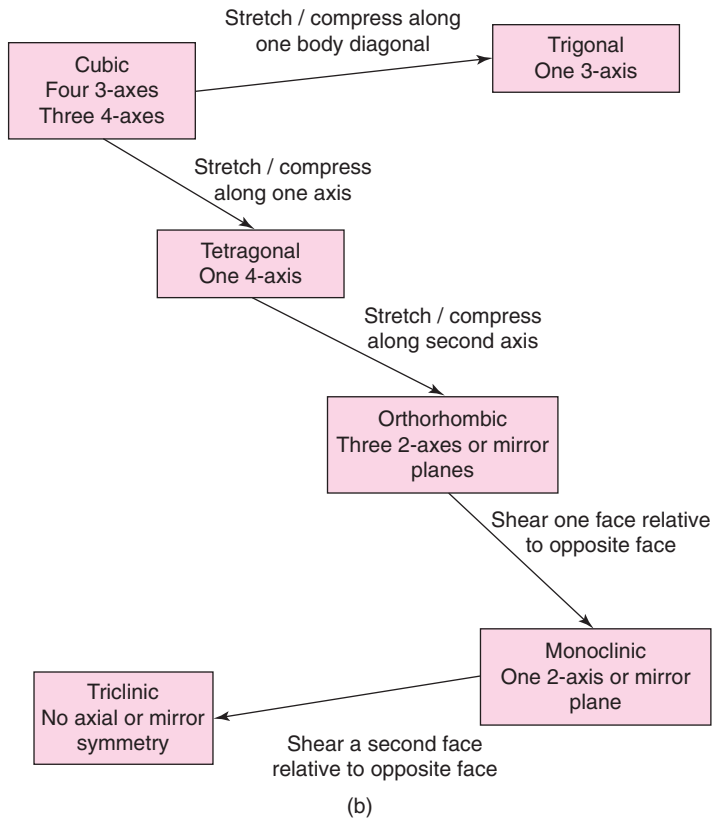
1.2 Symmetry

1.2.1 Rotational symmetry; symmetry elements and operations

Symmetry is most easily defined using examples. Consider the silicate tetrahedron shown in Fig. 1.4(a). If it is rotated about an axis passing along the vertical Si–O bond, then every 120° the tetrahedron finds itself in an identical position. Effectively, the three basal oxygens change position with each other every 120°. During a complete 360° rotation, the tetrahedron passes through three such identical positions. The fact that different (i.e. >1) identical orientations are possible means that the SiO_4 tetrahedron possesses symmetry. The axis



(a)



(b)

Figure 1.3 ***(a) The seven crystal systems and their unit cell shapes; (b) five of the seven crystal systems can be derived from cubic by structural distortions.***

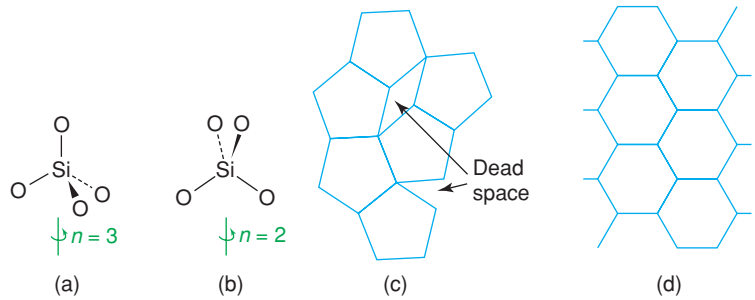


Figure 1.4 ***(a) Threefold and (b) twofold rotation axes; (c) the impossibility of forming a complete layer of pentagons; (d) a complete layer of hexagons.***

about which the tetrahedron may be rotated is called a *rotation axis*; it is an example of a *symmetry element*. The process of rotation is an example of a *symmetry operation*.

The symmetry elements that are important in crystallography are listed in Table 1.2. There are two nomenclatures for labelling them, the Hermann–Mauguin system used in crystallography and the Schönflies system used in spectroscopy. Ideally, there would be only one system which everybody uses, but this is unlikely to come about since (a) both systems are very well established, (b) crystallographers require elements of *space symmetry* that spectroscopists do not, and vice versa, (c) spectroscopists use a more extensive range of *point symmetry* elements than crystallographers.

The symmetry element described above for the silicate tetrahedron is a rotation axis, with symbol n. Rotation about this axis by $360/n$ degrees gives an identical orientation and the operation is repeated n times before the original configuration is regained. In this case, $n = 3$ and the axis is a *threefold rotation axis*. The SiO_4 tetrahedron possesses four threefold rotation axes, one in the direction of each Si–O bond.

When viewed from another angle, SiO_4 tetrahedra possess *twofold rotation axes* [Fig. 1.4(b)] which pass through the central Si and bisect the O–Si–O bonds. Rotation by 180° leads to indistinguishable orientations of the tetrahedra. The SiO_4 tetrahedron possesses three of these twofold axes.

Crystals may display rotational symmetries 2, 3, 4 and 6. Others, such as $n = 5, 7$, are never observed in 3D crystal structures based on a regular periodic repetition of the unit cell and its contents. This is shown in Fig. 1.4(c), where a fruitless attempt has been made to pack pentagons to form a complete layer; thus, individual pentagons have fivefold symmetry but the array of pentagons does not. For hexagons with sixfold

Table 1.2 ***Symmetry elements***

	Symmetry element	Hermann–Mauguin symbols (crystallography)	Schönflies symbols (spectroscopy)
Point symmetry	Mirror plane	m	σ_v, σ_h
	Rotation axis	$n = 2, 3, 4, 6$	C_n (C_2, C_3, etc.)
	Inversion axis	$\bar{n}$ ($= \bar{1}, \bar{2}$, etc.)	–
	Alternating axis[a]	–	S_n (S_1, S_2, etc.)
	Centre of symmetry	$\bar{1}$	i
Space symmetry	Glide plane	a, b, c, d, n	–
	Screw axis	2_1, 3_1, etc.	–

[a]The *alternating axis* is a combination of rotation (n-fold) and reflection perpendicular to the rotation axis. It is little used in crystallography.

rotation axes (d), a complete layer is easily produced; both the individual hexagons and the overall array exhibit sixfold symmetry. This is not to say that molecules which have pentagonal symmetry, $n = 5$, cannot exist in the crystalline state. They can, of course, but their fivefold symmetry cannot be exhibited by the crystal as a whole.

1.2.2 Quasicrystals

The discovery of a new state of matter, the *quasicrystalline state*, by Schechtman and colleagues in 1982 (and which led to the Nobel Prize in Chemistry in 2011) appeared at first sight to violate the rules concerning allowable rotational symmetries in crystal lattices. From their single-crystal diffraction patterns, rotational symmetries such as $n = 5$ but also $n = 10$ and 12 were observed whereas, as shown in Fig. 1.4(c), a regular crystal lattice exhibiting fivefold rotational symmetry cannot exist. The answer to this conundrum is that quasicrystals do not have regularly repeating crystal structures based on a single unit cell motif. Instead, they have fully ordered but non-periodic arrays constructed from more than one motif or building block.

Elegant examples of quasisymmetry are found in so-called Penrose tiling, as shown in Fig. 1.5. In this example, space is filled completely by a combination of red and blue diamonds; such a tiling pattern has many local areas of fivefold symmetry but the structure as a whole is not periodic, does not exhibit fivefold symmetry and a regular repeat unit cannot be identified. Quasicrystals have since been discovered in a wide range of alloy systems and also in organic polymer and liquid crystal systems; they have been discovered in Nature in an Al–Cu–Fe alloy named icosahedrite that was believed to have been part of a meteorite and had existed on Earth for billions of years. It is probably just a matter of time before they are discovered also in inorganic oxide materials, natural or synthetic.

In the early days of work on quasicrystals, an alternative explanation for possible fivefold symmetry was based on *twinning*, as shown schematically in Fig. 1.6. Five identical crystalline segments are shown, each of which has twofold rotational symmetry in projection. Pairs of crystal segments meet at a coherent interface or *twin plane* in which the structures on either side of the twin plane are mirror images of each other. The five crystal segments meet at a central point which exhibits fivefold symmetry as a macroscopic element of point symmetry but the individual crystal segments clearly do not exhibit any fivefold symmetry. Schechtman showed conclusively that twinning such as shown in Fig. 1.6 could not explain the quasicrystalline state.

1.2.3 Mirror symmetry

A *mirror plane*, *m*, exists when two halves of, for instance, a molecule can be interconverted by carrying out the imaginary process of reflection across the mirror plane. The silicate tetrahedron possesses six mirror planes, one of which, running vertically and perpendicular to the plane of the paper, is shown in Fig. 1.7(a). The silicon and two oxygens, 1 and 2, lie on the mirror plane and are unaffected by reflection. The other two oxygens, 3 and 4, are interchanged on reflection. A second mirror plane lies in the plane of the paper; for this, Si and oxygens 3, 4 lie on the mirror but oxygen 2, in front of the mirror, is the image of oxygen 1, behind the mirror.

1.2.4 Centre of symmetry and inversion axes

A *centre of symmetry*, $\bar{1}$, exists when any part of a structure can be reflected through this centre of symmetry, which is a point, and an identical arrangement found on the other side. An AlO_6 octahedron has a centre of symmetry, Fig. 1.7(b), located on the Al atom. If a line is drawn from any oxygen, e.g. 1, through the centre and extended an equal distance on the other side, it terminates at another oxygen, 2. A tetrahedron, e.g. SiO_4, does not have a centre of symmetry (a).

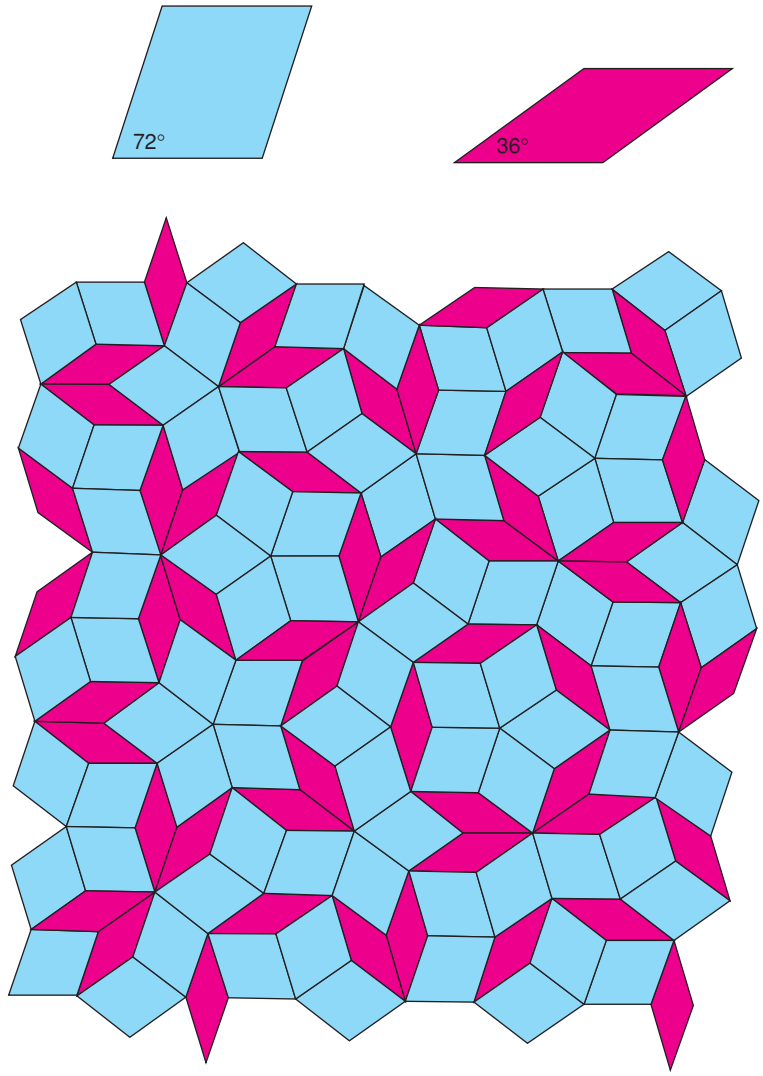


Figure 1.5 ***Two-dimensional Penrose tiling constructed by packing together two different sets of parallelograms. Adapted with permission from C. Janot,* Quasicrystals*, © 1997 Oxford University Press.***

The *inversion axis*, $\bar{n}$, is a combined symmetry operation involving rotation (according to n) and inversion through the centre. A $\bar{4}$ (*fourfold inversion*) axis is shown in (c). The first stage involves rotation by $360/4 = 90^\circ$ and takes, for example, oxygen 2 to position $2'$. This is followed by inversion through the centre, at Si, and leads to position 3. Oxygens 2 and 3 are therefore related by a $\bar{4}$ axis. Possible inversion axes in crystals are limited to $\bar{1}$, $\bar{2}$, $\bar{3}$, $\bar{4}$ and $\bar{6}$ for the same reason that only certain pure rotation axes are allowed. The onefold inversion axis is not a separate symmetry element, but is simply equivalent to the centre of symmetry; also, the twofold inversion axis is the same as a mirror plane perpendicular to that axis.

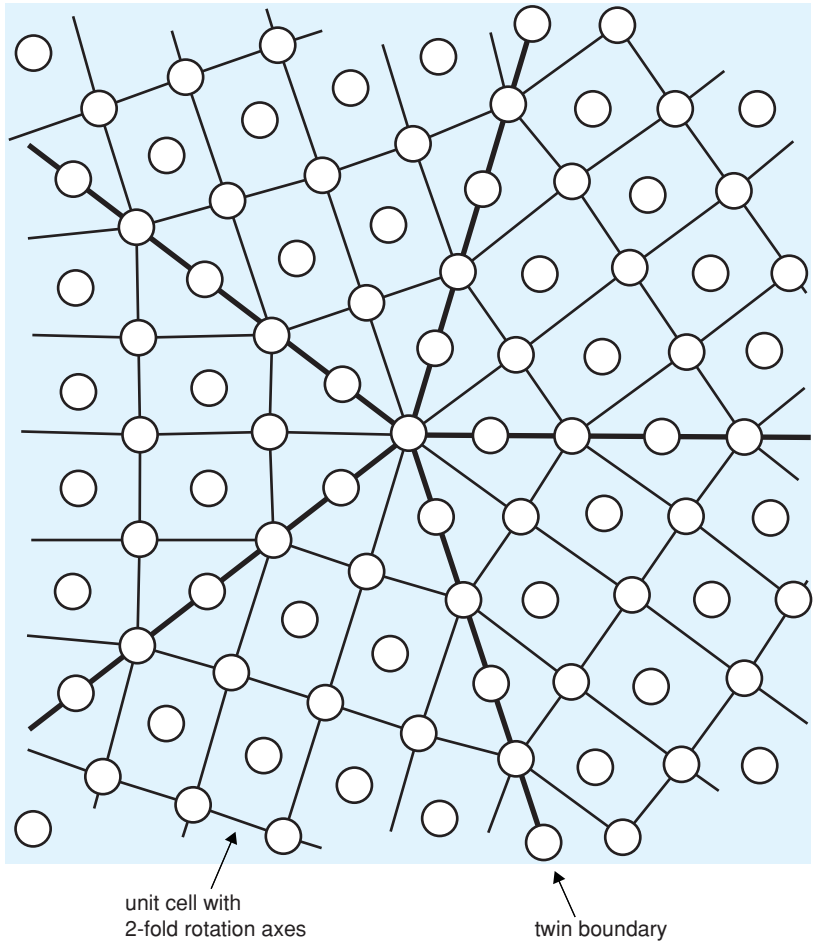


Figure 1.6 ***Hypothetical twinned structure showing fivefold symmetry. Adapted with permission from J. M. Dubois,* Useful Quasicrystals*, pg 10, © 2005 World Scientific Publishing Company.***

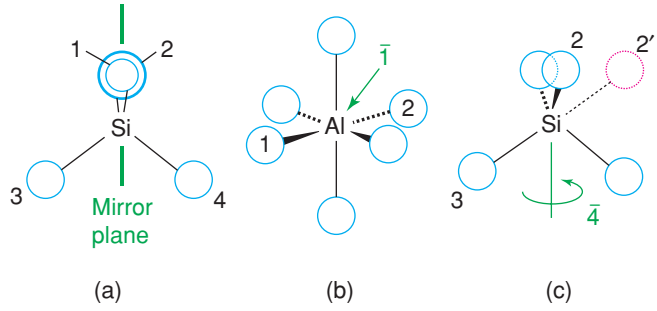


Figure 1.7 ***Symmetry elements: (a) mirror plane; (b) centre of symmetry; (c) fourfold inversion axis.***

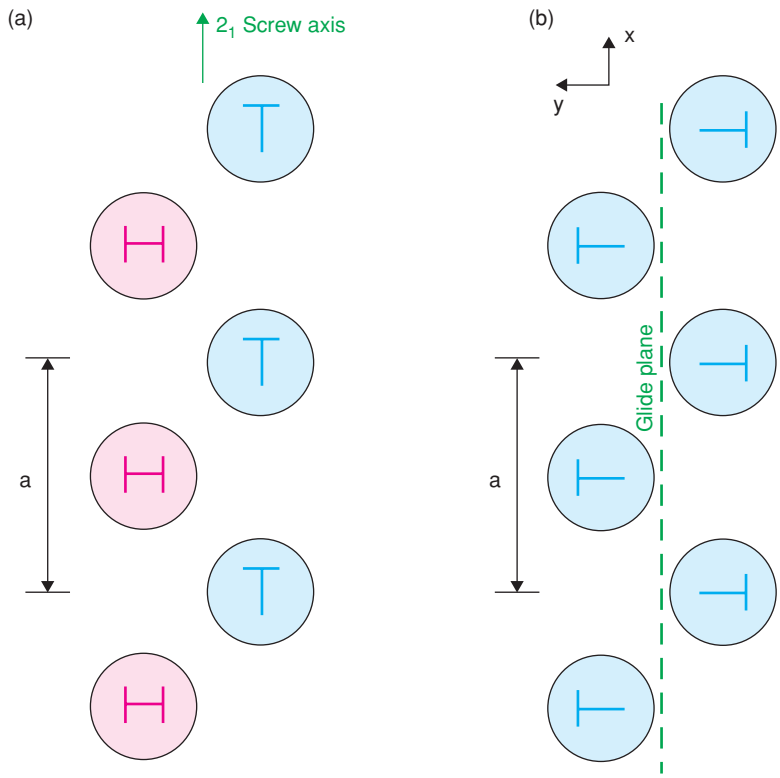


Figure 1.8 ***Arrangement of coins with heads (H) and tails (T) illustrating (a) a 2_1 screw axis parallel to a and (b) an a glide plane perpendicular to b; in each case, translation between symmetry-related objects is by a/2.***

1.2.5 Point symmetry and space symmetry

The symmetry elements discussed so far are elements of *point symmetry*. For each, at least one point stays unchanged during the symmetry operation, i.e. an atom lying on a centre of symmetry, rotation axis or mirror plane does not move during the respective symmetry operations. Finite-sized molecules can only possess point symmetry elements, whereas crystals may have extra symmetries that include translation steps as part of the symmetry operation. These are elements of *space symmetry*, of which there are two types.

The *screw axis* combines translation and rotation; the atoms or ions in a crystal which possesses screw axes appear to lie on spirals about these axes. A schematic example is shown in Fig. 1.8(a). All you need to demonstrate a screw axis is a handful of coins which can be arranged on a flat surface with either their heads (H) or tails (T) facing upwards. The symbol for a screw axis, X_Y, indicates translation by the fraction Y/X of the unit cell edge in the direction of the screw axis together with simultaneous rotation by 360/X° about the screw axis. Thus, a 4_2 axis parallel to *a* involves translation by *a*/2 and rotation by 90°; this process is repeated twice along *a* for every unit cell.

The *glide plane* combines translation and reflection, as shown schematically in Fig. 1.8(b). Translation may be parallel to any of the unit cell axes (*a*, *b*, *c*), to a face diagonal (*n*) or to a body diagonal (*d*). The *a*, *b*, *c* and *n* glide planes all have a translation step of half the unit cell in that direction; by definition, the *d* glide has a translation step which is $^1/_4$ of the body diagonal. For the *axial glide planes a*, *b* and *c*, it is important

to know both the direction of the translation and the reflection plane, e.g. an *a* glide may be perpendicular to *b* (i.e. reflection across the *ac* plane) or perpendicular to *c*.

Most crystal structures contain examples of space symmetry elements which are often difficult to visualise unless one has 3D models available. A simple example of both a screw axis and a glide plane is shown in Fig. 1.21 and discussed in Section 1.12. It is always amusing to look for symmetries in everyday objects. Next time you are in a car park, look at the arrangement of parking lots, especially if they are arranged diagonally, and see if they are displaced relative to each other in the form of a glide plane!

One difference between *crystals* and *quasicrystals* concerns space symmetry. Crystals, as a bare minimum, exhibit a periodic translation from one unit cell to the next. In many cases, translational symmetry elements, i.e. screw axes and glide planes, are also present. By contrast, quasicrystals do not exhibit space symmetry: they do not have a regular repeat unit, nor can translational symmetry elements be identified.

1.3 Symmetry and Choice of Unit Cell

The geometric shapes of the various crystal systems (unit cells) are listed in Table 1.1 and are shown in Fig. 1.3. These shapes do not *define* the unit cell; they are merely a consequence of the presence of certain symmetry elements.

A *cubic* unit cell is defined as one having four threefold symmetry axes. These run parallel to the cube body diagonals, one of which is shown in Fig. 1.9(a); an automatic consequence of this condition is that $a = b = c$ and $\alpha = \beta = \gamma = 90^\circ$. The *essential* symmetry elements by which each crystal system is defined are listed in Table 1.1. In most crystal systems, additional symmetry elements are also present. For instance, cubic crystals may have many others, including three fourfold axes passing through the centres of each pair of opposite cube faces (a) and mirror planes in two sets of orientations (b, c).

The *tetragonal* unit cell, Table 1.1, has a single fourfold axis and may be regarded as a cube that is either squashed or elongated along one axis. Consequently, all threefold rotation axes and two of the fourfold axes are lost. An example is given by the structure of CaC_2. This is related to the NaCl structure but, because the carbide ion is cigar-shaped rather than spherical, one of the cell axes becomes longer than the other two, Fig. 1.10(a). A similar tetragonal cell may be drawn for NaCl; it occupies half the volume of the true, cubic unit cell, Fig. 1.10(b). The choice of a tetragonal unit cell for NaCl is rejected, however, because it does not show the full cubic symmetry, i.e. it does not show the threefold axes. A tetragonal unit cell has one axis, the *unique axis*, which is of different length to the other two. By convention, this is chosen as the *c* axis, Table 1.1.

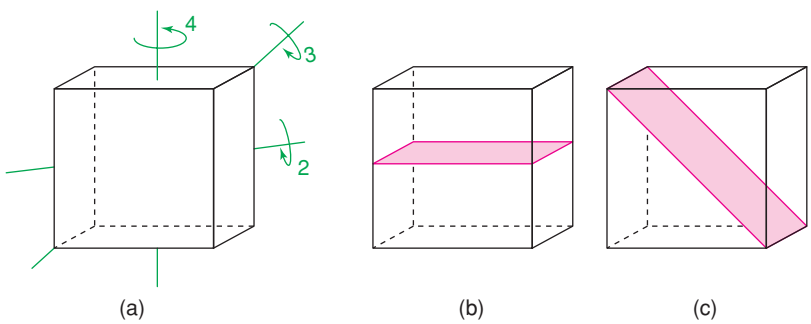


Figure 1.9 ***(a) Two-, three-, and fourfold axes and (b, c) mirror planes of a cube.***

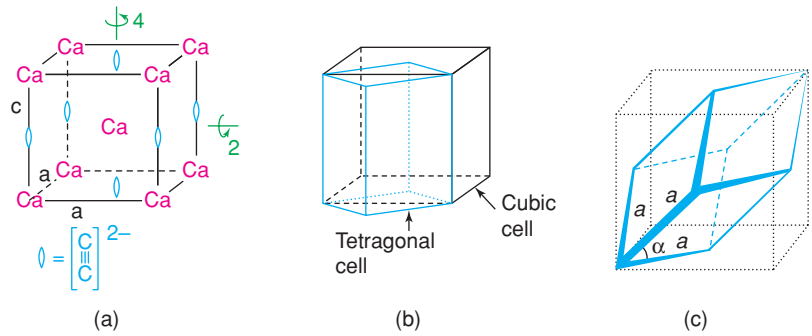


Figure 1.10 ***(a) Tetragonal unit cell of CaC_2: note the cigar-shaped carbide ions are aligned parallel to c; (b) relation between 'tetragonal' and cubic cells for NaCl; (c) derivation of a primitive trigonal unit cell for NaCl from the cubic cell.***

The *trigonal* system is characterised by a single threefold axis. Its shape is derived from a cube by stretching or compressing the cube along one of its body diagonals, Fig. 1.10(c). The threefold axis parallel to this direction is retained but those along the other body diagonals are destroyed. All three cell edges remain the same length; all three angles stay the same but are not equal to 90°. It is possible to describe such a trigonal cell for NaCl with $\alpha = \beta = \gamma = 60°$ with Na at the corners and Cl in the body centre (or vice versa), but this is again unacceptable because NaCl has symmetry higher than trigonal.

The structure of $NaNO_3$ is a trigonal distortion of the NaCl structure: instead of spherical Cl^- ions, it has triangular nitrate groups. These lie on planes perpendicular to one of the unit cell body diagonals. This causes a compression along one body diagonal (or rather an expansion in the plane perpendicular to the diagonal). All fourfold symmetry axes and all but one of the threefold axes are destroyed.

The *hexagonal* crystal system is discussed later (Fig. 1.21).

The *orthorhombic* unit cell may be regarded as a shoebox in which the angles are 90° but the sides are of unequal length. It usually possesses several mirror planes and twofold axes; the minimum requirement for orthorhombic symmetry is the presence of three mutually-perpendicular mirror planes or twofold axes.

The *monoclinic* unit cell may be derived from our orthorhombic shoebox by partially shearing the top face relative to the bottom face and in a direction parallel to one of the box edges. As a consequence, one angle departs from 90° and most of the symmetry is lost, apart from a mirror plane and/or a single twofold axis. One of the monoclinic unit cell axes is unique since it is perpendicular to the other two. The unique axis is usually defined as b, with $\alpha = \gamma = 90°$ (but see Table 1.1, footnote).

The *triclinic* system possesses no essential symmetry. Its unit cell has sides of unequal length and the angles are unequal and differ from 90°.

We have specified the essential symmetry elements for each crystal system. Consequently, the unit cell shapes follow automatically. The converse does not apply, however, in crystals which exhibit *pseudo-symmetry*. For example, the unit cell of a particular structure may be geometrically cubic but the atomic arrangement may not exhibit the requisite symmetry elements and, instead, the true crystal system may have symmetry lower than cubic.

1.4 Lattice, Bravais Lattice

It is very useful to be able to represent the manner of repetition of atoms, ions or molecules in a crystal by an array of points; the array is called a *lattice* and the points *lattice points*; each lattice point has exactly the

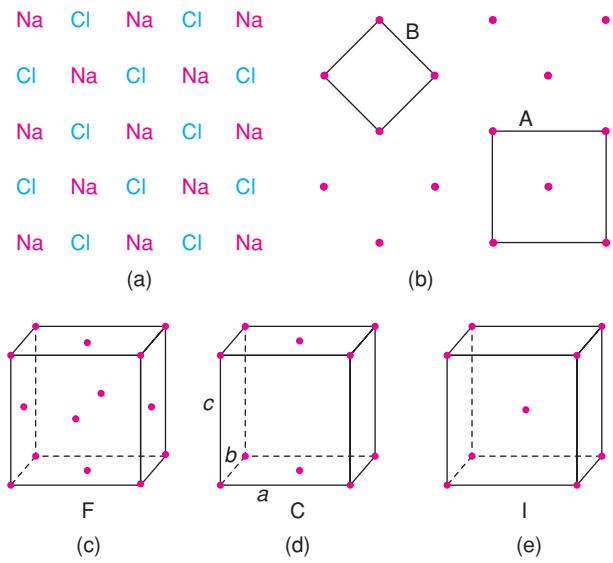


Figure 1.11 ***Representation of (a) the NaCl structure in two dimensions by (b) an array of lattice points; (c) face centred, (d) side centred and (e) body centred lattices.***

same environment and in the same orientation. The section of the NaCl structure shown in Fig. 1.11(a) may be represented by an array of points (b); each point represents one Na and one Cl but whether the point is located at Na, at Cl, or in between is irrelevant. The unit cell is constructed by linking the lattice points; two ways of doing this, A and B, are shown in Fig. 1.11(b). A cell such as B which has lattice points only at the corners is *primitive*, P, whereas a cell such as A which has additional lattice points is *centred*. Several types of centred lattice are possible.

The *face centred lattice*, F, contains additional lattice points in the centre of each face [Fig. 1.11(c)]. An example of a face centred cubic, *fcc*, structure is Cu metal. A *side centred lattice* contains extra lattice points on only one pair of opposite faces. It is labelled a C-centred lattice if the extra lattice points are on the *ab* face of the unit cell, as in (d). Similarly, an A-centred lattice has lattice points on the *bc* face.

A *body centred lattice*, I, has an extra lattice point at the body centre (e). α-Iron is *body centred cubic, bcc,* because it has a cubic unit cell with Fe atoms at the corner and body centre positions.

CsCl is also cubic with Cs at corners and Cl at the body centre (or vice versa), but it is *primitive*, P. This is because, for a lattice to be body centred, the atom or group of atoms located at or near the corner must be identical with those at or near the body centre.

In the simplest cases of monatomic metals such as Cu and α-Fe, mentioned above, the arrangement of metal atoms in the structure is simply the same as the arrangement of lattice points. In more complex structures such as NaCl, the lattice point represents an ion pair. This is still a very simple example, however, and in most inorganic structures the lattice point represents a considerable number of atoms. In crystals of organic molecules such as proteins, the lattice point represents an entire protein molecule. Obviously the lattice point gives no information whatsoever as to the atoms and their arrangements which it represents; what the lattice does show is how these species are packed together in 3D.

The combination of crystal system and lattice type gives the *Bravais lattice* of a structure. There are 14 possible Bravais lattices. They are given in Table 1.1, and shown in Fig. 1.12, by combining crystal system, column 1, and lattice type, column 4. For example, P-monoclinic, C-centred monoclinic and P-triclinic are three

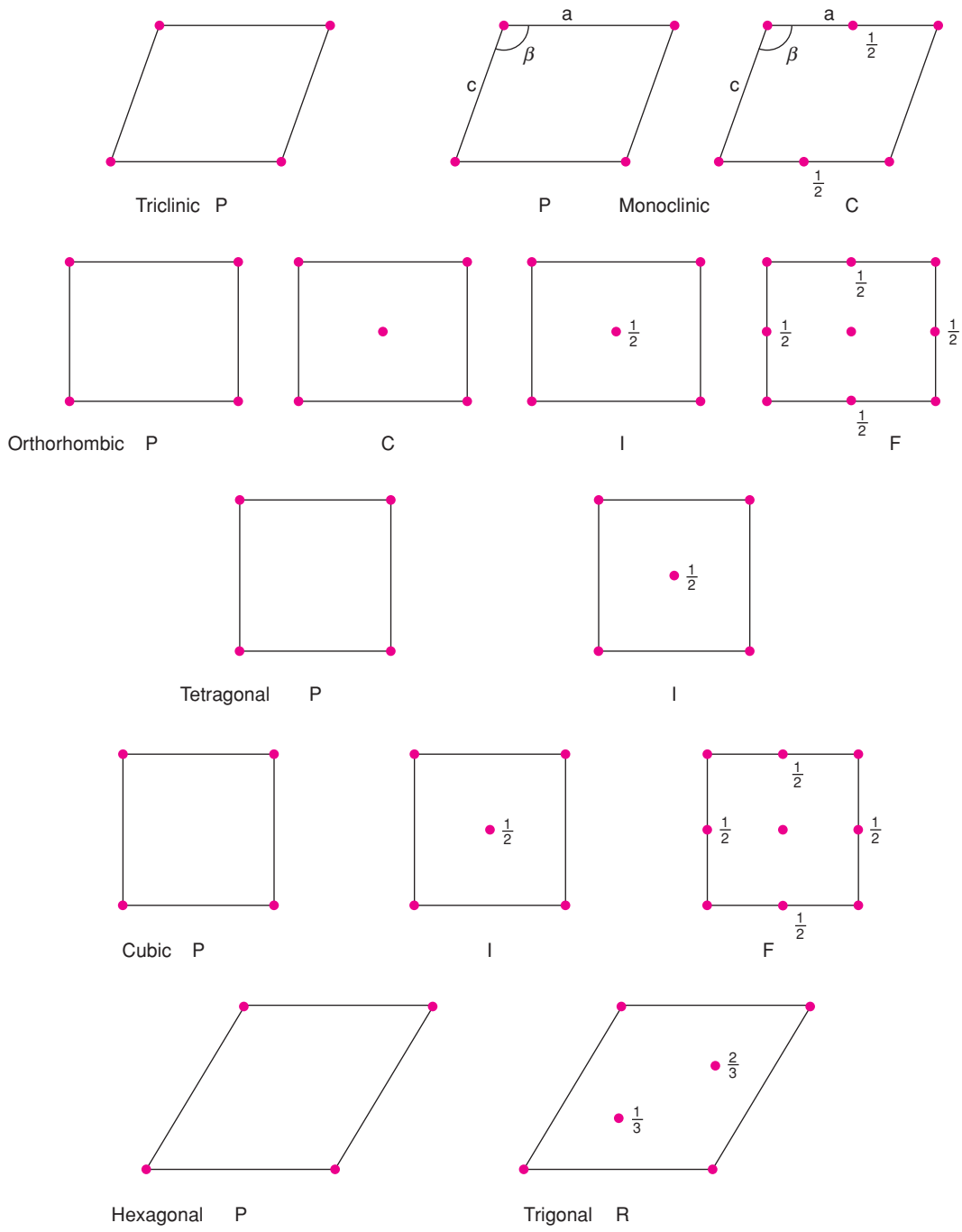


Figure 1.12 ***The unit cells of the 14 Bravais lattices: axes refer to the*** **ab** ***plane, unless specified. Heights of lattice points are 0,1, unless specified. Adapted with permission from*** **Essentials of Crystallography** ***© 1986 McKie and McKie, Blackwell.***

of the 14 possible Bravais lattices. The lattice type plus unit cell combinations which are absent either (a) would violate symmetry requirements, e.g. a C-centred lattice cannot be cubic because it would not have the necessary threefold axes or (b) may be represented by a smaller, alternative cell, e.g. a face centred tetragonal cell can be redrawn as a body centred tetragonal cell; the symmetry is still tetragonal but the volume is halved, Fig. 1.10(b).

1.5 Lattice Planes and Miller Indices

The concept of lattice planes causes considerable confusion because there are two separate ideas which can easily become mixed. Any simple structure, such as a metal or an ionic structure, may, in certain orientations, be regarded as built of layers or planes of atoms stacked to form a 3D structure. These layers are often related in a simple manner to the unit cell; for example, a unit cell face may coincide with a layer of atoms. The reverse is not necessarily true, however, especially in more complex structures and unit cell faces or simple sections through the unit cell often do not coincide with layers of atoms in the crystal. *Lattice planes*, a concept introduced with Bragg's law of diffraction (Chapter 5), are defined purely from the shape and dimensions of the unit cell. Lattice planes are entirely imaginary and simply provide a reference grid to which the atoms in the crystal structure may be referred. Sometimes, a given set of lattice planes coincides with layers of atoms, but not usually.

Consider the 2D array of lattice points shown in Fig. 1.13(a). This array may be divided into many different sets of rows and for each set there is a characteristic perpendicular distance, d, between pairs of adjacent rows. In three dimensions, these rows become *lattice planes* and adjacent planes are separated by the *interplanar d-spacing*, d. (Bragg's law treats X-rays as being diffracted from these various sets of lattice planes and the Bragg diffraction angle, θ, for each set is related to the d-spacing by Bragg's law.)

Lattice planes are labelled by assigning three numbers known as *Miller indices* to each set. The derivation of Miller indices is illustrated in Fig. 1.13(b) (and those for a hexagonal lattice are shown in Fig. 1.14). The origin of the unit cell is at point 0. Two planes are shown which are parallel and pass obliquely through the unit cell. A third plane in this set must, by definition, pass through the origin. Each of these planes continues out to the surface of the crystal and in so doing cuts through many more unit cells; also, there are many more planes in this set parallel to the two shown, but which do not pass through this particular unit cell.

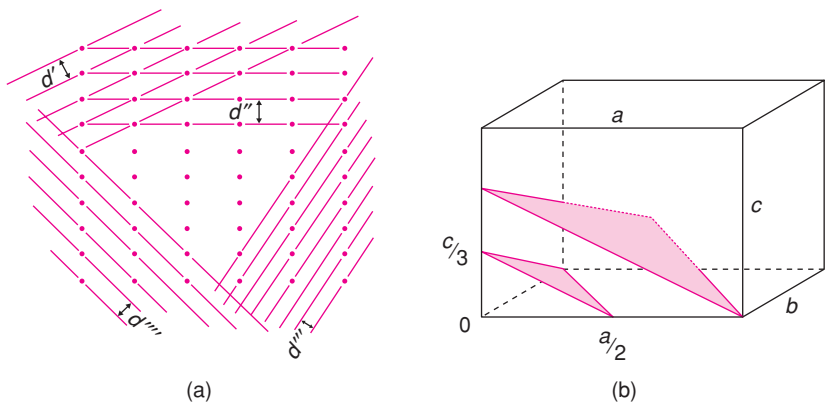


Figure 1.13 ***(a) Lattice planes (in projection); (b) derivation of Miller indices.***

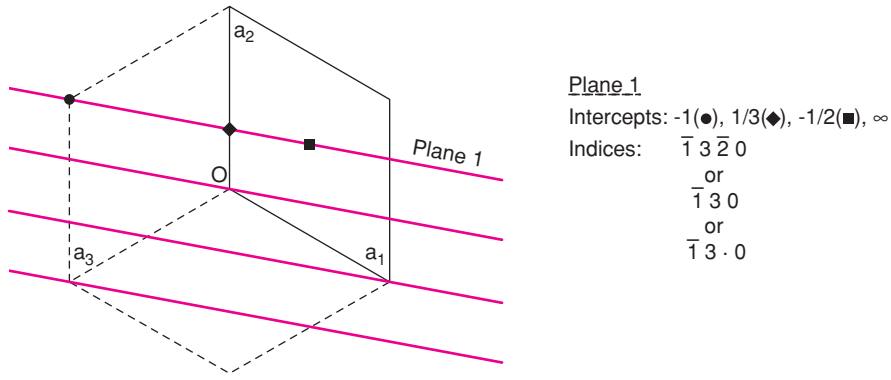


Figure 1.14 ***Miller indices for a hexagonal lattice.***

In order to assign Miller indices to a set of planes, there are four stages:

1. From the array of lattice points, or the crystal structure, identify the unit cell, choose the origin and label the axes a, b, c and angles α (between b and c), β (between a and c) and γ (between a and b).
2. For a particular set of lattice planes, identify that plane which is adjacent to the one that passes through the origin.
3. Find the intersection of this plane on the three axes of the cell and write these intersections as fractions of the cell edges. The plane in question, Fig. 1.13(b), cuts the x axis at $a/2$, the y axis at b and the z axis at $c/3$; the fractional intersections are therefore $^1/_2$, 1, $^1/_3$.
4. Take reciprocals of these fractions and write the three numbers in parentheses; this gives (213). These three integers, (213), are the Miller indices of the plane and all other planes parallel to it and separated from adjacent planes by the same d-spacing.

Some other examples are shown in Fig. 1.15. In (a), the shaded plane cuts x, y and z at $1a$, ∞b and $1c$, i.e. the plane is parallel to b. Taking reciprocals of 1, ∞ and 1 gives us (101) for the Miller indices. A Miller index of 0 means, therefore, that the plane is parallel to that axis. In Fig. 1.15 (b), the planes of interest comprise opposite faces of the unit cell. We cannot determine directly the indices of plane 1 as it passes through the origin. Plane 2 has intercepts of $1a$, ∞b and ∞c and Miller indices of (100).

Fig. 1.15(c) is similar to (b) but there are now twice as many planes as in (b). To find the Miller indices, consider plane 2, which is the one that is closest to the origin but without passing through it. Its intercepts are $^1/_2$, ∞ and ∞ and the Miller indices are (200). A Miller index of 2 therefore indicates that the plane cuts the relevant axis at half the cell edge. This illustrates an important point. After taking reciprocals, do not divide through by the highest common factor. A common source of error is to regard, say, the (200) set of planes as those planes interleaved between the (100) planes, to give the sequence (100), (200), (100), (200), (100), ... The correct labelling is shown in Fig. 1.15(d). If extra planes are interleaved between adjacent (100) planes then *all* planes are labelled as (200).

The general symbol for Miller indices is (*hkl*). It is not necessary to use commas to separate the three letters or numbers and the indices are enclosed in parentheses, (). The brackets {} are used to indicate sets of planes that are equivalent; for example, the sets (100), (010) and (001) are equivalent in cubic crystals and may be represented collectively as {100}. In the examples chosen so far, all of the Miller indices are either 0 or positive but, as we shall see later, it is also important to define lattice planes that intersect axes in

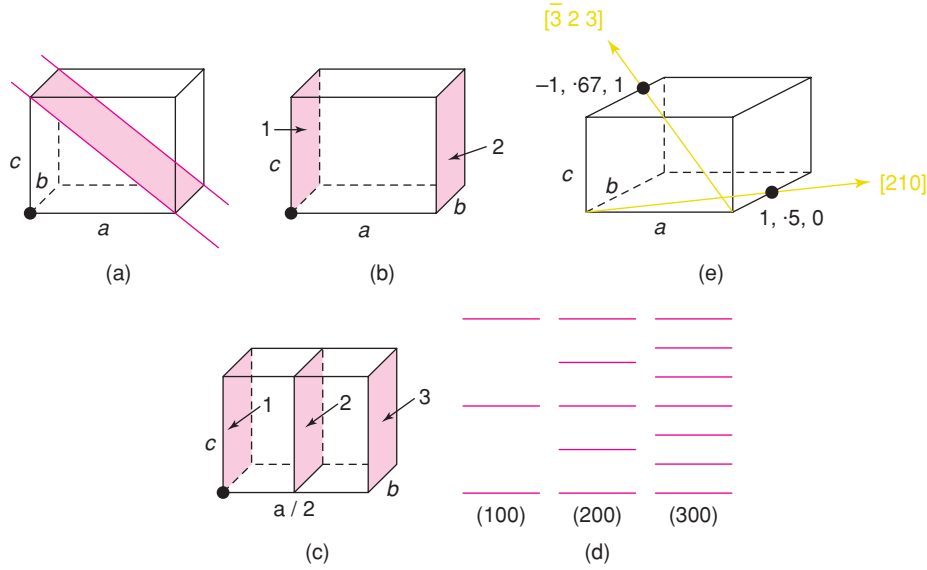


Figure 1.15 ***Examples of Miller indices: (a) (101); (b) (100); (c) (200); (d) (h00); (e) indices of directions [210] and [3̄23].***

their negative directions. In these cases the relevant index has a bar above the number; thus planes $(\bar{h},\bar{k},\bar{l})$ are referred to as 'bar h, bar k, bar l'.

To label lattice planes in hexagonal unit cells, a modification to the method is required. For hexagonal unit cells, with $a = a \neq c$, there are three possible choices for the two a axes in the basal plane, a_1, a_2 and a_3, as shown in Fig. 1.14. Let us choose a_1 and a_2 to define the unit cell with the origin at 0. Next, consider the set of planes shown, which are parallel to c; assign Miller indices in the usual way but in this case, find the intercepts of plane 1 on all three a axes as well as the c axis; since these planes in Fig. 1.14 are parallel to c, plane 1 does not intersect the c axis. This leads to four indices for this set, representing the intercepts on a_1, a_2, a_3 and c, i.e. $\bar{1}3\bar{2}0$. The third index is redundant since, writing the indices as $(hkil)$, the condition

$$h + k + i = 0$$

holds (i.e. $-1 + 3 - 2 = 0$). Sometimes, all four indices are specified; sometimes, only three are specified (i.e. $\bar{1}30$ in the usual way); sometimes, the third index may be represented as a dot (i.e. $\bar{1}3 \cdot 0$).

1.6 Indices of Directions

Directions in crystals and lattices are labelled by first drawing a line that passes through the origin and parallel to the direction in question. Let the line pass through a point with general fractional coordinates x, y, z; the line also passes through $2x$, $2y$, $2z$; $3x$, $3y$, $3z$, etc. These coordinates, written in square brackets, $[x, y, z]$, are the indices of the direction; x, y and z are arranged to be the set of smallest possible integers, by division or multiplication throughout by a common factor. Thus $[^1/_2\,^1/_20]$, [110], [330] all describe the same direction, but by convention [110] is used.

For cubic systems, an [*hkl*] direction is always perpendicular to the (*hkl*) plane of the same indices, but this is only sometimes true in non-cubic systems. Sets of directions which, by symmetry, are equivalent, e.g. cubic [100], [010], are written using angle brackets, <100>. Some examples of directions and their indices are shown in Fig. 1.15(e). Note that the [210] direction is defined by taking the origin at the bottom-left-front corner of the unit cell and taking the fractional coordinates: 1, 0.5, 0, to define the direction. For the direction $[\bar{3}23]$, where the bar sign indicates a negative direction, a different origin is chosen, at the bottom-right-front corner of the unit cell and passing through the fractional coordinates: −1, 0.67, 1.

1.7 *d*-Spacing Formulae

We have already defined the *d*-spacing of a set of planes as the perpendicular distance between any pair of adjacent planes in the set and it is this *d* value that appears in Bragg's law. For a cubic unit cell, the (100) planes simply have a *d*-spacing of *a*, the value of the cell edge, Fig. 1.15(b). For (200) in a cubic cell, $d = a/2$. For *orthogonal* crystals (i.e. $\alpha = \beta = \gamma = 90^\circ$), the *d*-spacing for any set of planes is given by

$$\frac{1}{d_{hkl}^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2} \tag{1.1}$$

The equation simplifies for tetragonal crystals, in which $a = b$, and still further for cubic crystals with $a = b = c$:

$$\frac{1}{d^2} = \frac{h^2 + k^2 + l^2}{a^2} \tag{1.2}$$

As a check, for cubic (200): $h = 2$, $k = l = 0$; $1/d^2 = 4/a^2$; $d = a/2$.

Monoclinic and, especially, triclinic crystals have much more complicated *d*-spacing formulae because each angle that is not equal to 90° is an additional variable. The formulae for *d*-spacings and unit cell volumes of all crystal systems are given in Appendix A.

1.8 Crystal Densities and Unit Cell Contents

The unit cell, by definition, must contain at least one formula unit, whether it be an atom, ion pair, molecule, etc. In centred cells, the unit cell contains more than one formula unit and more than one lattice point. A simple relation exists between cell volume, the number of formula units in the cell, the formula weight (FW) and the bulk crystal density (*D*):

$$D = \frac{\text{mass}}{\text{volume}} = \frac{\text{FW}}{\text{molar volume}} = \frac{\text{FW}}{\text{volume of formula unit} \times N}$$

where *N* is Avogadro's number. If the unit cell, of volume *V*, contains *Z* formula units, then

$$V = \text{volume of one formula unit} \times Z$$

Therefore,

$$D = \frac{\text{FW} \times Z}{V \times N} \tag{1.3}$$

V is usually expressed in Å^3 and must be multiplied by 10^{-24} to convert V to cm^{-3} and to give densities in units of g cm^{-3}. Substituting for N, the equation reduces to

$$D = \frac{\text{FW} \times Z \times 1.66}{V} \tag{1.4}$$

and, if V is in Å^3, the units of D are g cm^{-3}. This simple equation has a number of uses, as shown by the following examples:

(a) It can be used to check that a given set of crystal data is consistent and that, for example, an erroneous formula weight has not been assumed.
(b) It can be used to determine any of the four variables if the other three are known. This is most common for Z (which must be a whole number) but is also used to determine FW and D.
(c) By comparison of D_{obs} (the experimental density) and D_{calc} (calculated from the above equation), information may be obtained on the presence of crystal defects such as vacancies or interstitials, the mechanisms of solid solution formation and the porosity of ceramic pieces.

Considerable confusion often arises over the value of the contents, Z, of a unit cell. This is because atoms or ions that lie on corners, edges or faces are also shared between adjacent cells; this must be taken into consideration in calculating *effective* cell contents.

For example, α-Fe [Fig. 1.11(e)] has $Z = 2$. The corner Fe atoms, of which there are eight, are each shared between eight neighbouring unit cells. Effectively, each contributes only 1/8 to the particular cell in question, giving $8 \times 1/8 = 1$ Fe atom for the corners. The body centre Fe lies entirely inside the unit cell and counts as one. Hence $Z = 2$.

For Cu metal, Fig. 1.11(c), which is *fcc*, $Z = 4$. The corner Cu again counts as one. The face centre Cu atoms, of which there are six, count as 1/2 each, giving a total of $1 + (6 \times 1/2) = 4$ Cu in the unit cell.

NaCl is also *fcc* and has $Z = 4$. Assuming the origin is at Na (Fig. 1.2) the arrangement of Na is the same as that of Cu in Fig. 1.11(c) and therefore, the unit cell contains 4 Na. Cl occupies edge centre positions of which there are 12; each counts as 1/4, which, together with Cl at the body centre, gives a total of $(12 \times 1/4) + 1 = 4$ Cl. Hence the unit cell contains, effectively, 4 NaCl. If unit cell contents are not counted in this way, but instead all corner, edge- and face-centre atoms are simply counted as one each, then the ludicrous answer is obtained that the unit cell has 14 Na and 13 Cl!

1.9 Description of Crystal Structures

Crystal structures may be described in various ways. The most common, and one which gives all the necessary information, is to refer the structure to the unit cell. The structure is given by the size and shape of the cell and the positions of the atoms, i.e. *atomic coordinates*, inside the cell. However, a knowledge of the unit cell and atomic coordinates alone is often insufficient to give a revealing picture of what the structure looks like in 3D. The latter is obtained only by considering a larger part of the structure, comprising perhaps several unit cells, and by considering the arrangement of atoms relative to each other, their coordination numbers, interatomic distances, types of bonding, etc. It then becomes possible to find alternative ways of visualising structures and also to compare and contrast different types of structure.

Two of the most useful ways of describing structures are based on *close packing* and *space-filling polyhedra*. Neither can be applied to all types of structure and both have their limitations. They do, however, provide greater insight into crystal chemistry than is obtained using unit cells alone. It can be very useful to make your own crystal structure models, either from coloured spheres or polyhedra, and tips for constructing models are

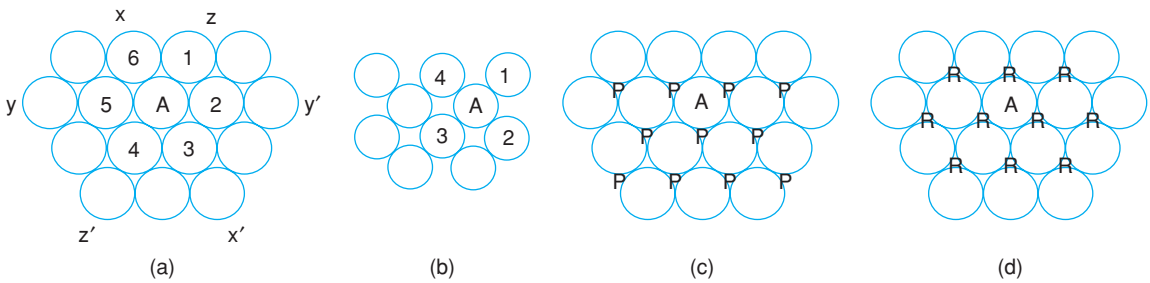


Figure 1.16 ***(a) A* cp *layer of equal-sized spheres; (b) a non-*cp *layer with coordination number 4; (c, d) alternative positions P and R for a second* cp *layer.***

given in Appendix B. In addition, all the structures described in this book, and many more, are available to view in the crystal viewer software package which can be downloaded, free of charge, from the Companion Wiley resource site. Using this, structures can be expanded, rotated, colours changed and key structural features highlighted by hiding selected atoms or polyhedra.

1.10 Close Packed Structures – Cubic and Hexagonal Close Packing

Many metallic, ionic, covalent and molecular crystal structures can be described using the concept of close packing. The guiding factor is that structures are usually arranged to have the maximum density. The principles involved can be understood by considering the most efficient way of packing equal-sized spheres in three dimensions.

The most efficient way to pack spheres in *two* dimensions is shown in Fig. 1.16(a). Each sphere, e.g. A, is surrounded by, and is in contact with, six others, i.e. each sphere has six *nearest neighbours* and its *coordination number, CN*, is six. By regular repetition, infinite sheets called *close packed layers* form. The coordination number of six is the maximum possible for a planar arrangement of contacting, equal-sized spheres. Lower coordination numbers are, of course, possible, as shown in Fig. 1.16(b), where each sphere has four nearest neighbours, but the layers are no longer close packed, *cp*. Note also that within a *cp* layer, three *close packed directions* occur. Thus, in Fig. 1.16(a) spheres are in contact in the directions xx', yy' and zz' and sphere A belongs to each of these rows.

The most efficient way to pack spheres in *three* dimensions is to stack *cp* layers on top of each other. There are two simple ways to do this, resulting in *hexagonal close packed* and *cubic close packed* structures as follows.

The most efficient way for two *cp* layers A and B to be in contact is for each sphere of one layer to rest in a hollow between three spheres in the other layer, i.e. at P or R in Fig. 1.16(c) and (d). Two layers in such a position relative to each other are shown in Fig. 1.17. Atoms in the second layer may occupy either P or R positions, but not both together, nor a mixture of the two. Any B (dashed) sphere is therefore seated between three A (solid) spheres, and vice versa.

Addition of a third *cp* layer to the two shown in Fig. 1.17 can also be done in two ways, and herein lies the distinction between hexagonal and cubic close packing. In Fig. 1.17, suppose that the A layer lies underneath the B layer and we wish to place a third layer on top of B. There is a choice of positions, as there was for the second layer: the spheres can occupy either of the new sets of positions S or T but not both together nor a mixture of the two. If the third layer is placed at S, then it is directly over the A layer. As subsequent layers are added, the following sequence arises:

. . . ABABAB . . .

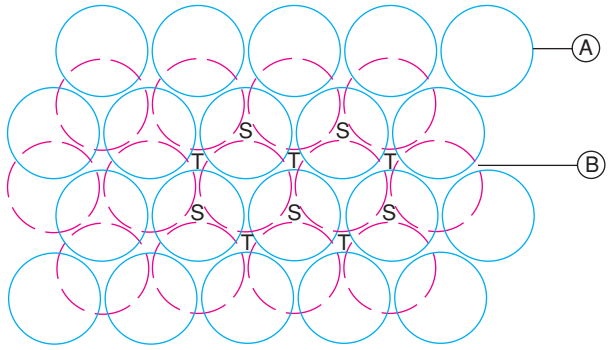


Figure 1.17 ***Two* cp *layers arranged in A and B positions. The B layer occupies the P positions of Fig. 1.16.***

This is known as *hexagonal close packing, hcp*. If, however, the third layer is placed at T, then all three layers are staggered relative to each other and it is not until the fourth layer is positioned (at A) that the sequence is repeated. If the position of the third layer is called C, this gives (Fig. 1.18)

. . . ABCABCABC . . .

This sequence is known as *cubic close packing* (*ccp*). The two simplest stacking sequences are *hcp* and *ccp* and these are by far the most important in structural chemistry. Other more complex sequences with larger repeat units, e.g. ABCACB or ABAC, occur in a few materials; some of these larger repeat units are responsible for the phenomenon of *polytypism*.

In a 3D *cp* structure, each sphere is in contact with *12* others, and this is the maximum coordination number possible for contacting and equal-sized spheres. [A common non-*cp* structure is the body centred cube, e.g. in α-Fe, with a coordination number of eight; see Fig. 1.11(e).] Six of these neighbours are coplanar with the central sphere, Fig. 1.16(a); from Fig. 1.17 and Fig. 1.18, the remaining six are in two groups of three spheres, one in the plane above and one in the plane below (Fig. 1.19); *hcp* and *ccp* differ in the relative orientations of these two groups of three neighbours.

Many structures, not just of metals and alloys, but also ionic, covalent and molecular structures, can be described using close packing ideas. Sometimes the atoms that form the *cp* array are as closely packed as possible, but in other cases their arrangement is as in *cp* but the atoms are clearly not touching. Such structures

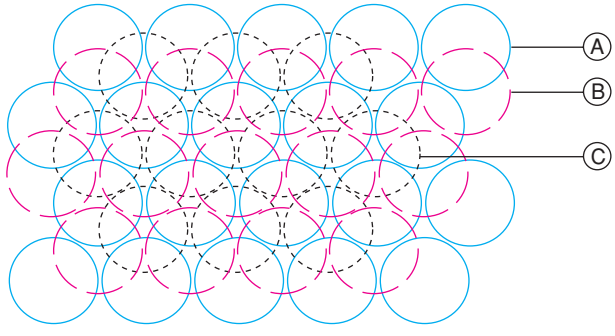


Figure 1.18 ***Three close packed layers in* ccp *sequence.***

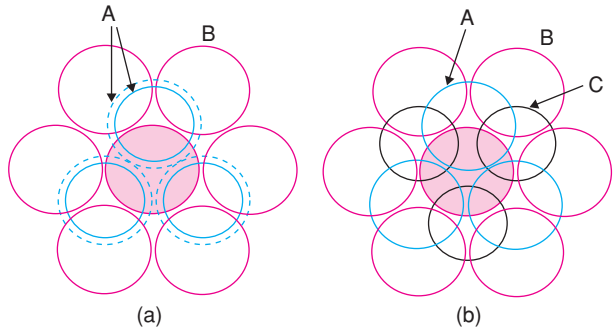


Figure 1.19 ***Coordination number 12 of shaded sphere in (a)* hcp *and (b)* ccp *structures. The shaded sphere is in the B layer, the layer underneath is A and the layer above is either (a) A or (b) C.***

are known as *eutactic* structures. Some guidelines as to whether it is appropriate to consider a structure in terms of a *cp* arrangement are given in Appendix D.

1.11 Relationship between Cubic Close Packed and Face Centred Cubic

The unit cell of a *ccp* arrangement is the familiar face centred cubic (*fcc*) unit cell, Fig. 1.11(c), with spheres at corner and face centre positions. The relation between *ccp* and *fcc* is not immediately obvious since the faces of the *fcc* unit cell do not correspond to *cp* layers. The *cp* layers are, instead, parallel to the {111} planes of the *fcc* unit cell. This is shown in Fig. 1.20 and Appendix B. The spheres labelled 2–7 in Fig. 1.20 (a) form part of a *cp* layer, as revealed by removing a corner sphere 1 in (b) and comparing (b) with Fig. 1.16(a). The orientations of (a) and (b) in Fig. 1.20 are the same but the spheres in (b) are shown larger. A similar arrangement to that shown in (b) would be seen on removing any corner sphere in (a) and, therefore, in a *ccp* structure, *cp* layers occur in four orientations. These orientations are perpendicular to the body diagonals of the cube (the cube has eight corners but only four body diagonals and, hence, four different orientations of the *cp* layers). The *cp* layers in one orientation are seen edge-on in (c) and in another orientation, perpendicular to the layers in (d); (c) is the same as (a) but rotated slightly; similarly, (d) is the same as (b) but also rotated slightly. In (d), all atoms of the unit cell, spanning four *cp* layers, ABCA, are shown in projection down a <111> direction.

1.12 Hexagonal Unit Cell and Close Packing

An *hcp* arrangement of spheres has a hexagonal unit cell (Fig. 1.21). The *basal plane* of the cell coincides with a *cp* layer of spheres (b). The unit cell contains two spheres, one at the origin (and hence at all corners) and one inside the cell at position $^1/_3, ^2/_3, ^1/_2$ (pink circle in (a) and (b)). Note that although the two a axes of the basal plane are equal, we need to distinguish them by a_1 and a_2 for the purpose of describing atomic coordinates of the position $^1/_3, ^2/_3, ^1/_2$. The use of such *fractional coordinates* to represent positions of atoms inside a unit cell is discussed later.

cp layers occur in only one orientation in an *hcp* structure. These are parallel to the basal plane, as shown for one layer in Fig. 1.21(b). The two axes in the basal plane are of equal length; $a = 2r$, if the spheres of radius r touch; the angle γ is 120° (Table 1.1).

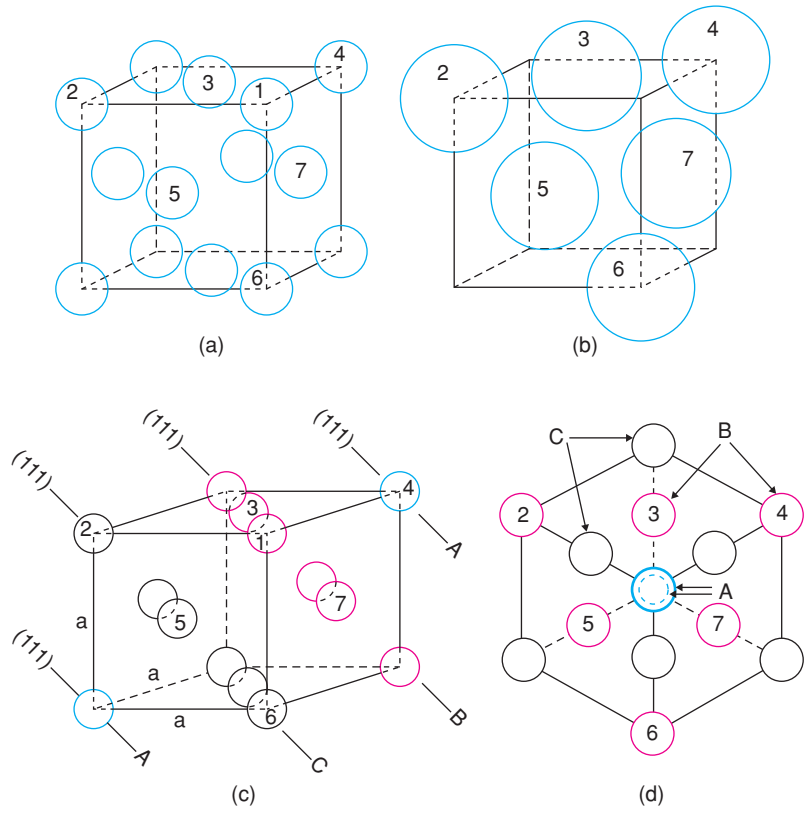


Figure 1.20 ***Face centred cubic,* fcc*, unit cell of a* ccp *arrangement of spheres.***

The symmetry of the hexagonal unit cell is deceptively simple. The basal plane in isolation has a sixfold rotation axis but the adjacent B layer along the *c* axis reduces this to threefold rotational symmetry, as shown in Fig. 1.21(c): note the crystallographic symbol for a threefold axis, which is a solid triangle.

The structure does, however, possess a 6_3 screw axis parallel to *c* and passing through the basal plane at the coordinate position $^2/_3$, $^1/_3$, 0, as shown in Fig. 1.21(d). This symmetry axis involves a combined step of translation by *c*/2 and rotation by 60°; atoms labelled 1–6 lie on a spiral with increasing *c* height above the basal plane; thus, atom 3 is on the top face of the unit cell whereas 4 and 5 are in the next unit cell in the *c* direction. Hence the *hcp* crystal structure has both a sixfold screw axis and a threefold rotation axis.

The *hcp* crystal structure has many other symmetry elements as well, including a nice example of a glide plane as shown in Fig. 1.21(e); the components of this *c*-glide involve displacement in the *c* direction by *c*/2 and reflection across the a_1c plane that passes through the unit cell with a_2 coordinate $^2/_3$, as shown by the dotted line (crystallographic symbol for a *c*-glide plane). Thus, atoms labelled 1, 2, 3, 4, etc. are related positionally to each other by this glide plane.

1.13 Density of Close Packed Structures

In *cp* structures, 74.05% of the total volume is occupied by spheres. This is the maximum density possible in structures constructed of spheres of only one size. This value may be calculated from the volume and contents

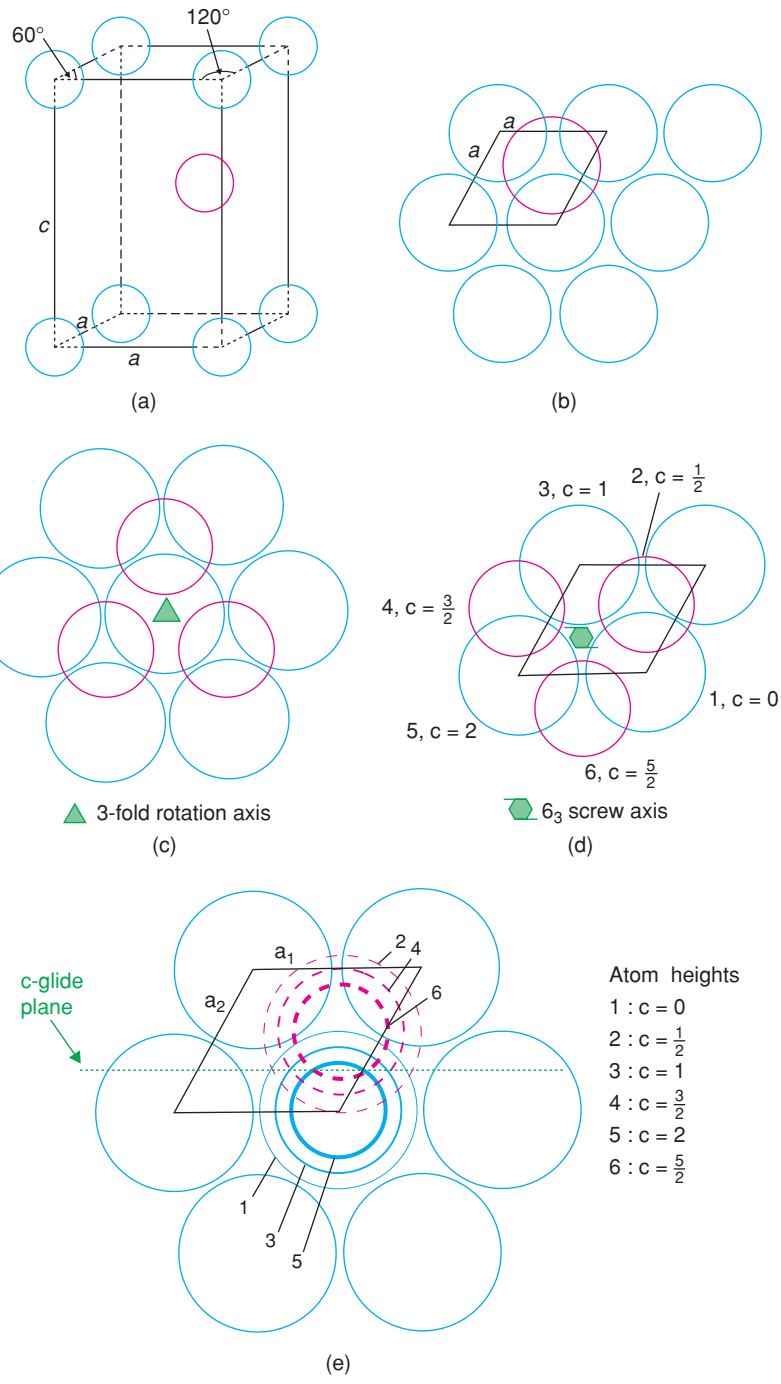


Figure 1.21 *(a, b) Hexagonal unit cell of an* **hcp** *arrangement of spheres showing (c) a threefold rotation axis, (d) a* 6_3 *screw axis and (e) a* **c**-*glide plane.*

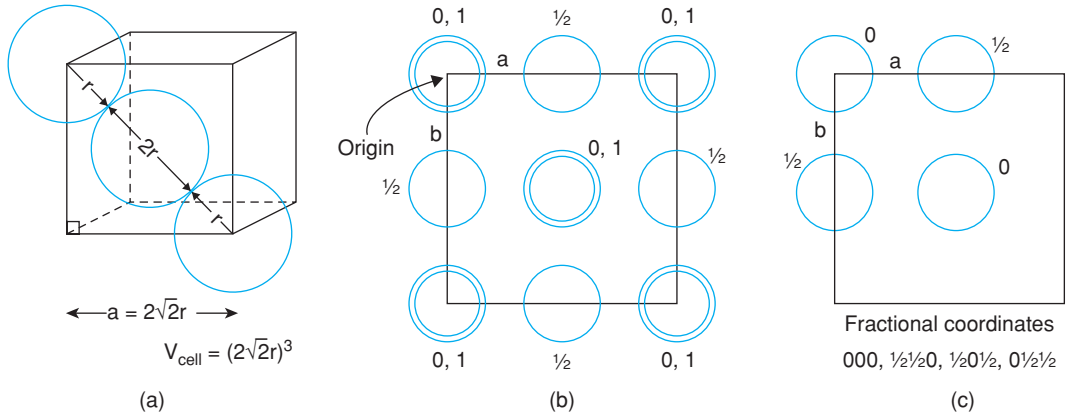


Figure 1.22 ***(a) Unit cell dimensions for a face centred cubic unit cell with spheres of radius r in contact along face diagonals. (b) Projection of a face centred cubic structure onto a unit cell face. (c) Unit cell contents.***

of the unit cell. In a *ccp* array of spheres, the *fcc* unit cell contains four spheres, one at a corner and three at face centre positions, Fig. 1.20 (this is equivalent to the statement that a *fcc* unit cell contains four lattice points). *cp* directions (xx′, yy′, zz′ in Fig. 1.16), in which spheres are in contact, occur parallel to the face diagonals of the unit cell, e.g. spheres 2, 5 and 6 in Fig. 1.20(b) form part of a *cp* row. The length of the face diagonal is therefore $4r$. From the Pythagoras theorem, the length of the cell edge is then $2\sqrt{2}r$ and the cell volume is $16\sqrt{2}r^3$, Fig. 1.22(a). The volume of each sphere is $1.33\pi r^3$ and so the ratio of the total sphere volume to the unit cell volume is given by

$$\frac{4 \times 1.33\pi r^3}{16\sqrt{2}r^3} = 0.7405 \tag{1.5}$$

Similar results are obtained for *hcp* by considering the contents and volume of the appropriate hexagonal unit cell, Fig. 1.21.

In non-*cp* structures, densities lower than 0.7405 are obtained, e.g. the density of body centred cubic, *bcc*, is 0.6802 (to calculate this it is necessary to know that the *cp* directions in *bcc* are parallel to the body diagonals, <111>, of the cube).

1.14 Unit Cell Projections and Atomic Coordinates

In order to give 3D perspective to crystal structures, they are often drawn as oblique projections, as in Fig. 1.20(a). For accurate and unambiguous descriptions, it is necessary, however, to project them down particular crystallographic directions and/or onto crystallographic faces, as shown for a face centred cube in Fig. 1.22(b) projected down the z axis onto the ab unit cell face. In this representation, the structure is projected onto a plane and so all sense of vertical perspective is lost. In order to restore some vertical perspective and to specify atomic positions fully, their vertical height in the unit cell is given, in this case as a fraction of c, beside each atom. It is not necessary to specify the x, y coordinates of each atom if the structure is drawn to scale; in such projections, the origin is shown here as the top left-hand corner.

In the projection shown, there are two atoms at each corner, with z coordinates of 0 and 1. Similarly, atoms at the top and bottom face centres are shown with $z = 0$, 1 in the middle of the projection. The four side face centre positions are shown as single circles, each with $z = {}^1/_2$.

It is important to be able to use diagrams such as that in Fig. 1.22(b) and to be able to relate these to listings of fractional atomic coordinates. Thus, a face centred cube contains, effectively, four positions in the unit cell, one corner and three face centres; their coordinates are 000, ${}^1/_2{}^1/_20$, ${}^1/_20{}^1/_2$, $0{}^1/_2{}^1/_2$: each coordinate specifies the fractional distance of the atom from the origin in the directions a, b and c, respectively of the unit cell. These four positions are shown in Fig. 1.22(c) and it should be clear that the more complete structure shown in (b) is obtained simply by the addition of extra, *equivalent positions* in adjacent unit cells.

The diagrams in Fig. 1.22(b) and (c) both, therefore, represent the unit cell of a face centred cube; if we wish to use the cell shown in (b), we must remember that only 1/8 of each corner atom belongs to this unit cell with the remaining 7/8 associated with surrounding unit cells. Similarly, edge centre atoms belong to four adjacent unit cells and, therefore, only 1/4 of each is in the unit cell shown. Finally, face centred atoms are split between two adjacent unit cells and therefore, each counts as half an atom in (b). If we use diagram (c) to represent our unit cell contents, then all of the four atoms shown belong entirely to this unit cell and, for instance, other corner atoms are simply the corner atoms of adjacent unit cells. Diagram (c) shows the bare minimum that is necessary to illustrate the crystal structure whereas diagram (b) is more informative and gives a much clearer perspective of the structure in 3D. Further details on describing unit cell contents, atomic coordinates and the possibility of negative atomic coordinates are given in Appendix E.

1.15 Materials That Can Be Described as Close Packed

1.15.1 Metals

Most metals crystallise in one of the three arrangements, *ccp*, *hcp* and *bcc*, the first two of which are *cp* structures. The distribution of structure type among the metals is irregular (Table 1.3), and no clear-cut trends are observed. It is still not well understood why particular metals prefer one structure type to another. Calculations reveal that the lattice energies of *hcp* and *ccp* metal structures are comparable and, therefore, the structure observed in a particular case probably depends on fine details of the bonding requirements or band structures of the metal. Structural details of elements not included in Table 1.3 are given in Appendix F.

1.15.2 Alloys

Alloys are intermetallic phases or solid solutions and, as is the case for pure metals, many have *cp* structures. For example, Cu and Au have *ccp* structures, both as pure elements and when mixed to form Cu–Au alloys.

Table 1.3 ***Structures and unit cell dimensions of some common metals***

ccp		*hcp*			*bcc*	
Metal	a/Å	Metal	a/Å	c/Å	Metal	a/Å
Cu	3.6147	Be	2.2856	3.5842	Fe	2.8664
Ag	4.0857	Mg	3.2094	5.2105	Cr	2.8846
Au	4.0783	Zn	2.6649	4.9468	Mo	3.1469
Al	4.0495	Cd	2.9788	5.6167	W	3.1650
Ni	3.5240	Ti	2.9506	4.6788	Ta	3.3026
Pb	4.9502	Zr	3.2312	5.1477	Ba	5.019

At high temperatures, a complete range of *solid solutions* between Cu and Au forms. In these the Cu and Au atoms are distributed statistically over the lattice points of the *fcc* unit cell and therefore the *ccp* layers contain a random mixture of Cu and Au atoms. On annealing the compositions AuCu and $AuCu_3$ at lower temperatures, the Au and Cu atoms *order* themselves; *ccp* layers still occur but the arrangement within the layers is no longer statistical. Such *order–disorder phenomena* occur commonly in both metallic and ionic structures.

1.15.3 Ionic structures

The structures of materials such as NaCl, Al_2O_3, Na_2O and ZnO, in which the anion is larger than the cation, are built of *cp* layers of anions with the cations placed in *interstitial sites*. Many structures are possible in which the variables are the anion stacking sequence, either *hcp* or *ccp*, and the number and type of interstitial sites occupied by cations. The cations are, however, often too large for the prescribed interstitial sites and the structure can accommodate them only by expanding the anion array. Consequently, the anion arrangement is the same as in *cp*, but the anions may not be in contact. O'Keeffe suggested the term *eutactic* for structures such as these. In the discussions that follow, use of the terms *hcp* and *ccp* for the anion arrays does not necessarily imply that the anions are in contact but rather that the structures are eutactic. A further complication, as we shall see later, is that the rigid sphere model is an oversimplification of reality since, in ionic structures, it can be difficult to specify ion sizes exactly.

1.15.3.1 Tetrahedral and octahedral sites

Two types of interstitial site, tetrahedral and octahedral, are present in *cp* structures, Fig. 1.23. These may be seen by considering the nature of the interstitial space between any pair of adjacent *cp* layers. For the tetrahedral sites, three anions that form the base of the tetrahedron belong to one *cp* layer with the apex of the tetrahedron either in the layer above (a) or below (b). This gives two orientations of tetrahedral sites, T_+ and T_-, in which the apex is up and down, respectively, and is illustrated further in (c) and (d) with the layers seen edge-on. Because the centre of gravity of a tetrahedron is nearer to the base than the apex (see Appendix C), cations in tetrahedral sites are not located midway between adjacent anion layers but are nearer one layer than the other. In *ccp* structures, there are four equivalent orientations of *cp* layers and therefore, four equivalent orientations of T_+, T_- sites

Octahedral sites, O, are coordinated to three anions in each layer, Fig. 1.23(e) and are midway between the anion layers (f). A more common way to regard octahedral coordination is as four coplanar atoms with one atom at each apex above and below the plane. In (e), atoms 1, 2, 4 and 6 are coplanar; 3 and 5 form apices of the octahedron. Also, atoms 2, 3, 4, 5 and 1, 3, 5, 6 are coplanar. This is further illustrated in Fig. 1.23(g) and (h), which are similar to (e) but seen from a different and more conventional perspective.

The distribution of interstitial sites between any two adjacent layers of *cp* anions is shown in Fig. 1.23(i). We can see that below every sphere in the upper layer is a T_+ site; likewise above every sphere in the lower layer is a T_- site. There are also as many O sites as either T_+ or T_- sites. A similar distribution to that in (i) is seen between each pair of *cp* anion layers. Counting up the numbers of each, then, for every anion there is one octahedral site and two tetrahedral sites, one T_+ and one T_-.

It is rare that all the interstitial sites in a *cp* structure are occupied; often one set is full or partly occupied and the remaining sets are empty. A selection of *cp* ionic structures, classified according to the anion layer stacking sequence and the occupancy of the interstitial sites, is given in Table 1.4. Individual structures are described later. Here we simply note how a wide range of structures are grouped into one large family; this helps to bring out similarities and differences between them. For example:

(a) The rock salt and nickel arsenide structures both have octahedrally coordinated cations and differ only in the anion stacking sequence. Similarly, there are other pairs of structures with similar cation coordination

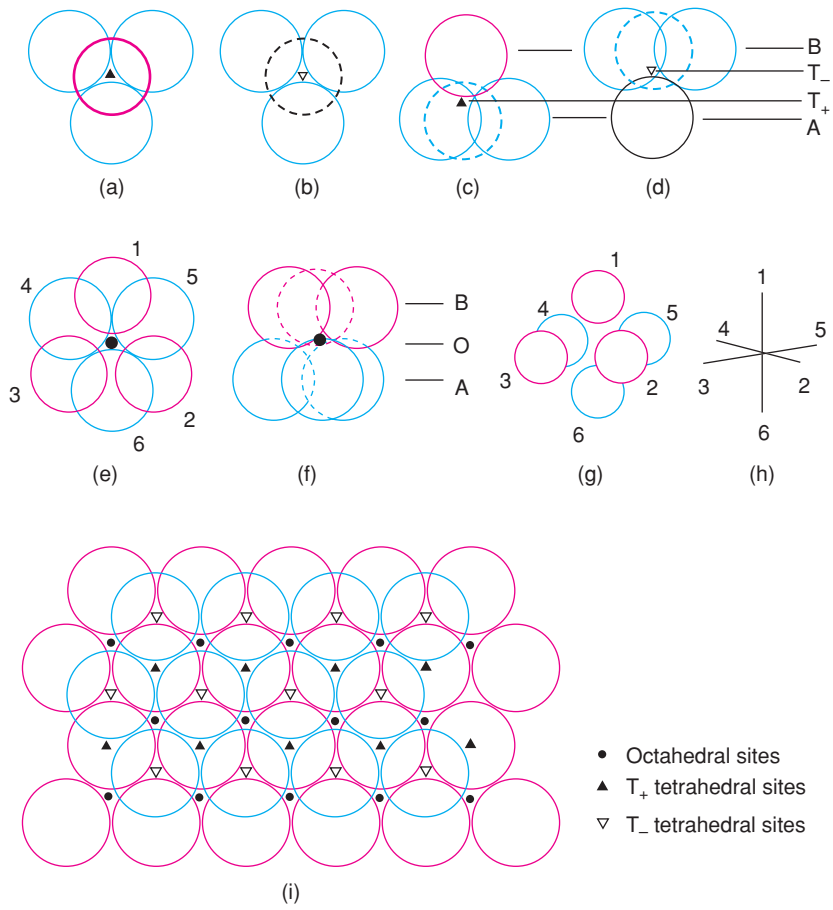


Figure 1.23 ***Tetrahedral and octahedral sites between two* cp *anion layers, seen from different perspectives. (a, b) Projection down threefold axis of T_+, T_- sites. (c, d) Tetrahedral sites edge-on. (e) Projection down threefold axis of octahedral site and (f) seen edge-on. (g, h) Conventional representation of octahedral site. (i) Distribution of T_+, T_-, O sites between two* cp *layers.***

numbers that differ only in the anion stacking sequence, e.g. olivine and spinel, wurtzite and zinc blende, CdI_2 and $CdCl_2$.

(b) Rutile, TiO_2, and CdI_2 both have *hcp* anions (although the layers are buckled in rutile) with half the octahedral sites occupied by cations but they differ in the manner of occupancy of these octahedral sites. In rutile, half the octahedral sites between any pair of *cp* anion layers are occupied by Ti. In CdI_2, layers of fully occupied octahedral sites alternate with layers in which all sites are empty; this gives CdI_2 a layered structure.

(c) In a few structures, it is useful to regard the *cation* as forming the *cp* layers with the anions occupying the interstitial sites. The fluorite structure, CaF_2, may be regarded as *ccp* Ca with all T_+ and T_- sites occupied by F. The antifluorite structure of K_2O is the exact inverse of fluorite with *ccp* layers of O and K in T_+, T_- sites; it is included in Table 1.4.

Table 1.4 ***Some close packed structures***

Anion arrangement	Interstitial sites			Examples
	T_+	T_-	O	
ccp	–	–	1	NaCl, rock salt
	1	–	–	ZnS, blende or sphalerite
	1/8	1/8	1/2	$MgAl_2O_4$, spinel
	–	–	1/2	$CdCl_2$
	–	–	1/3	$CrCl_3$
	1	1	–	K_2O, antifluorite
hcp	–	–	1	NiAs
	1	–	–	ZnS, wurtzite
	–	–	1/2	CdI_2
	–	–	1/2	TiO_2, rutile[a]
	–	–	2/3	Al_2O_3, corundum
	1/8	1/8	1/2	Mg_2SiO_4, olivine
ccp 'BaO_3' layers	–	–	1/4	$BaTiO_3$, perovskite

[a]The *hcp* oxide layers in rutile are not planar but are buckled; the oxide arrangement may alternatively be described as *tetragonal* packed, tp.

(d) The *cp* concept may be extended to structures in which a mixture of anions and large cations form the packing layers and smaller cations occupy interstitial sites. In perovskite, $BaTiO_3$, *ccp* layers of composition 'BaO_3' occur and one-quarter of the octahedral sites between these layers are occupied by Ti, although only those sites in which all six corners are O^{2-} ions.

(e) Some structures are anion-deficient *cp* structures in which the anions form an incomplete *cp* array. The ReO_3 structure has *ccp* oxide layers with one-quarter of the O sites empty. It is analogous to perovskite just described in which Ti is replaced by Re and Ba sites are left vacant. The structure of β-alumina, nominally of formula $NaAl_{11}O_{17}$, contains *cp* oxide layers in which every fifth layer has approximately three-quarters of the O^{2-} ions missing.

1.15.3.2 Relative sizes of tetrahedral and octahedral sites

A general guideline is that tetrahedra are smaller than octahedra which, in turn, are smaller than polyhedra of higher coordination number; for a given anion array, the relative sizes of sites may be obtained quantitatively by geometric calculations (see the next section). In ionic structures, cations M occupy those sites or polyhedra of most suitable size. Cations may occupy sites which, at first sight, appear too small, by pushing apart anions and expanding the structure (as in eutactic structures). By contrast, cations tend not to occupy sites that are too large, unless the structure can adjust itself, by twisting or distorting the anion array so as to reduce the size of the sites; this happens in many distorted perovskite structures in which cations are two small to occupy large 12-coordinate sites and a partial structural collapse occurs (see Section 1.17.7). Thus, cations do not occupy large sites in which they can 'rattle'. An interesting intermediate situation occurs in some structures

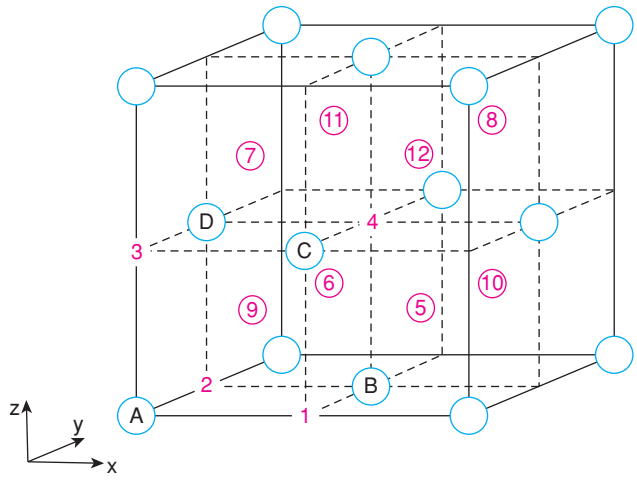


Figure 1.24 ***Available cation sites, 1–12, in an* fcc *anion array.***

in which a cation site is marginally too large. The cation can then undergo small off-centre displacements giving rise to high polarisability, high permittivity and the phenomenon of ferroelectricity (Section 8.7).

1.15.3.3 Location of tetrahedral and octahedral sites in an fcc *unit cell; bond length calculations*

We have seen how T_+, T_- and O sites occur between pairs of *cp* layers, Fig. 1.23, and how *fcc* and *ccp* structures are often equivalent, Fig. 1.20. We can now see where the T_+, T_- and O sites are located within an *fcc* unit cell that has anions, X, at corners and face centres (A, B, C, and D), Fig. 1.24. The octahedral sites are easiest to locate; they are at edge-centre 1, 2, 3 and body-centre 4 positions. If the unit cell has length a, the M–X distance for octahedral sites is $a/2$.

In order to see the T_+, T_- sites clearly, it is convenient to divide the unit cell into eight minicubes by bisecting each cell edge (dashed lines). These minicubes contain anions at only four of the eight corners; in the middle of each minicube is a tetrahedral site, either T_+ or T_-. There are two orientations for the minicubes, and therefore for the T sites, as shown in Fig. 1.25; relative to the parent unit cell, Fig. 1.24, the T_- tetrahedra have their apices pointing in the $[\bar{1}\bar{1}\bar{1}]$ direction and the T_+ sites in the [111] direction.

At this point, we note that it is exceptionally useful to be able to regard a tetrahedral site as occupying the centre of a cube with alternate corners absent, Fig. 1.25(a) and (b). For such an arrangement, it is easy to calculate bond lengths and bond angles (c) and to demonstrate the various symmetry elements associated with a tetrahedron. Thus, for the minicube of edge dimension $a/2$,

$$\mathrm{M-X} = \frac{1}{2}\left[3(a/2)^2\right]^{\frac{1}{2}} = a\sqrt{3}/4$$

Twofold rotation axes and fourfold inversion axes run parallel to minicube edges but pass through pairs of opposite faces; there are three of these in total (see also Fig. 1.7c); mirror planes occur parallel to the set

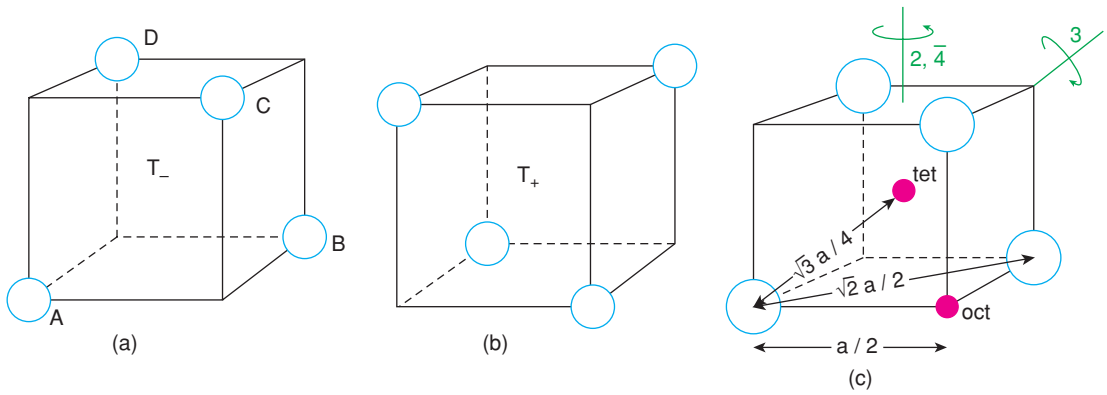


Figure 1.25 ***(a, b) Tetrahedral sites T_+, T_- and their relation to a cube. (c) Bond length calculations and some symmetries of a tetrahedron.***

of lattice planes with indices {110}, and there are six of these in total (as in Fig. 1.9c); threefold axes run parallel to <111> directions, four in total, Fig. 1.9a; further details are in Appendix C.

1.15.3.4 Description of crystal structures; fractional atomic coordinates

In order to describe crystal structures, it is necessary to specify (a) the unit cell type, (b) its dimensions and (c) the fractional coordinates of the atoms in the unit cell. As described earlier, an *fcc* unit cell containing anions at corner and face centre positions has, effectively, four anions in the unit cell, one at a corner and three at face centres. The fractional coordinates of the four anions A–D in the unit cell, Fig. 1.24, are

$$000,\ {}^1\!/_2{}^1\!/_2 0,\ {}^1\!/_2 0{}^1\!/_2,\ 0{}^1\!/_2{}^1\!/_2$$

Only one corner, at 000, is included since the other seven, e.g. at 100, 010, etc., are equivalent and can be regarded as the corner atoms of adjacent cells. The alternative would be to list all eight corner atoms, with the qualification that only 1/8 of each belonged to the unit cell in question. This would be equivalent to, but much more cumbersome than, regarding one corner only as specifically belonging, completely, to our unit cell. Likewise, for each pair of opposite faces, e.g. ${}^1\!/_2{}^1\!/_2 0$ and ${}^1\!/_2{}^1\!/_2 1$, it is most convenient to regard ${}^1\!/_2{}^1\!/_2 0$ as belonging completely to our unit cell; ${}^1\!/_2{}^1\!/_2 1$ then becomes the bottom face centre in the cell above.

The various cation positions in Fig. 1.24 have the following coordinates:

octahedral	1: ${}^1\!/_2 00$	2: $0{}^1\!/_2 0$	3: $00{}^1\!/_2$	4: ${}^1\!/_2\ {}^1\!/_2\ {}^1\!/_2$
tetrahedral, T_+	5: ${}^3\!/_4\ {}^1\!/_4\ {}^1\!/_4$	6: ${}^1\!/_4\ {}^3\!/_4\ {}^1\!/_4$	7: ${}^1\!/_4\ {}^1\!/_4\ {}^3\!/_4$	8: ${}^3\!/_4\ {}^3\!/_4\ {}^3\!/_4$
tetrahedral, T_-	9: ${}^1\!/_4\ {}^1\!/_4\ {}^1\!/_4$	10: ${}^3\!/_4\ {}^3\!/_4\ {}^1\!/_4$	11: ${}^1\!/_4\ {}^3\!/_4\ {}^3\!/_4$	12: ${}^3\!/_4\ {}^1\!/_4\ {}^3\!/_4$

Note that there are four of each type of cation site, O, T_+, T_-, in the unit cell, together with four anions. When different sites are fully or partially occupied by the cations, different structures are generated, as indicated in Table 1.4 and discussed later.

1.15.4 Covalent network structures

Materials such as diamond and silicon carbide, which have very strong, directional, covalent bonds, can also be described as *cp* structures or eutactic structures; many have the same structures as ionic compounds. Thus, one polymorph of SiC has the wurtzite structure and it is immaterial whether Si or C is regarded as the packing atom since the net result, a 3D framework of corner-sharing tetrahedra, is the same. Diamond can be regarded as a sphalerite structure in which half of the C atoms form a *ccp* array and the other half occupy T_+ sites, but again the two types of atom are equivalent. Classification of diamond as a eutactic structure is useful since in diamond all atoms are of the same size and it is unrealistic to distinguish between packing atoms and interstitial atoms.

Many structures have mixed ionic–covalent bonding, e.g. ZnS and $CrCl_3$; one advantage of describing them in terms of *cp* is that this can be done, if necessary, without reference to the type of bonding that is present.

1.15.5 Molecular structures

Since *cp* structures provide an efficient means of packing atoms, many molecular compounds crystallise as *cp* structures even though the bonding forces between adjacent molecules are weak van der Waals forces. If the molecules are roughly spherical or become spherical because they can rotate or occupy different orientations at random, then simple *hcp* or *ccp* structures result, e.g. in crystalline H_2, CH_4 and HCl. Non-spherical molecules, especially if they are built of tetrahedra and octahedra, can also fit into a *cp* arrangement. For example, Al_2Br_6 is a dimer with two $AlBr_4$ tetrahedra sharing a common edge, Fig. 1.26(a). In crystalline Al_2Br_6, the Br atoms form an *hcp* array and Al atoms occupy one-sixth of the available tetrahedral sites. One molecule, with the Br atoms in heavy outline is shown in Fig. 1.26(b); Al atoms occupy one pair of adjacent T_+ and T_- sites. Br atoms 3 and 5 are common to both tetrahedra and are the bridging atoms in (a). Adjacent Al_2Br_6 molecules are arranged so that each Br in the *hcp* array belongs to only one molecule. $SnBr_4$ is a tetrahedral molecule and also crystallises with an *hcp* Br array, but only one-eighth of the tetrahedral sites are occupied.

1.15.6 Fullerenes and fullerides

The simplest, best known and symmetrical *fullerene* is C_{60}. It is a hollow, spherical cage of diameter 7.1 Å formed from 3-coordinate (approximately sp^2 hybridised) C atoms which link to form a network

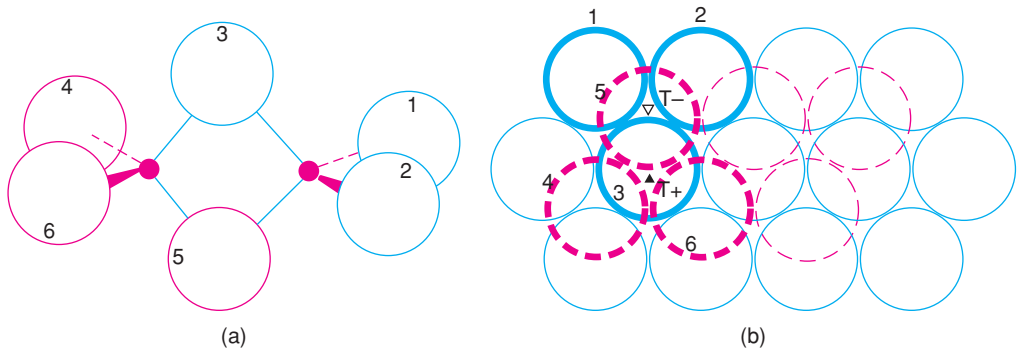


Figure 1.26 ***(a)*** **hcp** ***arrangement of Br atoms in crystalline Al_2Br_6; (b) Al atoms occupy T_+ and T_- sites. Dashed circles are below the plane of the paper.***

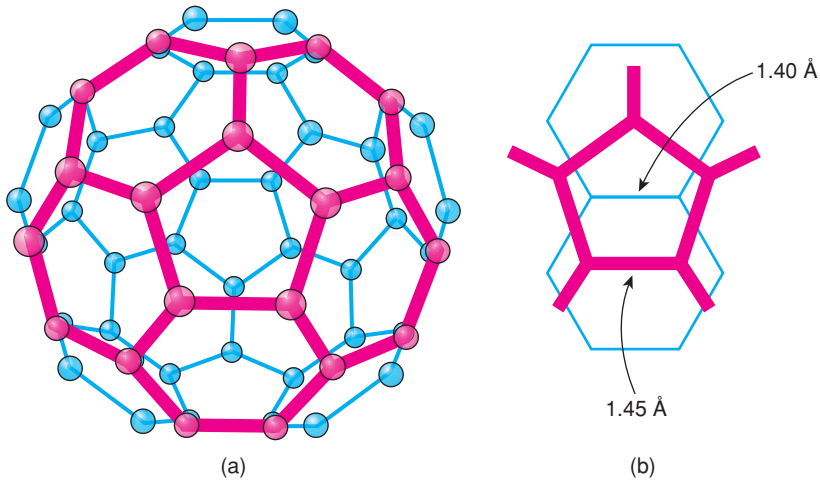


Figure 1.27 ***(a) The C_{60} molecule; one pentagon surrounded by five hexagons is shown in bold. (b) Parts of two C_{60} molecules in the region of closest contact showing a short C–C bond separating two hexagons in one molecule pointing towards the middle of a pentagon in an adjacent molecule. Adapted from David* et al.*,* Nature*, 353, 147, © 1991 Macmillan Publishers Ltd. Adapted under the terms of the STM Agreement.***

of 12 pentagons and 20 hexagons, Fig. 1.27(a), that match exactly the pattern on a soccer ball (hence the alternative, colloquial, American name 'buckyball'; the name 'fullerene' is in honour of Buckminster Fuller, who designed the geodesic dome that has the same network structure as C_{60}).

Not surprisingly, in crystalline C_{60}, the C_{60} molecules arrange themselves according to the principles of close packing. At room temperature, crystalline C_{60} is face centred cubic, $Z = 4$ (C_{60}), $a = 14.17$ Å and the C_{60} molecules are orientationally disordered. The separation of adjacent C_{60} molecules can be calculated with the aid of Fig. 1.22(a); a cube face diagonal corresponds to a *cp* direction from which we calculate that the centres of adjacent molecules are separated by 10.0 Å. The 'hard diameter' of a C_{60} molecule is calculated to be 7.1 Å which leaves a gap of 2.9 Å between adjacent molecules for van der Waals bonding.

Because the C_{60} molecules are not perfectly spherical, Fig. 1.27(a), they settle into an ordered arrangement below 249 K. The driving force for this ordering is optimisation of the bonding between adjacent C_{60} molecules. In particular, in a given [110] *cp* direction, the short electron-rich C–C bond that links two hexagons in one C_{60} molecule points directly at the centre of an electron-deficient pentagon, on an adjacent C_{60} molecule, as shown for two ring fragments in Fig. 1.27(b).

This arrangement minimises direct C–C overlap but maximises donor–acceptor electronic interactions between adjacent molecules. The C–C bonds in fullerene fall into two groups, 1.40 Å for C–C bonds linking two hexagons and 1.45 Å for bonds linking one hexagon and one pentagon; pentagons are isolated in C_{60}. Comparing these with typical single and double bond lengths in organic compounds, 1.54 and 1.33 Å, respectively, we can see that the C–C bonds in C_{60} are of intermediate bond order.

Some other fullerenes, such as C_{70}, also form *cp* structures, even though the molecules are not spherical. Thus, the C_{70} molecule is an ellipsoid (shaped like a rugby ball with a long axis of 8.34 Å and a short axis of 7.66 Å), but in crystalline C_{70} at room temperature the ellipsoids are able to rotate freely to give quasispherical molecules on a time average which form an *fcc* structure, $a = 15.01$ Å.

cp structures have, of course, tetrahedral and octahedral interstitial sites and in C_{60} these may be occupied by a range of large alkali metal cations to give materials known as *fullerides*. The most studied C_{60} fullerides

have general formula A_3C_{60} as in Rb_3C_{60} or K_2RbC_{60}, in which all T_+, T_- and O sites are occupied. These materials are metallic since the alkali metals ionise and donate their electrons to the conduction band of the C_{60} network, which is half full in A_3C_{60}. At low temperatures, many of these fullerides become superconducting. The highest T_c (the temperature at which the metal–superconductor transition occurs on cooling, Chapter 8) to date in the fullerides is 45 K in Tl_2RbC_{60}.

Other fullerides exist with different patterns of occupancy of interstitial sites, e.g. rock salt or zinc blende analogues in AC_{60} and fluorite analogues in A_2C_{60}. It is also possible to increase the number of A ions per C_{60} beyond 3; in A_4C_{60} (A = Na), A_4 clusters form in the octahedral sites. In A_6C_{60}, the packing arrangement of the C_{60} molecules changes from *ccp* to *bcc*; the *bcc* structure has a large number of interstitial sites, e.g. six distorted tetrahedral sites, distributed over the cube faces in the *bcc* array of anions (discussed in Chapter 8, for α-AgI, Fig. 8.28). This wide range of formulae arises because the C_{60} molecule can accept a number of electrons to give anions C_{60}^{n-}: $n = 1, 2, 3, 4, \ldots$ Part of the excitement in C_{60} chemistry is that the electronic properties range from insulating to metallic/superconducting, depending on the value of n.

1.16 Structures Built of Space-Filling Polyhedra

This approach to describing crystal structures emphasises the coordination number of cations. Structures are regarded as built of polyhedra, formed by the cations with their immediate anionic neighbours, which link by sharing corners, edges or faces. For example, in NaCl each Na has six Cl nearest neighbours arranged octahedrally; this is represented as an octahedron with Cl at the corners and Na at the centre. A 3D overview of the structure is obtained by looking at how neighbouring octahedra link to each other. In NaCl, each octahedron edge is shared between two octahedra (Fig. 1.31, see later), resulting in an infinite framework of edge-sharing octahedra. In perovskite, $SrTiO_3$, TiO_6 octahedra link by corner-sharing to form a 3D framework (Fig. 1.41, see later). These are just two examples; there are very many others, involving mainly tetrahedra and octahedra, leading to a wide diversity of structures. Some are listed in Table 1.5.

Although it is legitimate to estimate the efficiency of packing spheres in *cp* structures, we cannot do this for space-filling polyhedra since the anions, usually the largest ions in the structure, are represented by points at the corners of the polyhedra. In spite of this obvious misrepresentation, the space-filling polyhedron approach shows the topology or connectivity of a framework structure and indicates clearly the location of empty interstitial sites.

Table 1.5 ***Some structures built of space-filling polyhedra***

Structure	Example
Octahedra only	
12 edges shared	NaCl
6 corners shared	ReO_3
3 edges shared	$CrCl_3$, BiI_3
2 edges and 6 corners shared	TiO_2
4 corners shared	$KAlF_4$
Tetrahedra only	
4 corners shared (between 4 tetrahedra)	ZnS
4 corners shared (between 2 tetrahedra)	SiO_2
1 corner shared (between 2 tetrahedra)	$Si_2O_7^{6-}$
2 corners shared (between 2 tetrahedra)	$(SiO_3)_n^{2n-}$, chains or rings

A complete scheme for classifying polyhedral structures has been developed by Wells and others. The initial problem is a geometric one: what types of network built of linked polyhedra are possible? The variables are as follows. Polyhedra may be tetrahedra, octahedra, trigonal prisms, etc. Polyhedra may share some or all of their corners, edges and faces with adjacent polyhedra, which may or may not be of the same type. Corners and edges may be common to two or more polyhedra (obviously only two polyhedra can share a common face). An enormous number of structures are feasible, at least theoretically, and it is an interesting exercise to categorise real structures on this basis.

The topological approach to arranging polyhedra in various ways takes no account of the bonding forces between atoms or ions. Such information must come from elsewhere. Also, description of structures in terms of polyhedra does not necessarily imply that such entities exist as separate species. Thus, in NaCl, the bonding is mainly ionic, and physically distinct $NaCl_6$ octahedra do not occur. Similarly, SiC has a covalent network structure and separate SiC_4 tetrahedral entities do not exist. Polyhedra *do* have a separate existence in structures of (a) molecular materials, e.g. Al_2Br_6 contains pairs of edge-sharing tetrahedra, and (b) compounds that contain complex ions, e.g. silicate structures built of SiO_4 tetrahedra which form complex anions ranging in size from isolated monomers to infinite chains, sheets and 3D frameworks.

In considering the polyhedral linkages that are likely to occur, Pauling's third rule for the structures of complex ionic crystals (see the next chapter for other rules) provides a useful guideline. This states that shared edges and, especially, shared faces decrease the stability of a structure, particularly for cations of high valence and small coordination number, i.e. for small polyhedra, especially tetrahedra, that contain highly charged cations. When polyhedra share edges and, particularly, faces, the cation–cation distances (i.e. the distances between the centres of the polyhedra) decrease and the cations repel each other electrostatically. In Fig. 1.28 are shown pairs of (a) corner-sharing and (b) edge-sharing octahedra. The cation–cation distance is

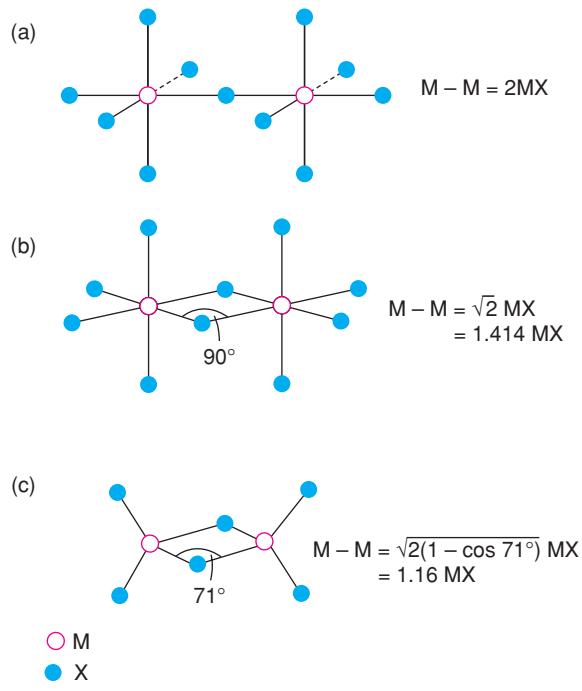


Figure 1.28 ***Cation–cation separation in octahedra which share (a) corners and (b) edges and in (c) tetrahedra which share edges.***

Table 1.6 ***The M–M distance between MX_4 or MX_6 groups sharing X atom(s)***

	Corner sharing	Edge sharing	Face sharing
Two tetrahedra	2.00 MX(tet.)[a]	1.16 MX(tet.)[a]	0.67 MX(tet.)
Two octahedra	2.00 MX(oct.)[a]	1.41 MX(oct.)[a]	1.16 MX(oct.)

[a]Maximum possible value.

clearly less in the latter and for face-sharing octahedra, not shown, the cations are even closer. Comparing edge-shared octahedra and edge-shared tetrahedra (b, c), the cation–cation distance M–M (relative to the cation–anion distance M–X) is less between the tetrahedra because the MXM bridging angle is 71° for them and 90° for the octahedra. A similar effect occurs for face-sharing tetrahedra and octahedra. In Table 1.6, the M–M distances for various polyhedral linkages are given as a function of the M–X distance. The M–M distance is greatest for corner-sharing tetrahedra or octahedra and is least for face-sharing tetrahedra. For the sharing of corners and edges, the M–M distances given are the maximum possible. Reduced distances occur if the polyhedra are rotated about the shared corner or edge (e.g. so that the M–X–M angle between corner-shared polyhedra is less than 180°).

From Table 1.6, the M–M distance in face-shared tetrahedra is considerably *less than* the M–X bond distance. This therefore represents an unstable situation, because of the large cation–cation repulsions which would arise, and does not normally occur. A good example of the non-occurrence of face-sharing tetrahedra is provided by the absence of any crystal whose structure is the *hcp* equivalent of the fluorite (or antifluorite) structure. In Na_2O, which has the antifluorite structure, the oxide ions are in a *ccp* array and the NaO_4 tetrahedra share edges. However, in a structure with an *hcp* anion array and all the tetrahedral cation sites occupied, the MX_4 tetrahedra would share faces. Edge sharing of tetrahedra, in which the M–M separation is only 16% larger than the M–X bond distance (Table 1.6) is energetically acceptable, as shown by the common occurrence of compounds with a fluorite structure, but face sharing of tetrahedra is generally unacceptable.

For tetrahedra containing cations of high charge, edge sharing may be energetically unacceptable and only corner sharing occurs. For example, in silicate structures which are built of SiO_4 tetrahedra, edge sharing of SiO_4 tetrahedra never occurs. (*Note*: In an ideally ionic structure, the charge on Si would be 4+, but the actual charges are considerably less due to partial covalency of the Si–O bonds.) In SiS_2, however, edge sharing of SiS_4 tetrahedra does occur; the Si–S bond is longer than the Si–O bond and consequently the Si–Si distance in edge-shared SiS_4 tetrahedra is also greater and appears to be within the acceptable range for interatomic separations.

1.17 Some Important Structure Types

1.17.1 Rock salt (NaCl), zinc blende or sphalerite (ZnS), fluorite (CaF_2), antifluorite (Na_2O)

These structures are considered together because they all have *ccp/fcc* anions and differ only in the positions of the cations. In Fig. 1.24 are shown the anions in a *fcc* unit cell with all possible O, T_+ and T_- sites for the cations. There is no rule as to which sites should be labelled T_+ and T.; the choice is yours. The choice of origin is also arbitrary; for present purposes, it is more convenient to place the anions at the origin and also, therefore, at face centre positions. The different structures are generated as follows:

- rock salt: O occupied; T_+ and T_- empty
- zinc blende: T_+ (or T_-) occupied; O, T_- (or T_+) empty
- antifluorite: T_+, T_- occupied; O empty.

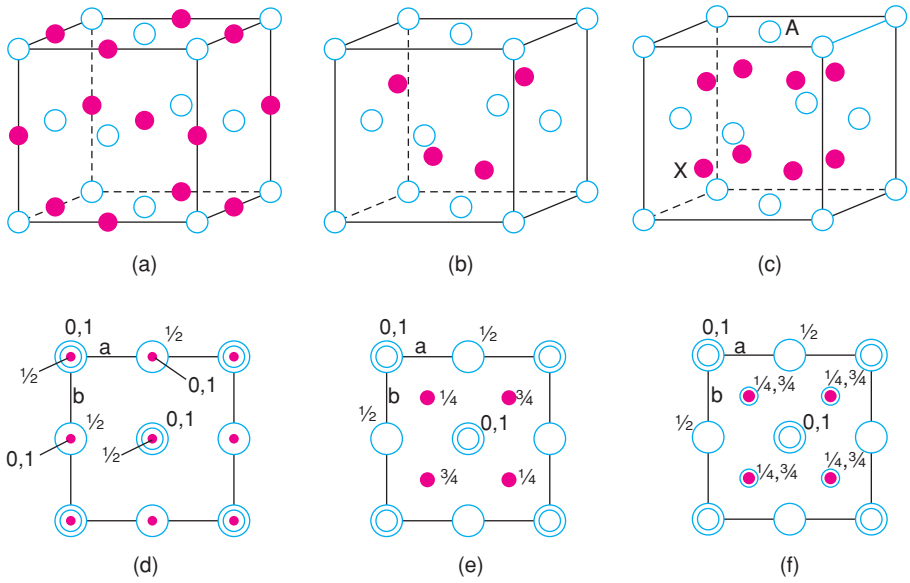


Figure 1.29 ***Unit cell of (a, d) NaCl, (b, e) ZnS, sphalerite, and (c, f) Na_2O. Closed circles are cations; open circles are anions.***

Unit cells are shown in Fig. 1.29, in oblique projection (a–c) and as projections on the *ab* face of the unit cell in (d–f). Each is described in more detail later.

A general rule regarding coordination numbers is that in any structure of formula A_xX_y, the coordination numbers of A and X must be in the ratio of y:x. In both rock salt and zinc blende, $x = y$ and therefore, in each, anions and cations have the same coordination number.

In antifluorite, of formula A_2X, the coordination numbers of cation and anion must be in the ratio of 1:2. Since the cations occupy tetrahedral sites, the anion coordination number must be eight. In order to see this, it is convenient to redefine the origin of the unit cell to coincide with a cation rather than an anion. This is done by displacing the unit cell along a body diagonal by one-quarter of the length of the diagonal. The cation at X in Fig. 1.29(c), with coordinates $^1/_4\ ^1/_4\ ^1/_4$, may be chosen as the new origin of the unit cell. The coordinates of all atoms in the new cell are given by subtracting $^1/_4\ ^1/_4\ ^1/_4$ from their coordinates in the old cell, as in Table 1.7.

In cases where negative coordinates occur as a result of this subtraction, e.g. $-^1/_4 - ^1/_4 - ^1/_4$, the position lies outside the new unit cell and it is necessary to find an equivalent position within the unit cell. In this particular case, 1 is added to each coordinate, giving $^3/_4\ ^3/_4\ ^3/_4$. Addition of 1 to, say, the x coordinate is equivalent to

Table 1.7 ***Two ways to describe the antifluorite structure***

	Old cell	New cell
Anions	000, $^1/_2\,^1/_2 0$, $^1/_2 0\,^1/_2$, $0\,^1/_2\,^1/_2$	$^3/_4\ ^3/_4\ ^3/_4$, $^1/_4\ ^3/_4\ ^3/_4$, $^1/_4\ ^3/_4\ ^1/_4$, $^3/_4\ ^1/_4\ ^1/_4$
Cations	$^1/_4\ ^1/_4\ ^1/_4$, $^1/_4\ ^1/_4\ ^3/_4$, $^1/_4\ ^3/_4\ ^1/_4$, $^3/_4\ ^1/_4\ ^1/_4$	000, $00\,^1/_2$, $0\,^1/_2 0$, $^1/_2 00$
	$^1/_4\ ^3/_4\ ^3/_4$, $^3/_4\ ^1/_4\ ^3/_4$, $^3/_4\ ^3/_4\ ^1/_4$, $^3/_4\ ^3/_4\ ^3/_4$	$0\,^1/_2\,^1/_2$, $^1/_2 0\,^1/_2$, $^1/_2\,^1/_2 0$, $^1/_2\,^1/_2\,^1/_2$

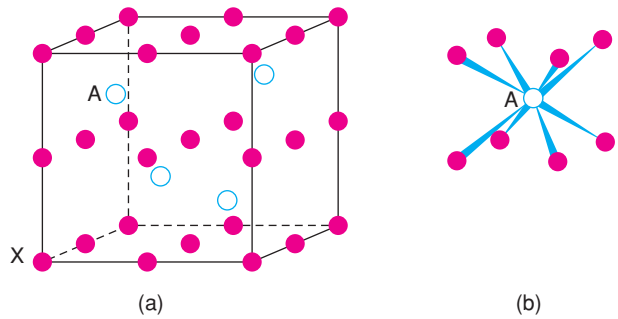


Figure 1.30 ***Alternative view of the antifluorite structure.***

moving to a similar position in the next unit cell in the *x* direction. The new unit cell of antifluorite with its origin at cation X, is shown in Fig. 1.30(a). It contains cations at corners, edge centres, face centres and body centre.

In order to see the anion coordination more clearly, the unit cell may be imagined as divided into eight minicubes (as in Fig. 1.24). Each minicube in Fig. 1.30(a) has cations at all eight corners and at the centre of each is an eight-coordinate site. Anions occupy four of these eight minicubes such that parallel to the cell axes the eight-coordinate sites are alternately occupied and empty. The eightfold coordination for one anion, A, is shown in Fig. 1.30(b).

In the antifluorite structure, the effect of changing the origin from an anion to a cation is to show the structure in a completely different light. This does not happen with the rock salt and zinc blende structures. In these, the cation and anion positions are interchangeable and it is immaterial whether the origin coincides with an anion or a cation.

So far, the NaCl, ZnS and Na_2O structures have been described in two ways: (a) as *cp* structures and (b) in terms of their unit cells. A third way is to regard them as built of space-filling polyhedra. Each ion and its nearest neighbours may be represented by the appropriate polyhedron, e.g. in zinc blende a tetrahedron represents one Zn with four S neighbours (or vice versa). It is then necessary to consider how neighbouring polyhedra are linked in 3D. Let us now consider each of these structures in more detail.

1.17.1.1 Rock salt structure

To summarise, the rock salt structure has *ccp/fcc* anions with octahedral sites fully occupied by cations and tetrahedral sites empty. Each cation is surrounded by six anions, as shown for cations 1, 2 and 3 in Fig. 1.31; similarly, each anion is octahedrally coordinated by cations (to see this, consider the anion at the top face centre, $^1/_2\,^1/_2\,1$, which is coordinated to the four edge centre cations in the top face, to the cation at the body centre, shown, and to a sixth cation at the body centre of the unit cell above).

The ($NaCl_6$) or ($ClNa_6$) octahedra share common edges, Fig. 1.31. Since each has 12 edges and each edge is common to two octahedra, it is difficult to represent this satisfactorily in a drawing; Fig. 1.31 shows just two such linkages. A perspective which focuses on the 3D array of octahedra is shown in Fig. 1.32; the structure may be regarded as layers of octahedra, which also have an ABC stacking sequence similar to the anions. Each octahedron face is parallel to a *cp* layer of anions, as emphasised by the numbering or shading of coplanar faces. Parts of four different sets of faces are seen, corresponding to the four *cp* orientations in a *ccp/fcc* array. Also shown, arrowed, in Fig. 1.32 are the empty tetrahedral sites.

A large number of AB compounds possess the rock salt structure. A selection is given in Table 1.8 together with values of the *a* dimension of the cubic unit cell. Most halides and hydrides of the alkali metals and Ag^+

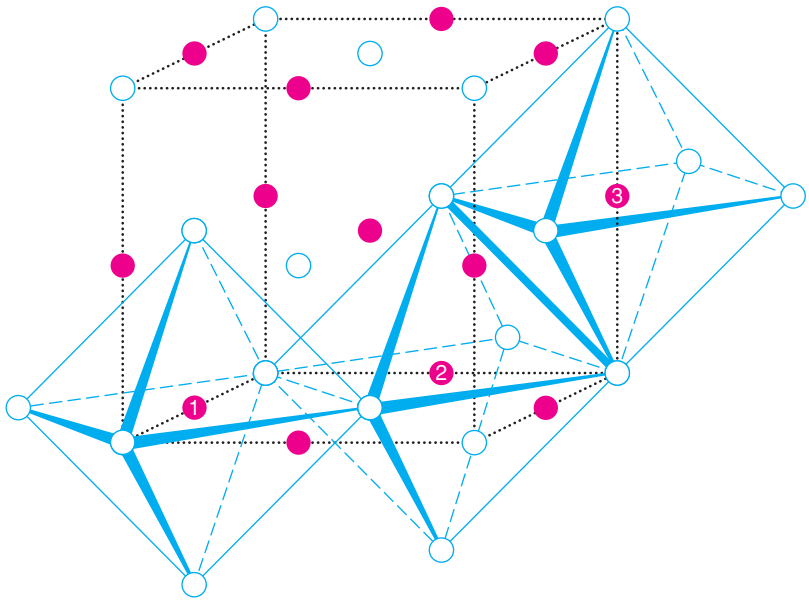


Figure 1.31 ***Unit cell of the rock salt structure showing edge-sharing octahedra.***

have this structure, as do a large number of chalcogenides (oxides, sulfides, etc.) of divalent metals. Many are ionic but others are either metallic, e.g. TiO, or covalent, e.g. TiC.

1.17.1.2 Zinc blende (sphalerite) structure

The zinc blende structure has *ccp/fcc* anions with cations in one set of tetrahedral sites, either T_+ or T_-. The ZnS_4 tetrahedra are linked at their corners and each corner is common to four such tetrahedra. The unit cell

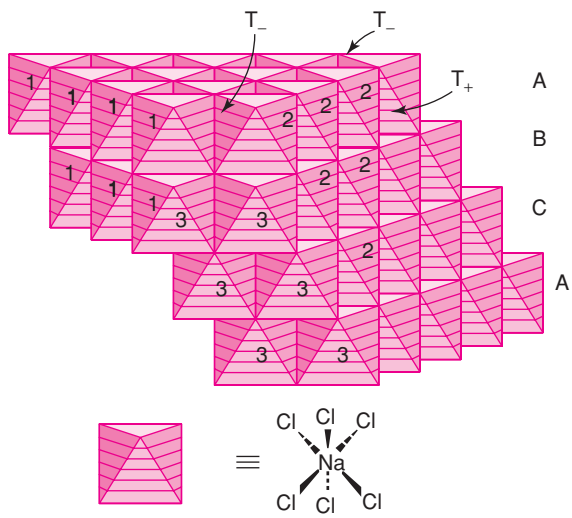


Figure 1.32 ***The rock salt structure as an array of edge-sharing octahedra.***

Table 1.8 ***Some compounds with the NaCl structure, a/Å***

MgO	4.213	MgS	5.200	LiF	4.0270	KF	5.347
CaO	4.8105	CaS	5.6948	LiCl	5.1396	KCl	6.2931
SrO	5.160	SrS	6.020	LiBr	5.5013	KBr	6.5966
BaO	5.539	BaS	6.386	LiI	6.00	KI	7.0655
TiO	4.177	α-MnS	5.224	LiH	4.083	RbF	5.6516
MnO	4.445	MgSe	5.462	NaF	4.64	RbCl	6.5810
FeO	4.307	CaSe	5.924	NaCl	5.6402	RbBr	6.889
CoO	4.260	SrSe	6.246	NaBr	5.9772	RbI	7.342
NiO	4.1769	BaSe	6.600	NaI	6.473	AgF	4.92
CdO	4.6953	CaTe	6.356	TiN	4.240	AgCl	5.549
TiC	4.3285	LaN	5.30	UN	4.890	AgBr	5.7745

of zinc blende, Fig. 1.29(b), is shown again in Fig. 1.33(a), but in terms of corner-sharing ZnS_4 tetrahedra. The faces of the tetrahedra are parallel to the *cp* anion layers, i.e. the $\{111\}$ planes; this is emphasised in a more extensive model in Fig. 1.33(b), which is oriented so that one set of tetrahedron faces is approximately horizontal. Conventionally, ZnS is regarded as built of *cp* layers of sulfide anions with the smaller zinc cations in tetrahedral sites. Since the same structure is generated by interchanging the Zn and S, the structure could also be described as a *cp* array of Zn with S occupying one set of tetrahedral sites. A third, equivalent description is as an array of *ccp* ZnS_4 (or SZn_4) tetrahedra.

Some compounds with the zinc blende structure are listed in Table 1.9. The bonding is less ionic than in compounds with the rock salt structure. Thus, oxides do not usually have the zinc blende structure (ZnO, not included in Table 1.9, is an exception; it is dimorphic with zinc blende and wurtzite structure polymorphs). Chalcogenides of the alkaline earth metals (but not Be) have the rock salt structure whereas the chalcogenides of Be, Zn, Cd and Hg have the zinc blende structure, as do Cu(I) halides and γ-AgI. A range of III–V compounds (i.e. elements from Groups III and V of the Periodic Table) have the zinc blende structure and some, e.g. GaAs, are important semiconductors.

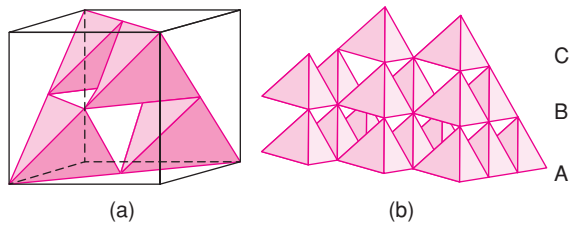


Figure 1.33 ***The sphalerite (zinc blende) structure showing (a) the unit cell contents and (b) a more extended network of corner-sharing tetrahedra.***

1.17.1.3 Antifluorite/fluorite structure

The antifluorite structure has *ccp/fcc* anions with cations in all (T_+ and T_-) tetrahedral sites. The difference between antifluorite and fluorite is that antifluorite refers to an anion array with tetrahedral cations, whereas fluorite has the inverse arrangement with a *ccp* cation array and tetrahedral anions. Since the cation:anion ratio is 2:1 in antifluorite and the cation coordination is 4, the anion coordination must be 8, Fig. 1.30.

The very different coordination environments of anions and cations leads to two entirely distinct descriptions of the structure in terms of a 3D network of either tetrahedra or cubes, Fig. 1.34; (a) corresponds to the

Table 1.9 ***Some compounds with the zinc blende (sphalerite) structure, a/Å***

CuF	4.255	BeS	4.8624	β-CdS	5.818	BN	3.616	GaP	5.448
CuCl	5.416	BeSe	5.07	CdSe	6.077	BP	4.538	GaAs	5.6534
γ-CuBr	5.6905	BeTe	5.54	CdTe	6.481	BAs	4.777	GaSb	6.095
γ-CuI	6.051	β-ZnS	5.4060	HgS	5.8517	AlP	5.451	InP	5.869
γ-AgI	6.495	ZnSe	5.667	HgSe	6.085	AlAs	5.662	InAs	6.058
β-MnS, red	5.600	β-SiC	4.358	HgTe	6.453	AlSb	6.1347	InSb	6.4782
C[a]	3.5667	Si	5.4307	Ge	5.6574	α-Sn (grey)	6.4912		

[a] Diamond structure.

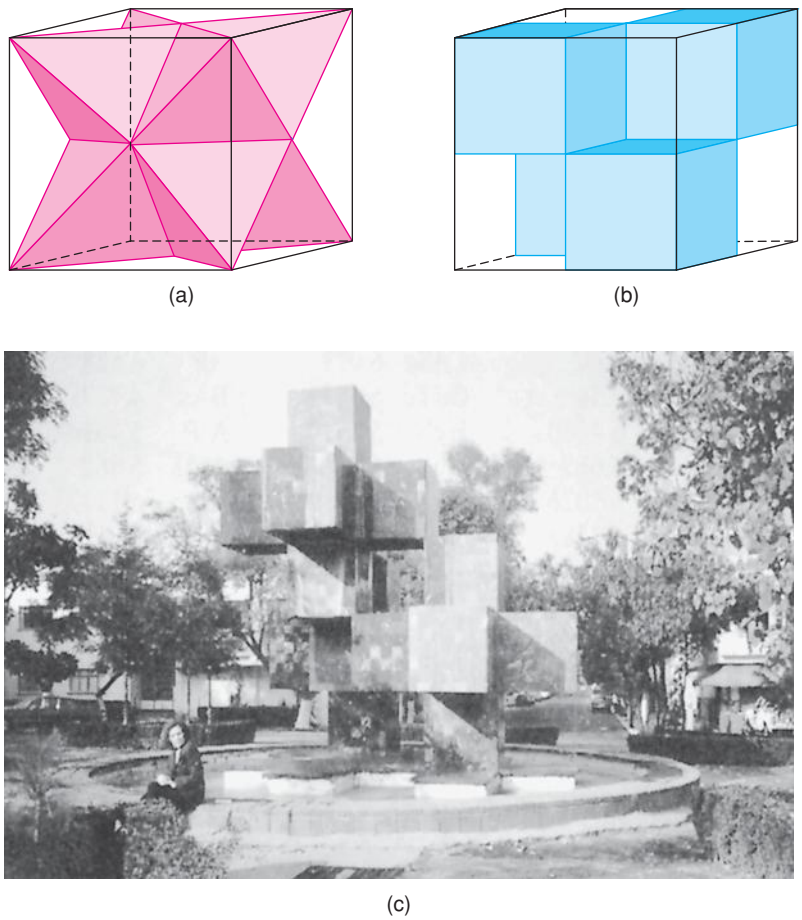


Figure 1.34 ***The antifluorite structure showing the unit cell in terms of (a) NaO_4 tetrahedra and (b) ONa_8 cubes. A more extended array of cubes is shown in (c); this model resides on a roundabout in Mexico City.***

Table 1.10 ***Some compounds with fluorite or antifluorite structure, a/Å***

Fluorite structure				Antifluorite structure			
CaF_2	5.4626	PbO_2	5.349	Li_2O	4.6114	K_2O	6.449
SrF_2	5.800	CeO_2	5.4110	Li_2S	5.710	K_2S	7.406
$SrCl_2$	6.9767	PrO_2	5.392	Li_2Se	6.002	K_2Se	7.692
BaF_2	6.2001	ThO_2	5.600	Li_2Te	6.517	K_2Te	8.168
CdF_2	5.3895	UO_2	5.372	Na_2O	5.55	Rb_2O	6.74
β-PbF_2	5.940	NpO_2	5.4334	Na_2S	6.539	Rb_2S	7.65

arrangement shown in Fig. 1.29(c) and the tetrahedra are highlighted; (b) corresponds to the arrangement in Fig. 1.30(b) in which the cubic coordination arrangement is highlighted. A more extended network of corner- and edge-sharing cubes is shown in Fig. 1.34(c). This must surely rate as one of the world's largest models of the antifluorite structure!

The *antifluorite* structure is shown by a large number of oxides and other chalcogenides of the alkali metals (Table 1.10), i.e. compounds of general formula $A_2^+X^{2-}$. A group of fluorides of large, divalent cations and oxides of large tetravalent cations have the inverse, *fluorite*, structure, i.e. $M^{2+}F_2$ and $M^{4+}O_2$.

From Fig. 1.34(b) and (c), an alternative way of describing the fluorite structure is as a primitive cubic array of anions in which the eight-coordinate sites at the cube body centres are alternately empty and occupied by a cation. It should be stressed that the true lattice type of fluorite is *fcc* and not primitive cubic, since the primitive cubes represent only a small part (one-eighth) of the *fcc* unit cell. Description of fluorite as a primitive cubic array of anions with alternate cube body centres occupied by cations shows a similarity to the CsCl structure (see later). This also has a primitive cubic array of anions, but, instead, cations occupy *all* the body centre sites.

Most fluorite structures are eutactic in terms of descriptions as both a primitive cubic array of anions and an *fcc* array of cations. Thus, the anions are usually too large to occupy tetrahedral holes in a fully-dense *ccp* cation array and, conversely, cations are too large to occupy eight-coordinate sites in a fully-dense primitive cubic anion array. A compound that approaches maximum density is Li_2Te containing the smallest alkali metal and largest chalcogen and for which the Te–Te distance, 4.6 Å, is only slightly greater than the diameter of the Te^{2+} ion, 4.4 Å.

In this section, we have described the ideal cubic fluorite structure, A_2X. A number of closely related structures with a range of formulae, including the pyrochlore structure, is described in Section 1.17.12.

1.17.1.4 Bond length calculations

It is very often desirable to be able to calculate interatomic distances in crystal structures. This is usually straightforward for crystals which have *orthogonal* unit cells (i.e. $\alpha = \beta = \gamma = 90^\circ$), and involves simple trigonometric calculations. For example, in the rock salt structure, the anion–cation distance is $a/2$ and the anion–anion distance is $a/\sqrt{2}$. The main interatomic distances involving cations in T/O sites in an *fcc/ccp* anion array are shown in Fig. 1.25(c) and are summarised in Table 1.11 for the important structure types. These may be used together with the tables of unit cell dimensions, Table 1.9, etc., for calculations on specific compounds. Typical values of bond distances (to oxygen and fluorine) for all the elements are given in Appendix F, where relevant values for different coordination numbers and oxidation states are also given.

Consideration of atom arrangements in the three structure types described above shows that the concept of *cp* anions with cations in interstitial sites begins to break down in the fluorite structure. Thus, while the

Table 1.11 ***Calculation of interatomic distances in some simple structures***

Structure type	Distance	Number of such distances	Magnitude of distance in terms of unit cell dimensions
Rock salt (cubic)	Na–Cl	6	$a/2 = 0.5a$
	Cl–Cl	12	$a/\sqrt{2} = 0.707a$
	Na–Na	12	$a/\sqrt{2} = 0.707a$
Zinc blende (cubic)	Zn–S	4	$a\sqrt{3}/4 = 0.433a$
	Zn–Zn	12	$a/\sqrt{2} = 0.707a$
	S–S	12	$a/\sqrt{2} = 0.707a$
Fluorite (cubic)	Ca–F	4 or 8	$a\sqrt{3}/4 = 0.433a$
	Ca–Ca	12	$a/\sqrt{2} = 0.707a$
	F–F	6	$a/2 = 0.5a$
Wurtzite[a] (hexagonal)	Zn–S	4	$a\sqrt{3}/8 = 0.612a = 3c/8 = 0.375c$
	Zn–Zn	12	$a = 0.612c$
	S–S	12	$a = 0.612c$
Nickel arsenide[a] (hexagonal)	Ni–As	6	$a/\sqrt{2} = 0.707a = 0.433c$
	As–As	12	$a = 0.612c$
	Ni–Ni	2	$c/2 = 0.5c = 0.816a$
	Ni–Ni	6	$a = 0.612c$
Caesium chloride (cubic)	Cs–Cl	8	$a\sqrt{3}/2 = 0.866a$
	Cs–Cs	6	a
	Cl–Cl	6	a
Cadmium iodide (hexagonal)	Cd–I	6	$a/\sqrt{2} = 0.707a = 0.433c$
	I–I	12	$a = 0.612c$
	Cd–Cd	6	$a = 0.612c$

[a]These formulae do not necessarily apply when *c/a* is different from the ideal value of 1.633.

antifluorite structure of Na_2O may be regarded as containing *ccp* O^{2-} ions with Na^+ ions in tetrahedral sites, in the fluorite structure of CaF_2, it is necessary to regard the Ca^{2+} ions as forming the *ccp* array with F^- ions in tetrahedral sites. In CaF_2 the Ca^{2+} ions have a eutactic *ccp* arrangement, but are well separated from each other; from Table 1.9 and Table 1.10, Ca–Ca = 3.86 Å, which is much larger than the diameter of a Ca^{2+} ion (depending on which table of ionic radii is consulted, the diameter of Ca^{2+} is in the range ~2.2–2.6 Å).

The F–F distance in CaF_2 is 2.73 Å, which indicates that the fluorines are approximately contacting (r_{F^-} = 1.2–1.4 Å). Although the array of F^- ions in CaF_2 is not *cp* but is primitive cubic, this is perhaps a more realistic way of describing the structure than as containing a eutactic *ccp* array of Ca^{2+} ions; nevertheless, both descriptions are used widely.

1.17.2 Diamond

This structure type, so important to the semiconductor industry, has already been described, as the zinc blende or sphalerite structure, Fig. 1.29 and Fig. 1.33. The diamond structure is obtained when the two elements in zinc blende are identical, as in diamond. It may therefore be described as a *ccp* array of carbon atoms, with one set of tetrahedral sites (either T_+ or T_-) occupied also by carbon atoms. It is, however, rather artificial to

make a distinction between packing and interstitial atoms since structurally they are identical. Most Group IV elements crystallise with the diamond structure, Table 1.9.

1.17.3 Wurtzite (ZnS) and nickel arsenide (NiAs)

These structures have in common an *hcp* arrangement of anions and differ only in the positions of the cations, as follows:

- wurtzite: T_+ (or T_-) sites occupied; T_- (or T_+), O empty
- nickel arsenide: O sites occupied; T_+,T_- empty.

These structures are the *hcp* analogues of the *ccp* sphalerite and rock salt structures, respectively. Note that there is no hexagonal equivalent of the fluorite and antifluorite structures.

Both wurtzite and nickel arsenide have hexagonal symmetry and unit cells. A unit cell containing *hcp* anions is shown in Fig. 1.35(a). It is less easy to visualise and draw on paper than a cubic cell because of the γ angle of 120°. The unit cell contains two anions, one at the origin and one inside the cell:

$$0, 0, 0;\ {}^{1}/_{3}, {}^{2}/_{3}, {}^{1}/_{2}$$

In Fig. 1.35(b) is shown a projection down c of the same structure. Close packed layers occur in the basal plane, i.e. at $c = 0$ (blue circles), at $c = 1$ (not shown) and at $c = {}^{1}/_{2}$ (pink circles). The layer stacking arrangement is repeated every other layer ... ABABA.... The contents of one unit cell are shown in (c). Dashed circles represent atoms at the top four corners of the unit cell, i.e. at $c = 1$.

In metals which have *hcp* structures, adjacent metal atoms are in contact, e.g. along a_1 and a_2 ($\equiv b$) edges of the unit cell (b,c). In eutactic *cp* ionic structures, however, the anions may be pushed apart by the cations in the interstitial sites. Assuming for the moment that the anions are in contact, then the hexagonal unit cell has a definite shape given by the ratio $c/a = 1.633$. This is because a is equal to the shortest distance X–X, i.e. the diameter of an anion, and c is equal to twice the vertical height of a tetrahedron comprising four anions. The ratio c/a may then be calculated by geometry (see Appendix C).

The interstitial sites available for cations in an *hcp* anion array are shown in Fig. 1.35(d) for the bottom half of the unit cell between $c = 0$ and $c = {}^{1}/_{2}$ and in (e) for the top half of the unit cell. Since the cell contains two anions, there must be two each of T_+, T_- and O.

A detailed description of the sites that are occupied in wurtzite and NiAs is now given for completeness; such a degree of detail may, however, not be necessary to gain an overview of these two structures and is not compulsory reading! In Fig. 1.35(d), a T_- site occurs along the c edge of the cell at height ${}^{3}/_{8}$ above the anion at the origin. This T_- site is coordinated to three anions at $c = {}^{1}/_{2}$ and one anion at the corner, $c = 0$. The tetrahedron so formed therefore points downwards. The position of the T_- site inside this tetrahedron is at the centre of gravity, i.e. at one-quarter of the vertical distance from base to apex (see Appendix C). Since the apex and base are at $c = 0$ and $c = {}^{1}/_{2}$, this T_- site is at $c = {}^{3}/_{8}$. In practice, the occupant of this T_- site in the wurtzite structure may not be at exactly $0.375c$. For those structures which have been studied accurately, Table 1.11, values range from 0.345 to 0.385; the letter w represents the fractional c value.

The three anions at $c = {}^{1}/_{2}$ that form the base of this T_- site also form the base of a T_+ site shown in Fig. 1.35(e) centred at 0, 0, ${}^{5}/_{8}$. The apex of this tetrahedron is the anion at the top corner with coordinates 0, 0, 1. Another T_+ site at ${}^{1}/_{3}$, ${}^{2}/_{3}$, ${}^{1}/_{8}$ is coordinated to three anions in the basal plane and an anion at ${}^{1}/_{3}$, ${}^{2}/_{3}$, ${}^{1}/_{2}$ (d). The triangular base of this site, at $c = 0$, is shared with a T_- site underneath (not shown) at ${}^{1}/_{3}$, ${}^{2}/_{3}$, $-{}^{1}/_{8}$. The equivalent T_- site that lies inside the unit cell is at ${}^{1}/_{3}$, ${}^{2}/_{3}$, ${}^{7}/_{8}$ (e).

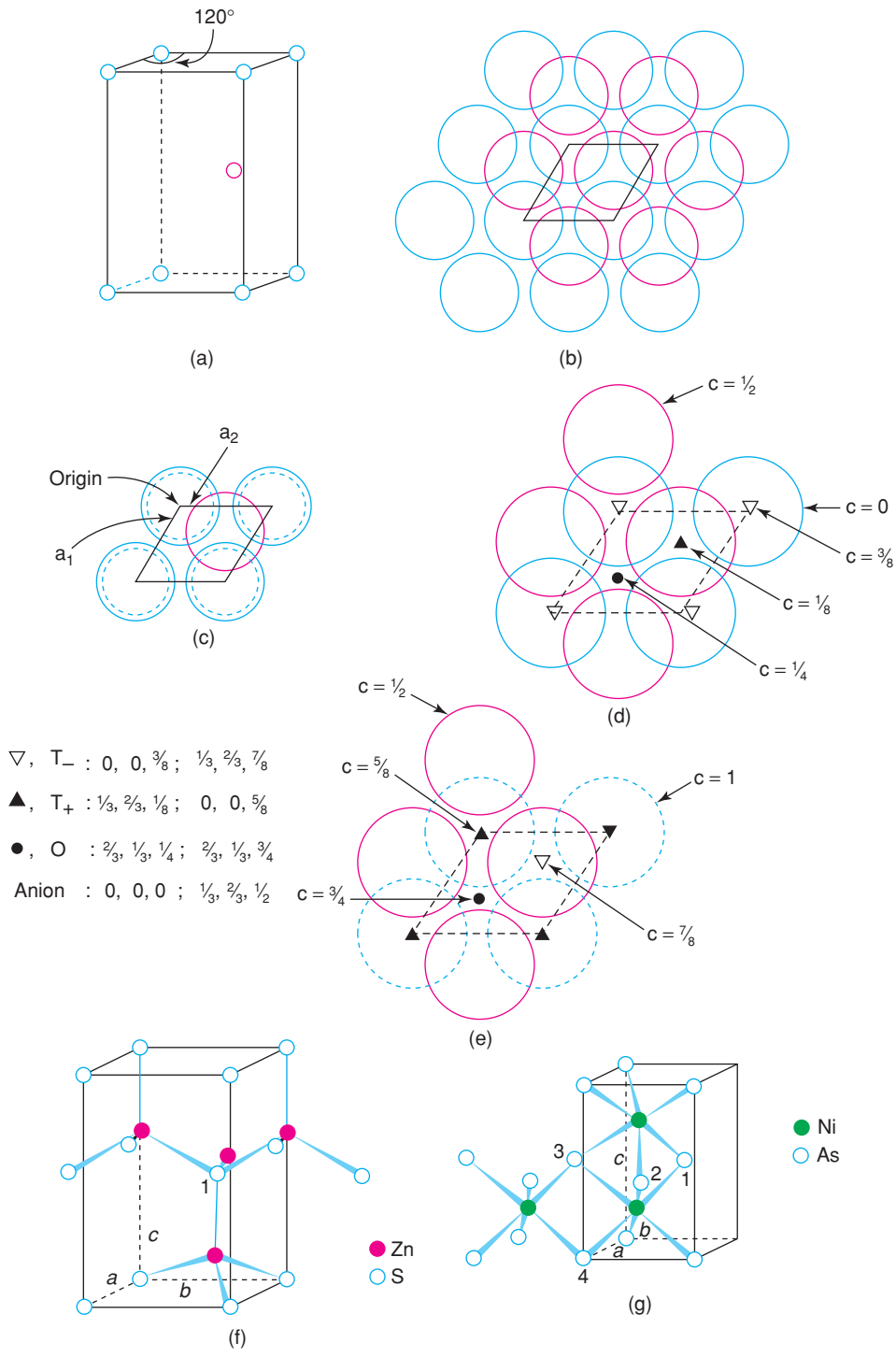


Figure 1.35 ***The wurtzite and nickel arsenide structures: (a), (b), (c) the hexagonal unit cell of an*** **hcp** ***anion array; (d), (e) interstitial sites in an*** **hcp** ***array; (f), (g) structures of wurtzite and NiAs;***

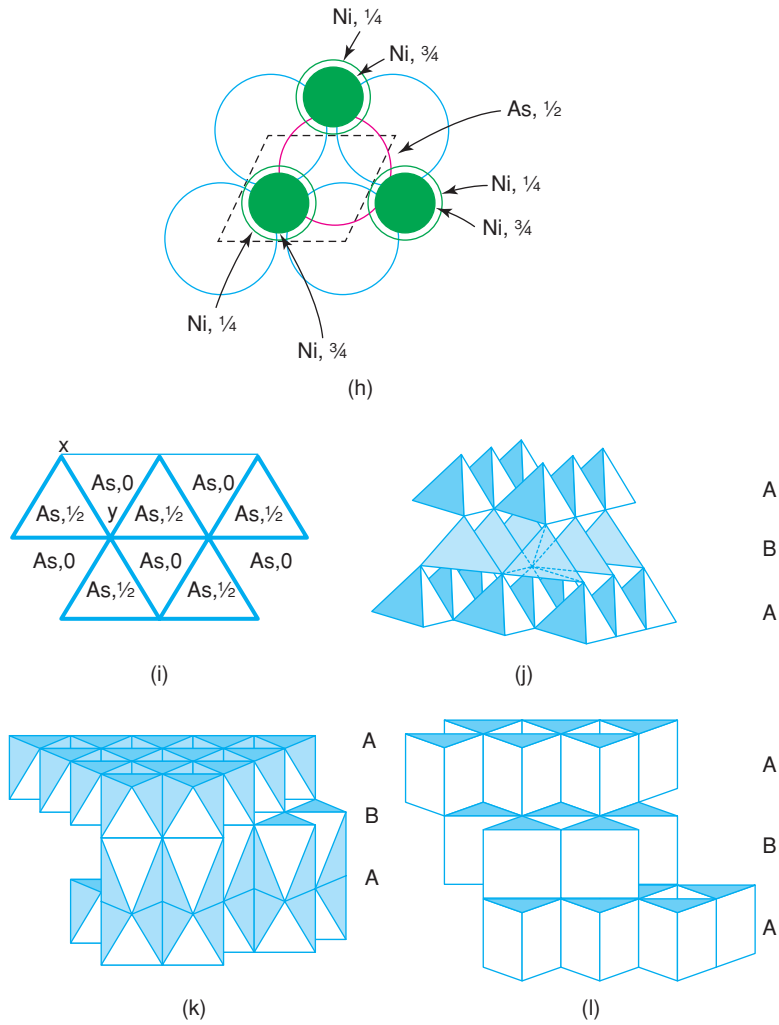


Figure 1.35 **(Continued)** ***(h), (i) trigonal prismatic coordination of arsenic in NiAs; (j), (k), (l) models of the ZnS and NiAs structures showing the arrangement and linkages of the polyhedra.***

The octahedral site in Fig. 1.35(d) is coordinated to three anions at $c = 0$ and three anions at $c = {}^1/_2$. The centre of gravity of the octahedron lies midway between these two groups of anions and has coordinates ${}^2/_3$, ${}^1/_3$, ${}^1/_4$. The second octahedral site lies immediately above the octahedral site shown in (d) and has coordinates ${}^2/_3$, ${}^1/_3$, ${}^3/_4$ (e). The three anions at $c = {}^1/_2$ are therefore common to the two octahedra, which means that octahedral sites share opposite faces.

The coordination environments of the cations in wurtzite and NiAs are emphasised in Fig. 1.35(f) and (g). Zinc is shown in T_+ sites and forms ZnS_4 tetrahedra (f), linked at their corners to form a 3D network, as in (j). A similar structure results on considering the tetrahedra formed by four Zn atoms around a S. The tetrahedral environment of S (1) is shown in (f). The SZn_4 tetrahedron which it forms points down, in contrast to the ZnS_4 tetrahedra, all of which point up; on turning the SZn_4 tetrahedra upside down, however, the same structure results.

Table 1.12 ***Some compounds with the wurtzite structure***

Compound	*a*/Å	*c*/Å	*u*	*c*/*a*	Compound	*a*/Å	*c*/Å	*u*	*c*/*a*
ZnO	3.2495	5.2069	0.345	1.602	AgI	4.580	7.494		1.636
ZnS	3.811	6.234		1.636	AlN	3.111	4.978	0.385	1.600
ZnSe	3.98	6.53		1.641	GaN	3.180	5.166		1.625
ZnTe	4.27	6.99		1.637	InN	3.533	5.693		1.611
BeO	2.698	4.380	0.378	1.623	TaN	3.05	4.94		1.620
CdS	4.1348	6.7490		1.632	NH_4F	4.39	7.02	0.365	1.600
CdSe	4.30	7.02		1.633	SiC	3.076	5.048		1.641
MnS	3.976	6.432		1.618	MnSe	4.12	6.72		1.631

Source: Reproduced from Wyckoff, Crystal Structures, Vol. 1 © 1971 Wiley Interscience. Reproduced under the terms of the STM Agreement.

Comparing larger scale models of zinc blende (Fig. 1.33b) and wurtzite (Fig. 1.35j), they are clearly very similar and both can be regarded as networks of tetrahedra. In zinc blende, layers of tetrahedra form an ABC stacking sequence and the orientation of the tetrahedra within each layer is identical. In wurtzite, the layers form an AB sequence and alternate layers are rotated by 180° about *c* relative to each other.

The $NiAs_6$ octahedra in NiAs are shown in Fig. 1.35(g). They share one pair of opposite faces (e.g. the face formed by arsenic ions 1, 2 and 3) to form chains of face-sharing octahedra that run parallel to *c*. In the *ab* plane, however, the octahedra share only edges: As atoms 3 and 4 are shared between two octahedra such that chains of edge-sharing octahedra form parallel to *b*. Similarly, chains of edge-sharing octahedra form parallel to *a* (not shown). A more extended view of the octahedra and their linkages is shown in (k).

The NiAs structure is unusual in that the anions and cations have the same coordination number but different coordination environment. Since the cation:anion ratio is 1:1 and the Ni coordination is octahedral, As must also be six-coordinate. However, the six Ni neighbours are arranged as in a trigonal prism and not octahedrally. This is shown for As at $c = {}^1/_2$ in Fig. 1.35(h), which is coordinated to three Ni at $c = {}^1/_4$ and three at $c = {}^3/_4$. The two sets of Ni are superposed in projection down *c* and give trigonal prismatic coordination for As. [Note that in a similar projection for octahedral coordination, the two sets of three coordinating atoms are staggered relative to each other, as in (e).]

The NiAs structure may also be regarded as built of $AsNi_6$ trigonal prisms, therefore, which link up by sharing edges to form a 3D array. In Fig. 1.35(i), each triangle represents a prism in projection down *c*. The prism edges that run parallel to *c*, i.e. those formed by Ni at $c = {}^1/_4$ and ${}^3/_4$ in (h), are shared between three prisms. Prism edges that lie in the *ab* plane are shared between only two prisms, however. In (i), the edge *xy* is shared between As at $c = {}^1/_2$ and $c = 0$. The structure therefore has layers of prisms arranged in an . . . ABABA . . . hexagonal stacking sequence, as shown further in (1).

A selection of compounds with wurtzite and NiAs structures is given in Table 1.12 and Table 1.13 with values of their hexagonal cell parameters *a* and *c*. The wurtzite structure is formed mainly by chalcogenides of divalent metals and is a fairly ionic structure. The NiAs structure is more metallic and is adopted by a variety of intermetallic compounds and some transition metal chalcogenides (S, Se, Te). The value of the ratio *c*/*a* is approximately constant in the wurtzite structures but varies considerably in compounds with the NiAs structure. This is associated with the presence of metallic bonding which arises from metal–metal interactions in the *c* direction, as follows. First consider the environment of Ni and As:

- Each As is surrounded by (Table 1.11):
  - 6 Ni in a trigonal prism at distance 0.707*a*
  - 12 As, *hcp* arrangement, at distance *a*.

Table 1.13 ***Some compounds with the NiAs structure***

Compound	*a*/Å	*c*/Å	*c*/*a*	Compound	*a*/Å	*c*/Å	*c*/*a*
NiS	3.4392	5.3484	1.555	CoS	3.367	5.160	1.533
NiAs	3.602	5.009	1.391	CoSe	3.6294	5.3006	1.460
NiSb	3.94	5.14	1.305	CoTe	3.886	5.360	1.379
NiSe	3.6613	5.3562	1.463	CoSb	3.866	5.188	1.342
NiSn	4.048	5.123	1.266	CrSe	3.684	6.019	1.634
NiTe	3.957	5.354	1.353	CrTe	3.981	6.211	1.560
FeS	3.438	5.880	1.710	CrSb	4.108	5.440	1.324
FeSe	3.637	5.958	1.638	MnTe	4.1429	6.7031	1.618
FeTe	3.800	5.651	1.487	MnAs	3.710	5.691	1.534
FeSb	4.06	5.13	1.264	MnSb	4.120	5.784	1.404
δ'-NbN[a]	2.968	5.549	1.870	MnBi	4.30	6.12	1.423
PtB*	3.358	4.058	1.208	PtSb	4.130	5.472	1.325
PtSn	4.103	5.428	1.323	PtBi	4.315	5.490	1.272

[a] Anti-NiAs structure.

- Each Ni is surrounded by:
  - 6 As, octahedrally, at distance 0.707*a*
  - 2 Ni, linearly, parallel to *c*, at distance 0.816*a* (i.e. *c*/2)
  - 6 Ni, hexagonally, in *ab* plane at distance *a*.

The main effect of changing the value of the c/a ratio is to alter the Ni–Ni distance parallel to *c*. Thus, in FeTe, $c/a = 1.49$, and the Fe–Fe distance is reduced to 0.745*a* [i.e. $c/2 = {}^1/_2(1.49a)$], thereby bringing these Fe atoms into close contact and increasing the metallic bonding in the *c* direction. Simple quantitative calculations of the effect of changing the *c*/*a* ratio are difficult to make since it is not readily possible to distinguish between, for example, an increase in *a* and a decrease in *c*, either of which could cause the same effect on the *c*/*a* ratio.

1.17.4 Caesium chloride (CsCl)

The unit cell of CsCl is shown in Fig. 1.36. It is a primitive cube, containing Cl at corners and Cs at the body centre, or vice versa (note that it is *not* body centred cubic since there are different ions at corner and body centre positions). The coordination numbers of both Cs and Cl are eight with interatomic distances of 0.866*a*, Table 1.11. The CsCl structure is not *cp*. In a *cp* structure, each anion has 12 other anions as next nearest neighbours whereas in CsCl, each Cl has only six Cl^- ions as next nearest neighbours (arranged octahedrally). Some compounds with the CsCl structure are given in Table 1.14. They fall into two groups, halides of large monovalent elements and a variety of intermetallic compounds.

Although CsCl is not a *cp* structure, there is a link between it and the fluorite structure, which can be described as a primitive cubic array of anions with cations in *alternate* cube body centres, Fig. 1.34; in CsCl, *all* body centres are occupied.

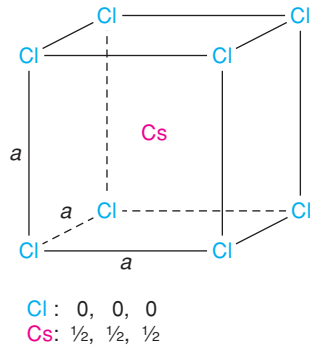


Figure 1.36 ***The primitive cubic unit cell of CsCl.***

Table 1.14 ***Some compounds with the CsCl structure***

Compound	*a*/Å	Compound	*a*/Å	Compound	*a*/Å	Compound	*a*/Å
CsCl	4.123	NH_4Br	4.0594	CuPd	2.988	AlNi	2.881
CsBr	4.286	TlCl	3.8340	AuMg	3.259	LiHg	3.287
CsI	4.5667	TlBr	3.97	AuZn	3.19	MgSr	3.900
CsCN	4.25	TlI	4.198	AgZn	3.156		
NH_4Cl	3.8756	CuZn	2.945	LiAg	3.168		

1.17.5 Other AX structures

There are five main AX structure types, rock salt, CsCl, NiAs, sphalerite and wurtzite, each of which is found in a large number of compounds. There are also several less common AX structures. Some are distorted variants of one of the main structure types, e.g.:

(a) FeO at low temperatures, <90 K, has a rock salt structure with a slight rhombohedral distortion (the α angle is increased from 90 to 90.07° by a slight compression along one threefold axis). This rhombohedral distortion is associated with magnetic ordering in FeO at low temperatures (see Chapter 9).
(b) TlF has a rock salt-related structure in which the *fcc* cell is distorted into a face centred orthorhombic cell by changing the lengths of all three cell axes by different amounts.
(c) NH_4CN has a distorted CsCl structure (as in NH_4Cl) in which the CN^- ions do not assume spherical symmetry but are oriented parallel to face diagonals. This distorts the symmetry to tetragonal and effectively increases *a* relative to *c*.

Other AX compounds have completely different structures, e.g.:

(a) Compounds of the d^8 ions, Pd and Pt (in PdO, PtS, etc.), often have a square planar coordination for the cation; d^9 ions also show this effect, e.g. Cu in CuO.
(b) Compounds of heavy *p*-block atoms in their lower oxidation states (e.g. Tl^+, Pb^{2+}, Bi^{3+}) often have distorted polyhedra in which the cation exhibits the *inert pair effect*. Thus, in PbO and SnO the M^{2+} ion has four O^{2-} neighbours to one side giving a square pyramidal arrangement, Fig. 3.14. InBi is similar with Bi^{3+} showing the inert pair effect and irregular coordination.

1.17.6 Rutile (TiO_2), cadmium iodide (CdI_2), cadmium chloride ($CdCl_2$) and caesium oxide (Cs_2O)

The title structures, together with fluorite, represent the main AX_2 structure types. The unit cell of rutile is tetragonal $a = b = 4.594$ Å, $c = 2.958$ Å, and is shown in Fig. 1.37(a). The Ti positions, two per cell, are fixed at the corner 0, 0, 0 and body centre $^1/_2{}^1/_2{}^1/_2$. The O positions, four per cell, have general coordinates x, x, 0; $1 - x$, $1 - x$, 0; $^1/_2 + x$, $^1/_2 - x$, $^1/_2$; $^1/_2 - x$, $^1/_2 + x$, $^1/_2$, with a variable parameter, x, whose value must be determined experimentally. Crystal structure determination and refinement gave the x values listed beneath Fig. 1.37(a) for the four oxygens in the unit cell, i.e. with $x \simeq 0.3$.

The body centre Ti at ($^1/_2$,$^1/_2$,$^1/_2$) is coordinated octahedrally to six oxygens. Four of these, two at $z = 0$ and two at $z = 1$ directly above the two at $z = 0$, are coplanar with Ti. Two oxygens at $z = {}^1/_2$ are collinear with Ti and form the axes of the octahedron. The corner Ti are also octahedrally coordinated but the orientation of their octahedra is different, Fig. 1.37(b). The oxygens are coordinated trigonally to three Ti, e.g. oxygen at 0 in (a) is coordinated to Ti at the corner, at the body centre and at the body centre of the cell below.

The TiO_6 octahedra link by sharing edges and corners to form a 3D framework. Consider the TiO_6 octahedron in the centre of the cell in Fig. 1.37(b); a similar octahedron in identical orientation occurs in the cells above and below such that octahedra in adjacent cells share edges to form infinite chains parallel to c. For example, Ti at $z = +^1/_2$ and $-^1/_2$ in adjacent cells are both coordinated to two oxygens at $z = 0$. Chains of octahedra are similarly formed by the octahedra centred at the corners of the unit cell. The two types of chains, which differ in orientation about c by 90° and which are $c/2$ out of step with each other, link by their corners to form a 3D framework, Fig. 1.37(c).

The rutile structure is also commonly described as a distorted *hcp* oxide array with half the octahedral sites occupied by Ti. A 3 × 3 block of unit cells is shown in Fig. 1.37(d) with only the oxygen positions marked. Corrugated *cp* layers occur, both horizontally and vertically. This contrasts with the undistorted *hcp* arrangement (e), in which the layers occur in one orientation only (horizontally).

The octahedral sites between two *cp* layers in an ideal *hcp* anion array are shown in projection in Fig. 1.38(a). While all these sites are occupied in NiAs (Fig. 1.35h), only half are occupied in rutile and in such a manner that alternate horizontal rows of octahedral sites are full and empty. The orientation of the tetragonal unit cell in rutile is shown. Parallel to the tetragonal c axis, horizontally, the TiO_6 octahedra share edges. This is shown in Fig. 1.38(b) for two octahedra with oxygens 1 and 2 forming the common edge.

An alternative, and more accurate, way to describe the packing arrangement of oxide ions in rutile is as a slightly distorted version of a different type of packing called *tetragonal packing* (*tp*), characterised by a sphere coordination number of *11* which contrasts with *hcp* and *ccp* which have a packing sphere coordination number of 12. The symmetry of *tp* is quite different to that of *hcp* since *tp* is characterised by fourfold rotational symmetry without any threefold symmetry whereas *hcp* has no fourfold symmetry but is characterised by threefold (or sixfold) symmetry. An [001] projection of the rutile structure is shown in Fig. 1.37(f); the fourfold symmetry shown is not a simple rotation axis but is a fourfold screw axis in which the structure rotates by 90° and translates by half the unit cell in the [001] direction.

The bond lengths in TiO_2 may be calculated readily; for the Ti–O bond between Ti at ($^1/_2$,$^1/_2$,$^1/_2$) and O at (0.3, 0.3, 0) the difference in both x and y coordinates of Ti and O is $(^1/_2 - 0.3)a = 0.92$ Å. From a right-angled triangle calculation, the Ti–O distance in projection down c [Fig. 1.37(a)] is $(0.92^2 + 0.92^2)^{1/2}$. However, Ti and O have a difference in c height of $(^1/_2 - 0)c = 1.48$ Å and the Ti–O bond length is therefore equal to $(0.92^2 + 0.92^2 + 1.48^2)^{1/2} = 1.97$ Å. The axial Ti–O bond length between, for example, Ti($^1/_2$, $^1/_2$, $^1/_2$) and O (0.8, 0.2, 0.5) is easier to calculate because both atoms are at the same c height. It is equal to $[2(0.3 \times 4.594)^2]^{1/2} = 1.95$ Å.

Two main groups of compounds exhibit the rutile structure, Table 1.15: oxides of tetravalent metals and fluorides of divalent metals. In both cases, the metals are too small to have eight coordination and form the fluorite structure. The rutile structure may be regarded as essentially ionic.

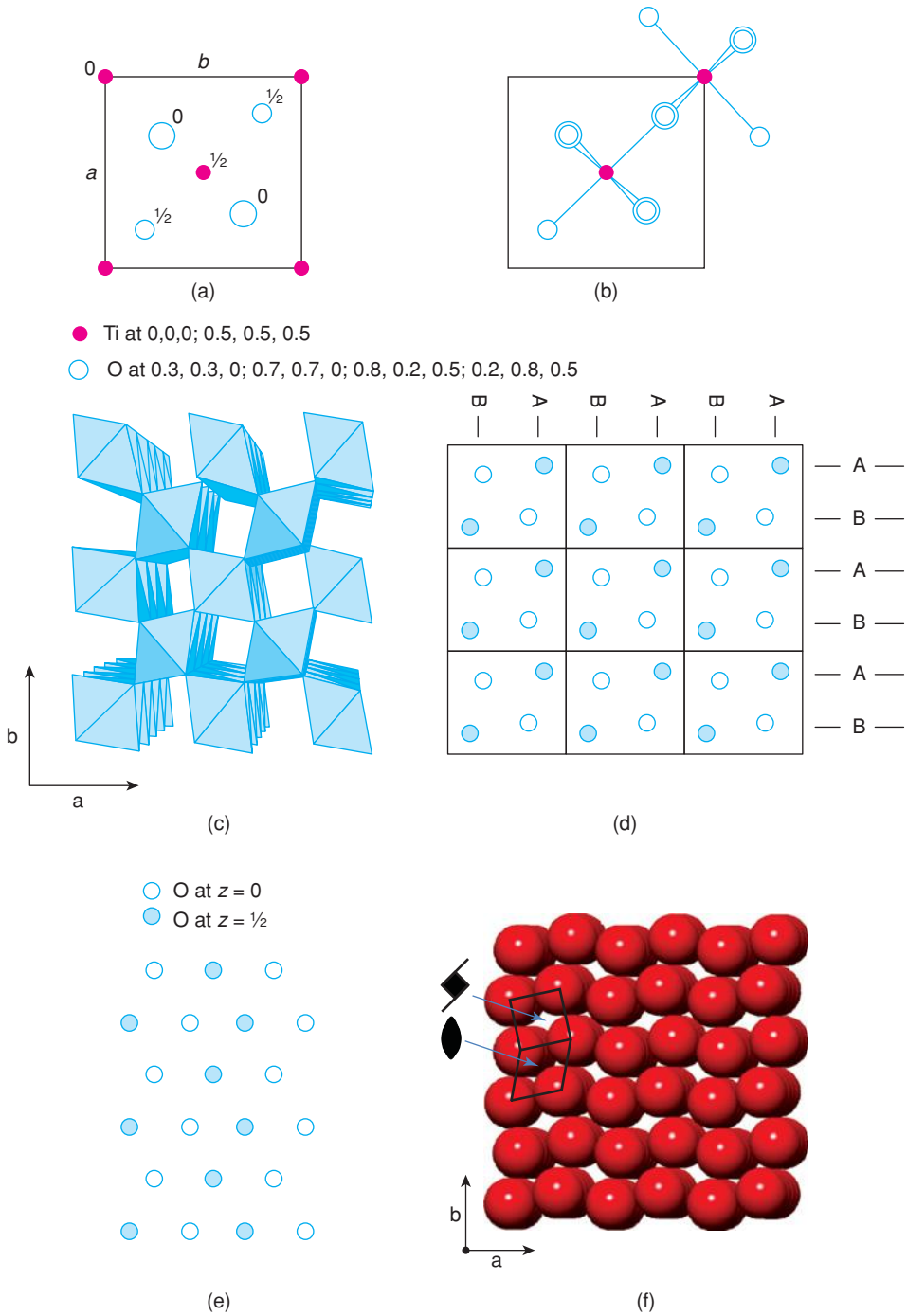


Figure 1.37 ***The rutile structure, TiO_2: (a) the unit cell; (b) TiO_6 octahedra in two orientations in the unit cell; (c) an array of octahedra in 3D; (d) oxygen atoms in zig-zag arrangement; (e) oxygen atoms in an ideal* hcp *structure in projection; (f) [001] projection of the structure showing fourfold screw axes and twofold rotation axes.***

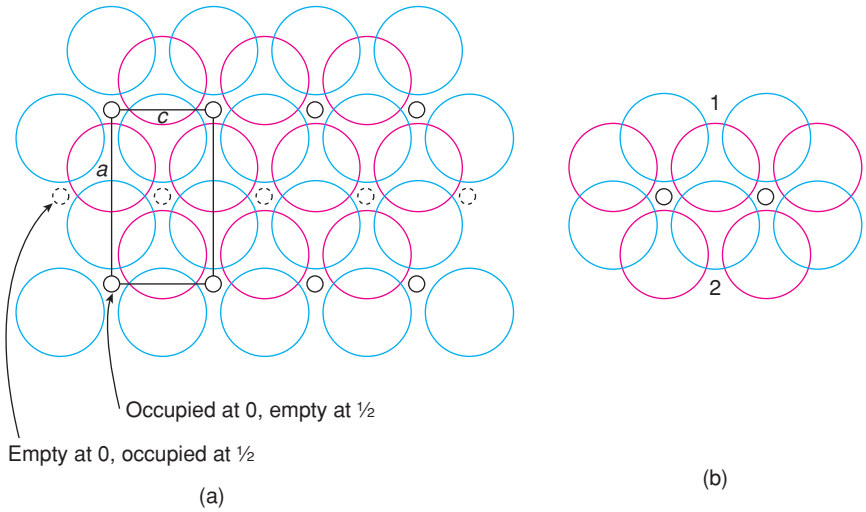


Figure 1.38 ***(a) Octahedral sites in an ideal* hcp *array; (b) edge-sharing octahedra.***

Table 1.15 ***Some compounds with the rutile structure***

Compound	*a*/Å	*c*/Å	*x*	Compound	*a*/Å	*c*/Å	*x*
TiO_2	4.5937	2.9581	0.305	CoF_2	4.6951	3.1796	0.306
CrO_2	4.41	2.91		FeF_2	4.6966	3.3091	0.300
GeO_2	4.395	2.859	0.307	MgF_2	4.623	3.052	0.303
IrO_2	4.49	3.14		MnF_2	4.8734	3.3099	0.305
β-MnO_2	4.396	2.871	0.302	NiF_2	4.6506	3.0836	0.302
MoO_2	4.86	2.79		PdF_2	4.931	3.367	
NbO_2	4.77	2.96		ZnF_2	4.7034	3.1335	0.303
OsO_2	4.51	3.19		SnO_2	4.7373	3.1864	0.307
PbO_2	4.946	3.379		TaO_2	4.709	3.065	
RuO_2	4.51	3.11		WO_2	4.86	2.77	

The CdI_2 structure is nominally similar to that of rutile because it has an *hcp* anion array with also, half of the octahedral sites occupied by M^{2+} ions. The manner of occupancy of the octahedral sites is quite different, however; entire layers of octahedral sites are occupied and these alternate with layers of empty sites, Fig. 1.39. CdI_2 is therefore a layered material in both its crystal structure and properties, in contrast to rutile, which has a more rigid, 3D character.

Two I layers in a *hcp* array are shown in Fig. 1.39(a) with the octahedral sites in between occupied by Cd. To either side of the I layers, the octahedral sites are empty. Compare this with NiAs [Fig. 1.35(d) and (h)], which has the same anion arrangement but with all octahedral sites occupied. The layer stacking sequence along *c* in CdI_2 is shown schematically in Fig. 1.39(b) and emphasises the layered nature of the CdI_2 structure: I layers form an . . . ABABA . . . stacking sequence. Cd occupies octahedral sites which may be regarded as the C positions relative to the AB positions for I. The CdI_2 structure is, effectively, a sandwich structure in which Cd^{2+} ions are sandwiched between layers of I^- ions; adjacent sandwiches are held together by weak

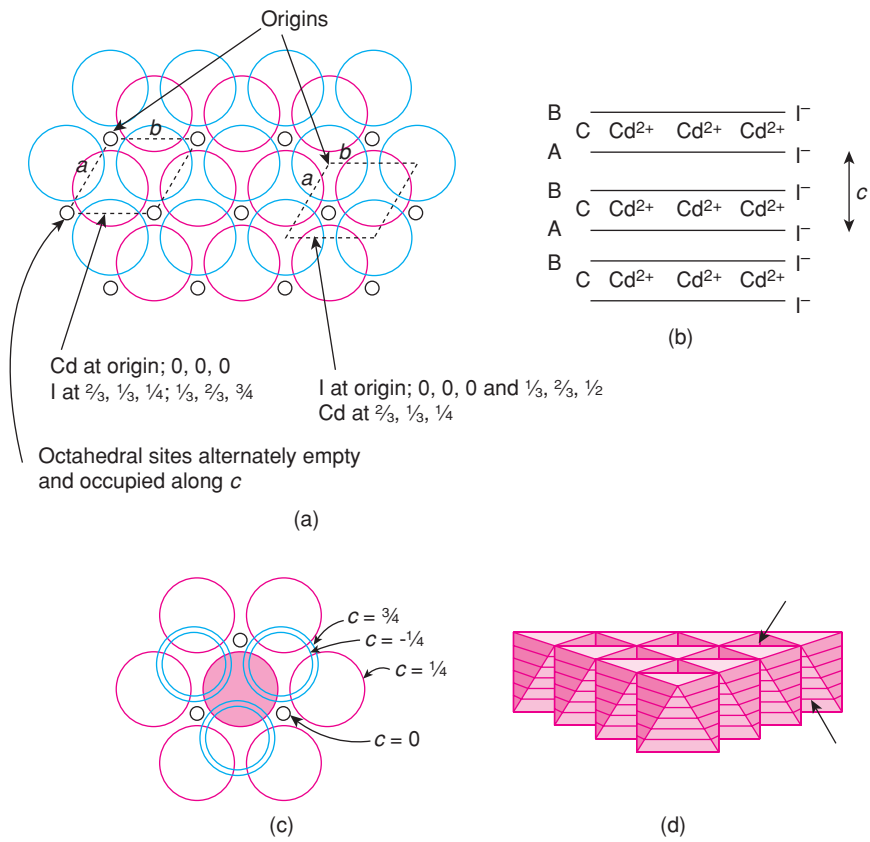


Figure 1.39 ***The CdI_2 structure: (a) the basal plane of the hexagonal unit cell is outlined, with two possible choices of origin; (b) the layer stacking sequence; (c) the coordination environment of I; (d) a layer of close packed octahedra; empty tetrahedral sites are arrowed.***

van der Waals bonds between the I layers. In this sense, CdI_2 has certain similarities to molecular structures. For example, solid CCl_4 has strong C–Cl bonds *within* the molecule but only weak Cl–Cl bonds *between* adjacent molecules. Because the intermolecular forces are weak, CCl_4 is volatile with low melting and boiling points. In the same way, CdI_2 may be regarded as an infinite sandwich 'molecule' in which there are strong Cd–I bonds within the molecule but weak van der Waals bonds between adjacent molecules.

The coordination of I in CdI_2 is shown in Fig. 1.39(c). An I at $c = {}^1/_4$ (shaded) has three close Cd neighbours to one side at $c = 0$. The next nearest neighbours are 12 I that form the *hcp* array: six are in the same plane, forming a hexagonal ring, at $c = {}^1/_4$; three are at $c = -{}^1/_4$ and three at $c = {}^3/_4$.

The layered nature of CdI_2 is emphasised further in a model of polyhedra: CdI_6 octahedra link at their edges to form infinite sheets, Fig. 1.39(d), but there are no direct polyhedral linkages between adjacent sheets. A self-supporting, 3D model of octahedra cannot be made for CdI_2, therefore. Some compounds which have the CdI_2 structure are listed in Table 1.16. It occurs mainly in transition metal iodides, bromides, chlorides and hydroxides. TiS_2 has the CdI_2 structure and was considered as a potential intercalation host cathode for use in lithium batteries (see Section 8.4): Li^+ ions are able to diffuse into the empty layers that separate adjacent TiS_2 sheets at the same time as electrons enter, and migrate through, the 3*d* band composed of d_{xy} orbitals on Ti.

Table 1.16 ***Some compounds with the CdI_2 structure***

Compound	*a*/Å	*c*/Å	Compound	*a*/Å	*c*/Å
CdI_2	4.24	6.84	VBr_2	3.768	6.180
CaI_2	4.48	6.96	$TiBr_2$	3.629	6.492
CoI_2	3.96	6.65	$MnBr_2$	3.82	6.19
FeI_2	4.04	6.75	$FeBr_2$	3.74	6.17
MgI_2	4.14	6.88	$CoBr_2$	3.68	6.12
MnI_2	4.16	6.82	$TiCl_2$	3.561	5.875
PbI_2	4.555	6.977	VCl_2	3.601	5.835
ThI_2	4.13	7.02	$Mg(OH)_2$	3.147	4.769
TiI_2	4.110	6.820	$Ca(OH)_2$	3.584	4.896
TmI_2	4.520	6.967	$Fe(OH)_2$	3.258	4.605
VI_2	4.000	6.670	$Co(OH)_2$	3.173	4.640
YbI_2	4.503	6.972	$Ni(OH)_2$	3.117	4.595
ZnI_2(l)	4.25	6.54	$Cd(OH)_2$	3.48	4.67

The $CdCl_2$ structure is closely related to that of CdI_2 and differs only in the nature of the anion packing: Cl^- ions are *ccp* in $CdCl_2$ whereas I^- is *hcp* in CdI_2.

The $CdCl_2$ structure may be represented by a hexagonal unit cell, although a smaller rhombohedral cell can also be chosen. The base of the hexagonal cell is of similar size and shape to that in CdI_2 but its *c* axis is three times longer than *c* in CdI_2. This is because in $CdCl_2$, the Cd positions, and the $CdCl_6$ octahedra, are staggered along *c* and give rise to a three-layer repeat for Cd (CBA) and a six-layer repeat for Cl (ABCABC), Fig. 1.40. In contrast, in CdI_2, the Cd positions and the CdI_6 octahedra are stacked on top of each other and the *c* repeat contains only two I layers (AB) and one Cd layer (C).

The unit cell of $CdCl_2$ in projection down *c* is shown in Fig. 1.40(b). Cl layers occur at $c = 0$ (A), 2/12 (B) and 4/12 (C), and this sequence is repeated at $c = 6/12$, 8/12 and 10/12. Between those Cl layers at 0 and 2/12, Cd occupies octahedral sites at 1/12. However, the octahedral sites between Cl layers at 2/12 and 4/12 are empty (these sites, at $c = 3/12$, are directly below Cd at 9/12).

The $CdCl_2$ structure is layered, similarly to CdI_2, and many of the comments made about structure and bonding in CdI_2 apply equally well. It also occurs with a variety of transition metal halides (Table 1.17).

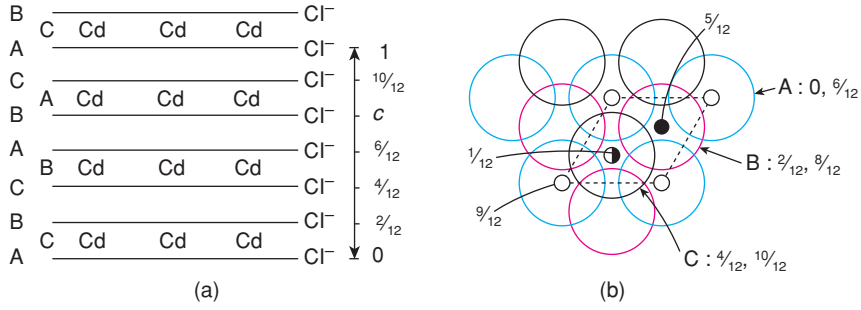


Figure 1.40 ***The $CdCl_2$ structure.***

Table 1.17 ***Some compounds with the $CdCl_2$ structure***

Compound	*a*/Å	*c*/Å	Compound	*a*/Å	*c*/Å
$CdCl_2$	3.854	17.457	$NiCl_2$	3.543	17.335
$CdBr_2$	3.95	18.67	$NiBr_2$	3.708	18.300
$CoCl_2$	3.544	17.430	NiI_2	3.892	19.634
$FeCl_2$	3.579	17.536	$ZnBr_2$	3.92	18.73
$MgCl_2$	3.596	17.589	ZnI_2	4.25	21.5
$MnCl_2$	3.686	17.470	Cs_2O[a]	4.256	18.99

[a] Cs_2O has an anti-$CdCl_2$ structure.

The structure of Cs_2O is most unusual as it is anti-$CdCl_2$. Cs forms *ccp* layers and O occupies the octahedral sites between alternate pairs of Cs layers. This raises some interesting questions because Cs is the most electropositive element and Cs salts are usually regarded as highly ionic. However, the structure of Cs_2O clearly shows that Cs is not surrounded by oxygens, as expected for an ionic structure, but has only three O neighbours, all located at one side. The structure is held together, in 3D, by bonding between Cs in adjacent layers.

It may be that the structure of Cs_2O does not reflect any peculiar type of bonding but rather that it is the only structural arrangement which is feasible for a compound of this formula and for ions of this size. Thus, from the formula, the coordination numbers of Cs and O must be in the ratio of 1:2; since Cs^+ is considerably larger than O^{2-}, the maximum possible coordination number of O by Cs may be six, which then leads to a coordination number of three for Cs.

A related question arises with the structures of the other alkali metal oxides, in particular K_2O and Rb_2O. These are antifluorites with coordination numbers of four and eight for M and O, respectively. These are unusual since Rb is normally far too large to enter into tetrahedral coordination with O. However, if there is no feasible alternative structure, then perhaps Rb has no choice but to enter the tetrahedral sites. With Cs_2O, tetrahedral coordination of Cs by O is probably impossible, hence its structure is anti-$CdCl_2$ rather than antifluorite. Thermodynamic data qualitatively support these observations; neither Cs_2O nor Rb_2O is very stable: instead, they oxidise readily to give peroxides, M_2O_2, and superoxides, MO_2, which contain much larger anions.

1.17.7 Perovskite ($SrTiO_3$)

This very important structure type, of general formula ABX_3, has a primitive cubic unit cell, shown in Fig. 1.41 as a projection down one axis (a, b) and as an oblique projection (c, d). There are two ways of drawing the unit cell of perovskite. One, (a), contains Ti at the cube corners (coordinates 0,0,0), Sr at the body centre ($^1/_2$, $^1/_2$, $^1/_2$) and O at the edge centres ($^1/_2$, 0, 0; 0, $^1/_2$, 0; 0, 0, $^1/_2$). The other, (b), contains Sr at the corners, Ti at the body centre and O at the face centres. These two descriptions are interchangeable and are simply related by translation of the origin half way along the cube body diagonal. The octahedral coordination environment of Ti is seen in (c, d) and interatomic distances calculated by simple geometry. The bond distance Ti–O = $a/2$ = 1.953 Å. Sr is equidistant from 12 oxygens and the Sr–O distance equals half the diagonal of any cell face, i.e. $a/\sqrt{2}$ or 2.76 Å.

Each O has two Ti as its nearest cationic neighbours, at 1.953 Å, and four Sr, coplanar with O at 2.76 Å. However, eight other oxygens are at the same distance, 2.76 Å, as the four Sr. It is debatable whether the O

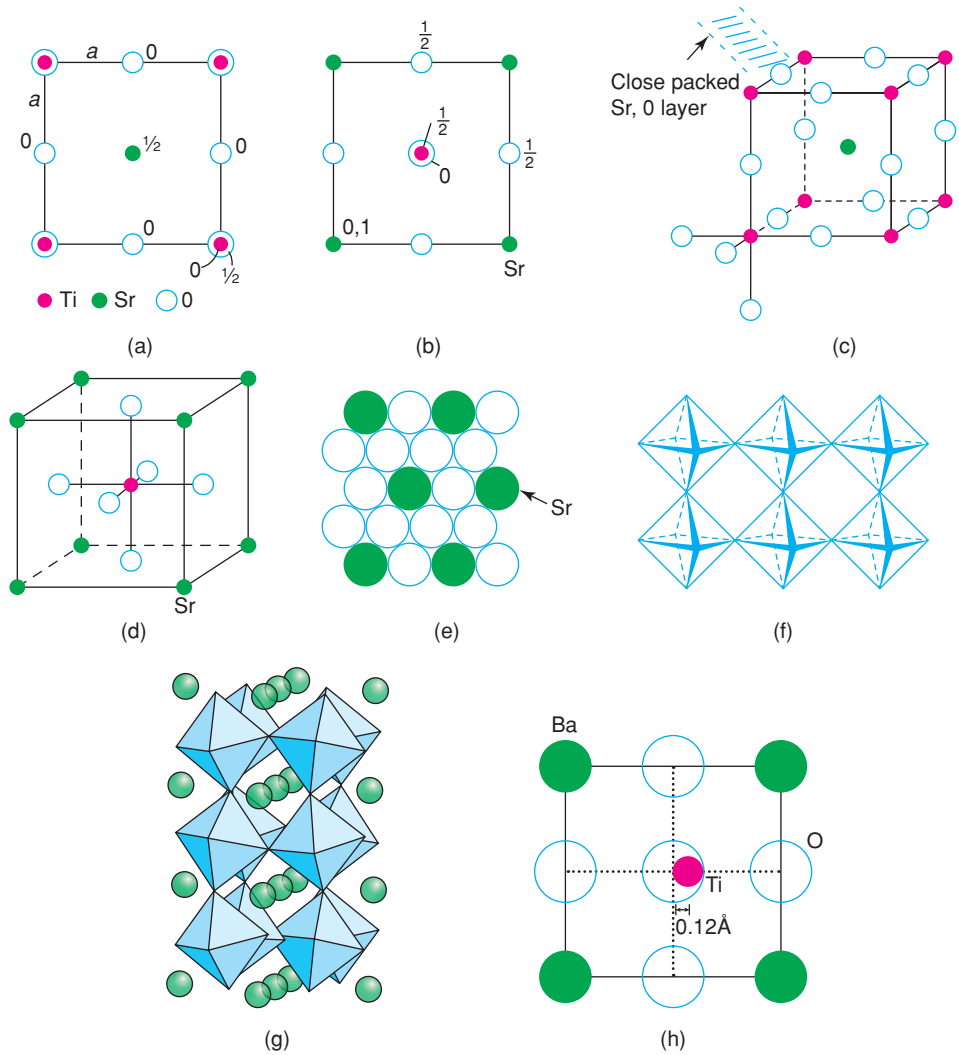


Figure 1.41 ***(a–d) The perovskite structure of $SrTiO_3$. (e) A close packed Sr,O layer. (f) A layer of corner-sharing octahedra. (g) $GdFeO_3$ structure. (h) Structure of tetragonal $BaTiO_3$ projected onto the* ac *plane. Note, the origin of the unit cell is shifted to coincide with Ba rather than with Ti as in (b). Adapted with permission from Weller,* Inorganic Materials Chemistry, *© 1994 Oxford University Press. (i) coupled rotation of octahedra in 2D corner-sharing sheets. (j, k) View looking down the* c *axis of $a^0a^0c^-$ and $a^0a^0c^+$ with the A-site cations shown as spheres and the B-site cations located at the centre of the octahedra. Adapted with permission from Lufaso and Woodward,* Acta Cryst., B57, *725–738, © 2001 International Union of Crystallography.***

coordination number is best regarded as two (linear) or as six (a grossly squashed octahedron with two short and four long distances) or as 14 (six cations and eight oxygens). No firm recommendation is made!

Having arrived at the unit cell of $SrTiO_3$, the atomic coordinates, coordination numbers and bond distances, we now wish to view the structure on a rather larger scale and ask the following questions. Does it have *cp* anions? Is it a framework structure? Answers are as follows.

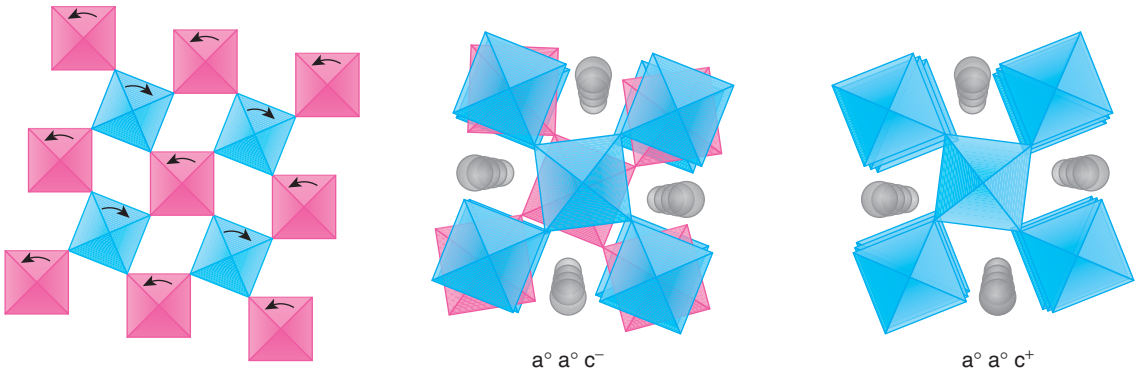


Figure 1.41 **(Continued)**

Perovskite does not contain *cp* oxide ions as such but O and Sr, considered together, do form a *ccp* array with the layers parallel to the {111} planes, Fig. 1.41(c) and (e). To see this, compare the perovskite structure with Sr at the origin, (d), with that of NaCl (Fig. 1.2). The latter contains Cl (or Na, depending on the choice of origin) at the corner and face centre positions of the cell and is *ccp*. By comparison, perovskite contains O at the face centres and Sr at the corner. The structure of the mixed Sr, O *cp* layers in perovskite is such that one-quarter of the atoms are Sr, arranged in a regular fashion, Fig. 1.41(e). It is quite common for fairly large cations, such as Sr^{2+} ($r = 1.1$ Å), to play apparently different roles in different structures, i.e. as 12-coordinate packing ions, as in $SrTiO_3$ perovskite, or as octahedrally coordinated cations within a *cp* oxide array, as in SrO (rock salt structure).

The formal relation between rock salt and perovskite also includes the Na and Ti cations as both occupy octahedral sites: in NaCl, all octahedral sites are occupied (corners and face centres), but in perovskite only one-quarter [the corner sites in (a)] are occupied.

Perovskite is also regarded as a framework structure with corner-sharing TiO_6 octahedra and with Sr in 12-coordinate interstices. The octahedral coordination of one Ti is shown in Fig. 1.41(c) and (d); each O of this octahedron is shared with one other octahedron, such that the Ti–O–Ti arrangement is linear. Thus, octahedra link at their corners to form sheets (f), and neighbouring sheets link similarly to form a 3D framework.

Several hundred oxides and halides form the perovskite structure; a selection is given in Table 1.18. The oxides contain two cations whose combined oxidation state is six. Thus, possible combinations are +I, +V

Table 1.18 ***Some compounds with the perovskite structure***

Compound	*a*/Å	Compound	*a*/Å	Compound	*a*/Å
$KNbO_3$	4.007	$LaFeO_3$	3.920		
$KTaO_3$	3.9885	$LaGaO_3$	3.875	$CsCaF_3$	4.522
KIO_3	4.410	$LaVO_3$	3.99	$CsCdBr_3$	5.33
$NaNbO_3$	3.915	$SrTiO_3$	3.9051	$CsCdCl_3$	5.20
$NaWO_3$	3.8622	$SrZrO_3$	4.101	$CsHgBr_3$	5.77
$LaCoO_3$	3.824	$SrHfO_3$	4.069	$CsHgCl_3$	5.44
$LaCrO_3$	3.874	$SrSnO_3$	4.0334		

as in $KNbO_3$, +II, +IV as in $CaTiO_3$ and +III, +III as in $LaGaO_3$. The 12-coordinate A site cations are, of course, much larger than the six-coordinate B site cations.

As well as the cubic perovskite structure, described so far, a variety of distorted, non-cubic structures exist. These lower-symmetry structures often form on cooling the high-temperature cubic structure and the framework of octahedra may be slightly twisted or distorted. An example is shown in Fig. 1.41(g) for the structure of $GdFeO_3$. The reasons for the structural distortions are associated with the size requirements of the 12-coordinate A and six-coordinate B sites and whether adjustments to the structure are required to accommodate different-sized cations. Also, more complex perovskite structures form in which two different cations may occupy either the A or B sites, giving a range of cation ordering possibilities.

1.17.7.1 Tolerance factor

The reason why structural distortions occur in many perovskites is that the A and/or B atoms are not exactly the right size to fit the sites generated by the remainder of the structure. In an oxide with the ideal, cubic perovskite structure, the bond lengths are related to the unit cell dimension, a, by

$$a = \sqrt{2} r_{A-O} = 2 r_{B-O} \tag{1.6}$$

Since bond lengths for each element, oxidation state and coordination number usually fall within closely defined ranges (Appendix F), it is possible to use equation (1.6) to see how well the sizes of a particular A, B combination meet the requirements for an undistorted, ideal perovskite. The degree to which the sizes depart from equation (1.6) is given by a *tolerance factor*, t:

$$t = \frac{\sqrt{2} r_{A-O}}{2 r_{B-O}} \tag{1.7}$$

In practice, there is some flexibility over bond lengths and usually, a cubic perovskite forms with t in the range $0.9 < t < 1.0$.

For $t > 1$, the B site is larger than required. If t is only slightly greater than 1.0, the structure distorts but is still basically a perovskite as in $BaTiO_3$, $t = 1.06$. For larger departures from $t = 1.0$, however, the B ion demands a smaller site, of lower coordination number, and the structure changes completely, as in $BaSiO_3$ which has tetrahedral Si.

For smaller tolerance factors, $0.85 < t < 0.90$, several different kinds of structural distortion occur because now, as in $GdFeO_3$, the A cation is too small for its site. These distortions generally involve tilting and rotation of the BO_6 octahedra as shown in Fig. 1.41(g). Consequently some, or all, of the B–O–B linkages are no longer linear but are zig-zag, which has the effect of reducing the size of the A cation site.

1.17.7.2 $BaTiO_3$

$BaTiO_3$ is tetragonal at room temperature, $a = 3.995$, $c = 4.034$ Å, with the structure shown in projection on the *ac* plane in Fig. 1.41(h). Since Ti is slightly too small for its octahedral site, it displaces by about 6% of the Ti–O distance towards one of the corner oxygens; Ba^{2+} ions also undergo a smaller displacement in the same direction. This reduces the coordination of Ti to five (square pyramidal) and, in order to have reasonable Ti–O bond lengths, the structure also contracts slightly in the *ab* plane [not shown in (h)].

Ti atoms in adjacent unit cells undergo a similar displacement in the same direction and the resulting structure has a large dipole moment due to the separation of positive and negative charge centres. It is possible to flip the orientation of the dipoles: under the action of an applied electric field, the Ti atoms move

through the centre of the octahedral site towards one of the other corner oxygens. This ready reversibility gives the structure high polarisability and a high permittivity (or dielectric constant) and is responsible for the property of ferroelectricity (see Section 8.7).

1.17.7.3 Tilted perovskites: Glazer notation

We saw in Fig. 1.41(g) how octahedra can tilt or rotate cooperatively leaving the BO_6 octahedra essentially unchanged but reducing the size of the A site and its coordination number. Thus, in $GdFeO_3$, the A site is eight-coordinate instead of 12-coordinate. Such structural distortions occur when the A cation is too small to occupy comfortably the 12 coordinate sites created by the array of corner-sharing BO_6 octahedra and, consequently, the octahedral rotations allow a reduction in the A–O bond length.

A wide variety of structural distortions occur in perovskites whose tolerance factor is less than unity. The most common distortion involves tilting or rotation of octahedra about one or more of the three axes of the octahedron. This is a cooperative process since octahedra link at their corners to adjacent octahedra in the 3D framework and, for instance, clockwise rotation of an octahedron about one axis causes anticlockwise rotation of the adjacent octahedra, Fig. 1.41(i). For the example shown, octahedra which form sheets in the *xy* plane undergo coupled rotations about the *z* axis.

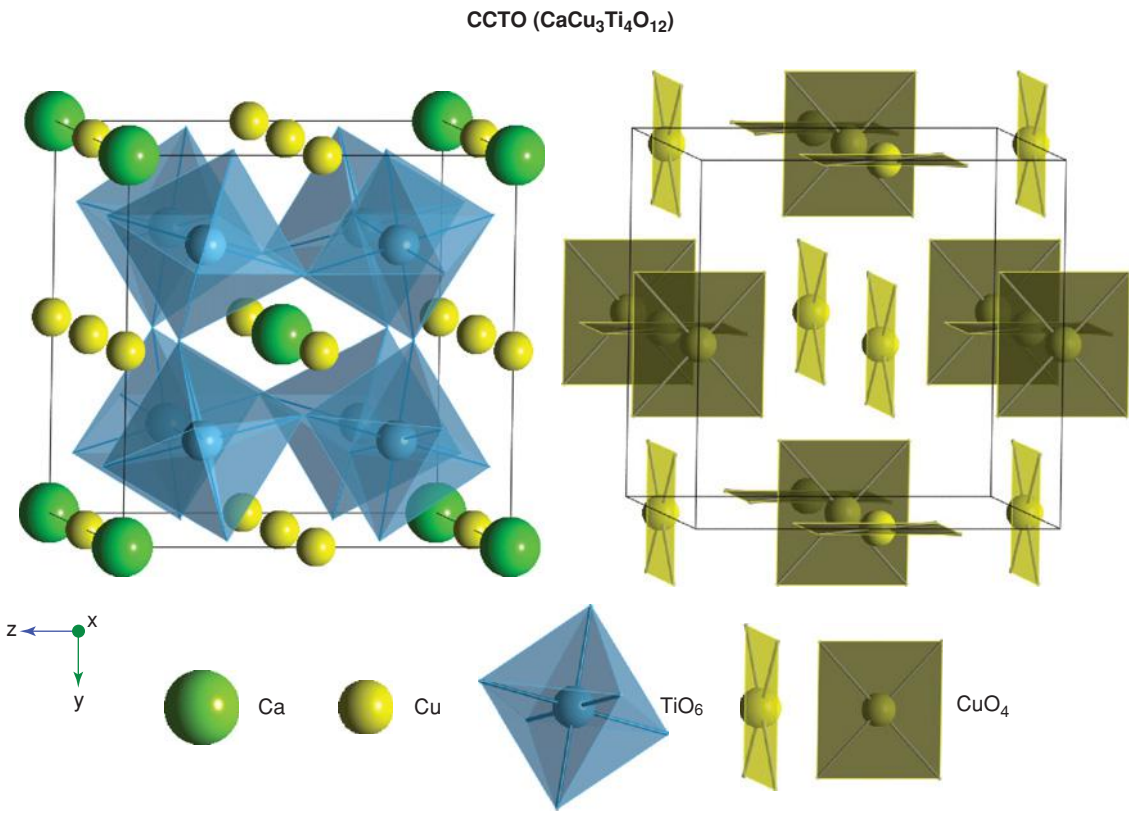
Similar coupled rotations may or may not also occur about *x* and *y* axes. Within a given sheet of octahedra, the octahedral rotations cannot occur independently of each other since clockwise rotation of one octahedron demands anticlockwise rotation of the four adjacent, corner-linked octahedra. However, between adjacent planes in, for instance, the *z* direction there is no automatic coupling between adjacent planes. This gives rise to two commonly observed possibilities in which the octahedral rotations in adjacent planes are identical or the exact opposite.

These two possibilities are illustrated in Fig. 1.41(j) and (k). Detailed analysis has shown that 15 different tilt systems are theoretically possibly and classification schemes have been developed by Glazer and Alexandrov. The widely used Glazer method is as follows:

1. Starting with the undistorted cubic perovskite, possible rotation about the three axes of the octahedron is indicated by three letters *a*, *b* and *c*. If the degree of rotation about all three axes is the same, the letters are the same, i.e. *aaa*; if rotation about one axis is different to that about the other two, the labelling is, e.g., *aac*.
2. To complete the label, each of the three letters carries a superscript; if there is no rotation about that axis, the superscript is ^{0}; if the rotation between adjacent planes of octahedra is the same, and in-phase, the superscript is $^+$; if the rotation between adjacent planes is in the opposite sense, the superscript is $^-$.

For the examples shown in Fig. 1.41(j) and (k), rotation of octahedra about only one axis is shown and therefore the two possibilities are labelled $a^0a^0c^-$ and $a^0a^0c^+$. These are easy to appreciate since the rotations are about one axis only. For cases involving rotation about all three axes, detailed scrutiny of the structures may be needed; thus for $GdFeO_3$, the tilt system is $a^-b^+a^-$ although this may not be readily apparent from diagram (g). The different tilt systems give rise to structures with unit cells that are usually no longer cubic. For instance, $GdFeO_3$ has a larger orthorhombic unit cell with $a \approx c \approx \sqrt{2}a_p$; $b = 2a_p$, where a_p refers to the parent cubic perovskite cell. Hence the unit cell of $GdFeO_3$ is four times larger than the cubic perovskite cell and contains four formula units. The $a^-b^+a^-$ tilt system is the most commonly observed distorted perovskite structure.

An interesting tilt system which retains cubic symmetry is $a^+a^+a^+$ and is shown by a family of phases with general formula $A'A''_3B_4O_{12}$ with $CaCu_3Ti_4O_{12}$ as the most studied example. The TiO_6 octahedra undergo similar coupled rotations about all three axes; a consequence of this is that three-quarters of the A sites change



Atom	Wyckoff positions	x	y	z
Ca	2a	0	0	0
Cu	6b	0	0.5	0.5
Ti	8c	0.25	0.25	0.25
O	24g	0.3029(1)	0.1790(1)	0

Cubic : a = 7.3818 Å

Space Group : Im $\bar{3}$ (No. 204)

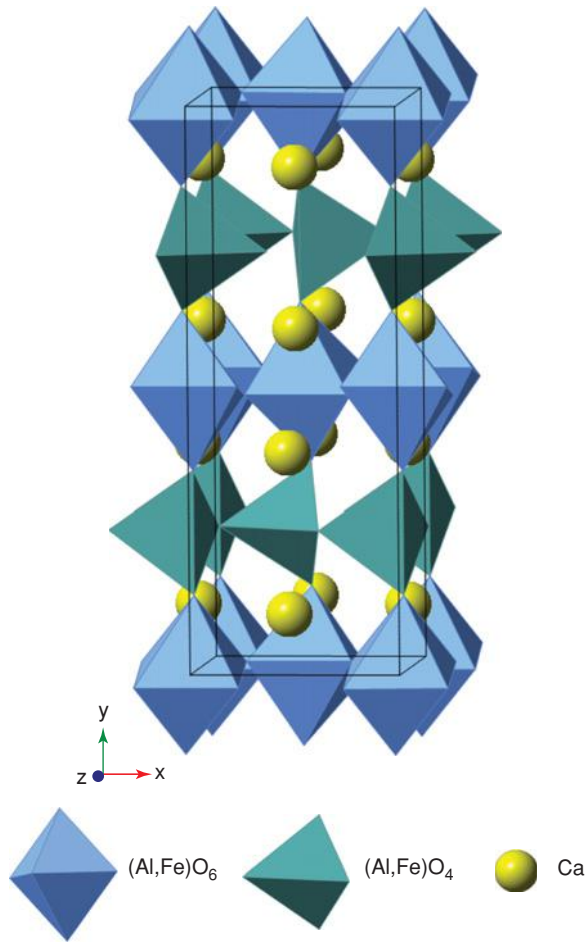
(a)

Figure 1.42 ***(a) Crystal structure of $CaCu_3Ti_4O_{12}$. Data taken from E. S. Bozin* et al.*, J. Phys. Condens. Matter, 16, S5091–S5102, © 2004 IOP Publishing.***

their coordination from 12-coordinate to square planar. This structure is shown by a number of compounds in which Cu, Mn and Ni occupy the square planar A″ sites, Fig. 1.42(a). Further comments are given in the next section.

Another common tilt system is $a^-a^-a^-$, which leads to a trigonal unit cell; it is shown by phases such as $LaBO_3$: B = Co, Ni, Cu, Al and Ga.

Brownmillerite [$Ca_2(Fe,Al)_2O_5$]



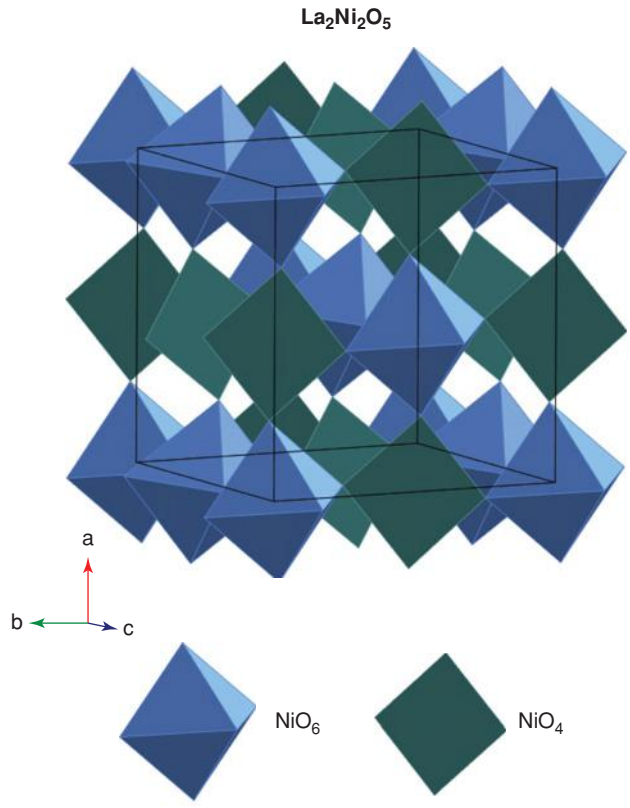
Atom	Wyckoff positions	x	y	z
Ca	8d	0.48141(6)	0.10802(3)	0.02331(6)
(Al, Fe) Site 1	4a	0	0	0
(Al, Fe) Site 2	4c	0.94628(6)	0.25	0.93351(6)
O1	8d	0.2622(2)	0.98432(8)	0.2375(2)
O2	8d	0.0236(2)	0.14090(9)	0.0723(2)
O3	4c	0.5992(3)	0.25	0.8736(3)

Orthorhombic : a = 5.4217, b = 14.7432, c = 5.5960 Å

Space Group : Pnma (No. 62)

(b)

Figure 1.42 **(Continued)** ***(b) The brownmillerite structure. Data taken from G. J. Redhammer* et al.*, Am. Miner.*, 89*, 405–420, © 2004 The Mineralogical Society of America (MSA).***



Atom	Wyckoff positions	x	y	z	f_{occ}
La	8f	0.2483(7)	0.2601(6)	0.2510(6)	1.0
Ni1	4a	0	0	0	0.5
Ni2	4b	0	0.5	0	0.5
O1	8f	0.2618(8)	0.0348(5)	0.0282(6)	1.0
O2	8f	–0.0295(6)	0.2669(9)	0.0429(5)	1.0
O3	4e	0	–0.054(1)	0.25	0.5
O4	4e	0	0.49(1)	0.25	0.06(1)

Monoclinic : a = 7.8426, b = 7.7988, c = 7.4720 Å, β = 93.859°

Space Group : C2/c (No. 15)

Tilt system:- $a^-a^-c^-$: antiphase tilt along the three axes.

(c)

Figure 1.42 **(Continued)** ***(c) The $La_2Ni_2O_5$ structure. Data taken from J. A. Alonso* et al.*, J. Phys. Condens. Matter, 9, 6417–6426, © 1997 IOP Publishing.***

1.17.7.4 $CaCu_3Ti_4O_{12}$, CCTO

There is much recent interest in a family of perovskite-related phases typified by $CaCu_3Ti_4O_{12}$ because of the possibility that they exhibit a giant dielectric constant. It is now known that the measured high permittivities are not a property of the bulk crystals but are associated with the grain–grain contacts in ceramics. The high permittivity is a consequence of the geometry of the regions responsible since the capacitance of a region is proportional to its area but inversely proportional to its thickness. CCTO has a most unusual crystal structure since the A sites of perovskite contain large Ca cations whose coordination number is reduced to eight by tilting of TiO_6 octahedra together with square planar CuO_4 units. The crystal structure is shown in Fig. 1.42(a) in which the left-hand diagram emphasises tilting of the TiO_6 octahedra and the right-hand diagram shows the square planar Cu-based units which have three equivalent orientations in the cubic unit cell. In Glazer notation, the tilt system for CCTO is $a^+a^+a^+$.

1.17.7.5 Anion-deficient perovskites

Anion-deficient perovskites occur commonly and, in cases where the anions are ordered, a variety of more complex structures are generated. As a consequence of anion deficiency, the B site coordination number must reduce from six on average. In the brownmillerite structure, $Ca(Fe,Al)O_{2.5}$, Fig. 1.42(b), the average B site coordination number is expected to be five and this is achieved in a layered structure with alternating sheets of tetrahedra and octahedra. In $LaNiO_{2.5}$, Fig. 1.42(c), the Ni coordination is a mixture of NiO_6 octahedra and NiO_4 square planes which are ordered to give a $2 \times 2 \times 2$ supercell relative to the perovskite subcell. In the *a* and *b* directions, octahedra and square planes alternate but in the *c* direction chains of corner-sharing octahedra form which are connected to adjacent chains by the square planar units. The Ni polyhedra exhibit antiphase tilting along the three unit cell axes and therefore, may be described in Glazer notation as $a^-a^-c^-$. This material was produced from $LaNiO_3$ by hydrogen reduction, but a residual amount of oxygen remained in the O(4) sites, giving a composition, for the particular material that was studied, of $LaNiO_{2.53(1)}$; possibly, using different preparative conditions, materials with different oxygen contents could be prepared.

The above are two examples of anion-ordered crystal structures which occur at specific compositions. Other anion ordering arrangements are known both at these and other compositions. In addition, solid solutions form in many systems in which the anion content is variable and can be represented by the general formula $ABO_{3-\delta}$. In these, the oxygen vacancies may be distributed at random through the perovskite structure or locally ordered structures may form in which small domains of a particular structure type are distributed at random through a disordered perovskite network.

1.17.7.6 Stoichiometry–property relations

The perovskite structure, with two different-sized cations and several possible cation charge combinations, occurs with a very wide range of compositions. In addition, defect perovskites form in which there are either cation or anion vacancies. Defects, solid solutions and various kinds of properties are all considered in later chapters. All we wish to note here is the incredible range of properties found in materials with perovskite-related structure whose composition has been adjusted to optimise a particular property. Almost every physical property imaginable has been found in materials with the perovskite structure by changing their composition and/or defect structure; for this reason, perovskite is sometimes referred to as an inorganic chameleon! A selection of perovskites and their properties is listed in Table 1.19.

Table 1.19 ***Perovskites: some composition–property correlations***

Composition	Property
$CaTiO_3$	Dielectric
$BaTiO_3$	Ferroelectric
$Pb(Mg_{1/3}Nb_{2/3})O_3$	Relaxor ferroelectric
$Pb(Zr_{1-x}Ti_x)O_3$	Piezoelectric
$(Ba_{1-x}La_x)TiO_3$	Semiconductor
$(Y_{1/3}Ba_{2/3})CuO_{3-x}$	Superconductor
Na_xWO_3	Mixed conductor (Na^+, e^-); electrochromic
$SrCeO_3$:H	Proton conductor
RE TM O_{3-x}	Mixed conductor (O^{2-}, e^-)
$Li_{0.5-3x}La_{0.5+x}TiO_3$	Li^+ ion conductor
A $MnO_{3-\delta}$	Giant magnetoresistance effect

RE = rare earth; TM = transition metal.

1.17.8 Rhenium trioxide (ReO_3), perovskite tungsten bronzes, tetragonal tungsten bronzes and tunnel structures

The structure of cubic ReO_3 is closely related to perovskite described above. It is the same as the 'TiO_3' framework of perovskite, $SrTiO_3$, but without the Sr atoms. Its unit cell is the same as that shown in Fig. 1.41(a) with Re at corners and O at edge centres. A few oxides and halides form the ReO_3 structure, Table 1.20, together with an example of the anti-ReO_3 structure in Cu_3N.

A wide variety of oxide and oxyfluoride structures, often with complex formulae, can be derived from the ReO_3 structure by coupled rotation of groups of octahedra. To see this, we start with the ideal cubic ReO_3 structure shown in Fig. 1.43(a). Each octahedron shares its six corners with other octahedra and in 2D, layers of corner-sharing octahedra are obtained. In the cubic perovskite structure, large 12-coordinate cavities are occupied by A cations, ideally without modification of the array of octahedra, (b).

The *perovskite tungsten bronze* structures are intermediate between ReO_3 and perovskite. They occur in series such as Na_xWO_3 and have a 3D framework of WO_6 octahedra, as in ReO_3, but with some ($0 < x < 1$) of the large A sites occupied by Na. To accommodate a variation in stoichiometry, x, the oxidation state of tungsten is a mixture of, or intermediate between, V and VI. The formula of the bronzes may be written more completely as

$$Na_x W^{V}_{x} W^{VI}_{1-x} O_3$$

These materials have interesting colours and electrical properties. At low x, they are pale green/yellow in colour and semiconducting. As x rises and electrons begin to occupy the 5*d* band of tungsten, they become brightly coloured and metallic, hence the name 'bronze'.

Table 1.20 ***Some compounds with the ReO_3 structure***

Compound	*a*/Å	Compound	*a*/Å
ReO_3	3.734	NbF_3	3.903
UO_3	4.156	TaF_3	3.9012
MoF_3	3.8985	Cu_3N	3.807

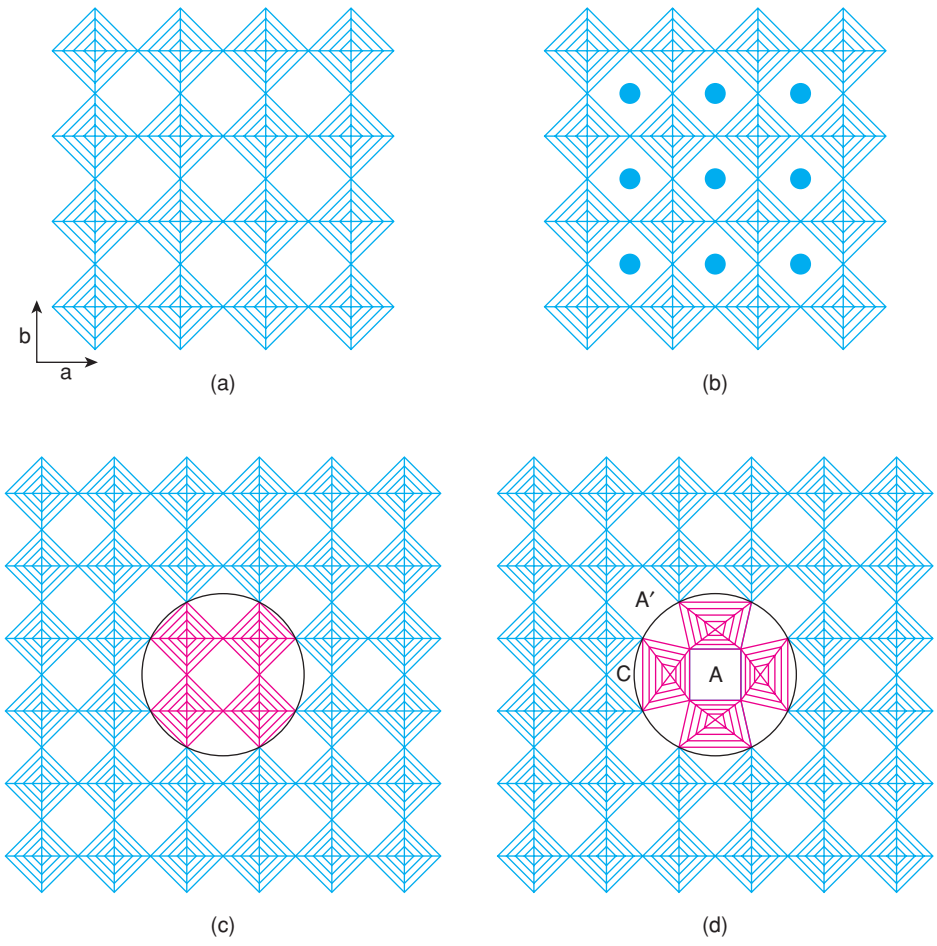


Figure 1.43 ***The structure of (a) ReO_3, (b) tungsten bronze Na_xWO_3, (c,d) bronze and tunnel structures derived from ReO_3 by rotation of blocks of four octahedra and (e,f) the tetragonal tungsten bronze structure. Adapted with permission from B. G. Hyde and M. O'Keeffe, Acta Cryst., A29, 243 © 1973 International Union of Crystallography. (g) The TTB structure of $Ba_2LaTi_2Nb_3O_{15}$. Adapted from G. C. Miles* et al. *(2005), J. Mater. Chem., 15, 798). (h, i) Tunnel structure of Mo_5O_{14}.***

A variety of monovalent cations enter the tungsten bronze structure; similar series also occur with MoO_3 in the *molybdenum bronzes*. Many other highly conducting bronzes based on oxides MO_2 or MO_3 form in which M is a transition metal, such as Ti, V, Nb, Ta, Mo, W, Re, Ru or Pt, that is capable of existing in mixed oxidation states.

Thus far, we have not modified the arrangement of octahedra. Different arrays of octahedra are, however, obtained by rotation of columns of four octahedra by 90° and reconnecting them with the parent structure as shown in Fig. 1.43(c) and (d). This has major consequences for the size and coordination number of the A cation sites. Some, A, retain their coordination number of 12 and in projection may be perceived as 'square tunnels'; a second set of sites, A′, has a coordination number which is increased to 15 and in projection may be regarded as "pentagonal tunnels'. A third set, C, has coordination number reduced to nine and is referred to as 'triangular tunnels'. We now have the building blocks for the *tetragonal tungsten bronze,* TTB, structure.

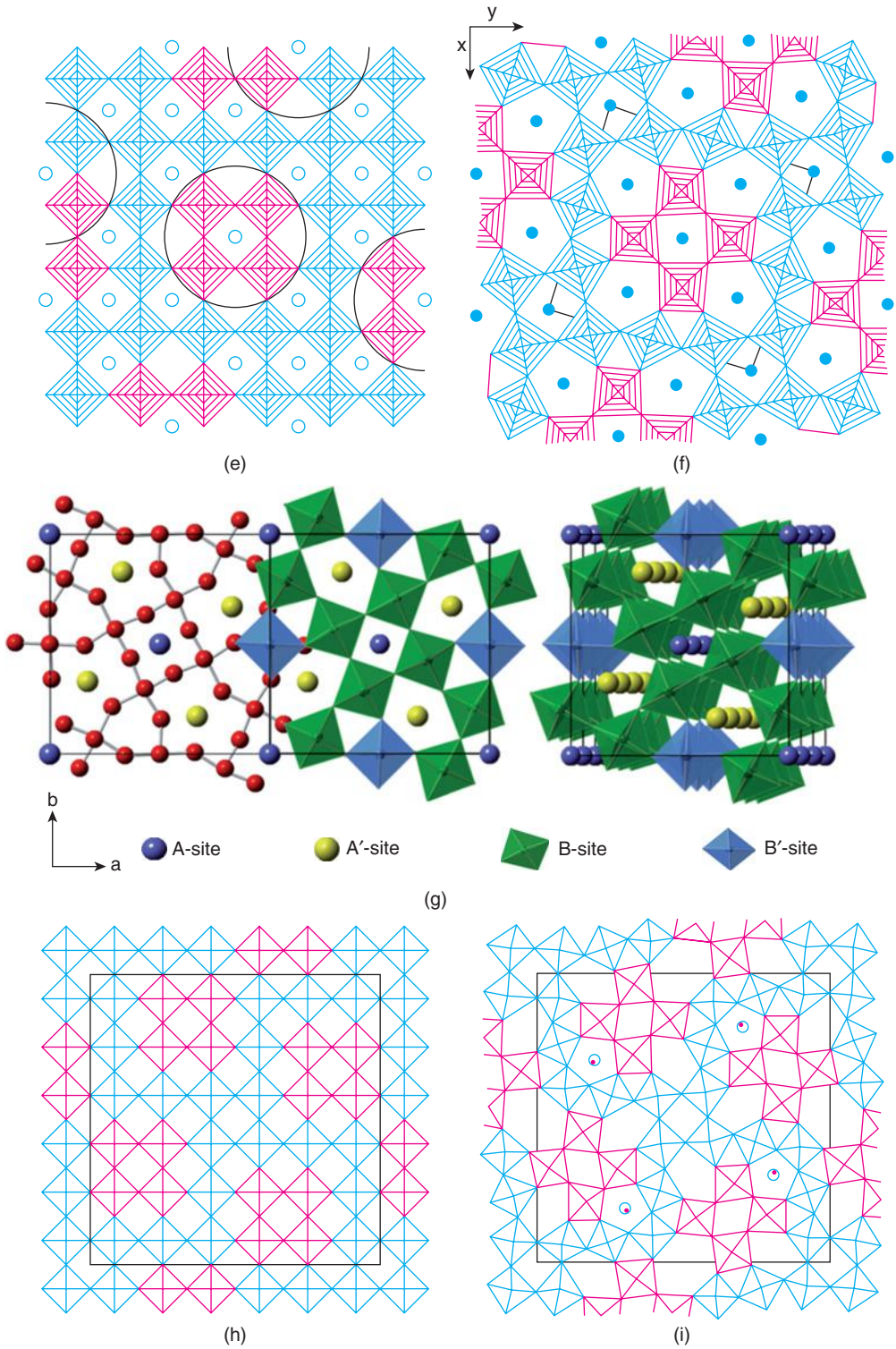


Figure 1.43 **(Continued)**

Table 1.21 ***Some compounds with the tetragonal tungsten bronze, TTB, structure. Unit cell: tetragonal a ≈ 12.4, c ≈ 3.9 Å***

A_2	A′	B_4B'	O_{15}
Ba_2	RE[a]	Ti_2, Nb_3	O_{15}
Sr_2	Na or K	Nb_5	O_{15}
Ba_2	Na	Nb_5	O_{15}
Ba_2 or Sr_2	Ba or Sr	Ti, Nb_4	O_{15}
$Sr_{2.5}$		Nb_5	O_{15}
$K_{0.4–0.6}$		W	O_3

[a]RE = La, . . . , Dy, Bi.

This contains an important family of ferroelectric materials although, since they are insulating, it is rather a misnomer to call them bronzes.

The general formula of the ideal TTB structure is $A_2A'B_4B'O_{15}$ or $A_2A'B_3B'_2O_{15}$, depending on the oxidation states of A, A′, B and B′. The structure is built of corner-sharing BO_6 and $B'O_6$ octahedra; the B,B′:O ratio is 1:3, as in ReO_3, but the linkage of the octahedra is different from that in ReO_3. In the TTB structure, blocks of octahedra which form four-membered rings are rotated by 90° relative to the parent ReO_3 structure to give a framework containing triangular, square and pentagonal columns, Fig. 1.43(e) and (f). In the *c* direction, the unit cell is only one octahedron thick. In principle, all the sites are available for cation occupancy. In practice, large cations such as K^+ and Ba^{2+} occupy the pentagonal tunnels with a coordination number to oxygen of 15, whereas somewhat smaller cations such as La^{3+} occupy square tunnels with coordination number 12; the triangular tunnels, C in the formula, remain either empty or occupied by smaller ions such as Li^+.

A selection of compounds with the TTB structure is shown in Table 1.21. The structure is versatile since cations on the A and A′ sites can be either fully ordered, disordered or only partially occupied. In addition, there are two crystallographically distinct, octahedral B and B′ sites, which can be occupied by one species or a mixture. As with all crystal structures, it may be useful to view them in different ways and the TTB structure is shown from three different perspectives in Fig. 1.43(g).

The structure of Mo_5O_{14} is based on the same principles but is rather different to the TTB structure, Fig. 1.43(h) and (i). Pentagonal columns are again present and these are occupied by rows of Mo and O atoms, as shown by open and closed circles in (i). Because of the extra oxygens in the columns, Mo occupies small sites whose coordination number is reduced greatly from 15 to 7. Other pentagonal columns, and triangular columns, remain empty. In addition, new hexagonal columns are created by the array of octahedra and these remain empty in Mo_5O_{14}. Many other complex structures are built using similar principles, for example, $NaNb_6O_{15}F$, $Nb_{16}W_{18}O_{94}$, $Bi_6Nb_{34}O_{94}$ and Ta_3O_7F.

1.17.9 Spinel

Several commercially important magnetic oxides have the spinel structure. The parent spinel is $MgAl_2O_4$. It has *ccp* oxide ions with Mg^{2+}, Al^{3+} in tetrahedral and octahedral interstices, respectively. Many oxides, sulfides and halides have the spinel structure and different cation charge combinations are possible, namely:

2, 3	as in	$MgAl_2O_4$
2, 4	as in	Mg_2TiO_4
1, 3, 4	as in	$LiAlTiO_4$
1, 3	as in	$Li_{0.5}Al_{2.5}O_4$
1, 2, 5	as in	$LiNiVO_4$
1, 6	as in	Na_2WO_4

Similar cation combinations occur with sulfides, e.g. 2, 3: $ZnAl_2S_4$ and 2, 4: Cu_2SnS_4. With halides, cations are limited to charges of 1 and 2, in order to give an overall cation:anion ratio of 3:4, e.g. Li_2NiF_4.

$MgAl_2O_4$ has a large, cubic unit cell with $a = 8.08$ Å. The cell contents are eight formula units ($Z = 8$) corresponding to '$Mg_8Al_{16}O_{32}$'. In the *ccp* oxide array, half of the octahedral sites are occupied by Al and one-eighth of the tetrahedral sites, both T_+ and T_-, by Mg.

The unit cell of the spinel structure is a large cube, eight times (i.e. $2 \times 2 \times 2$) the size of a typical face centred cube, and it is almost impossible to give a comprehensible 3D drawing of the complete unit cell and its contents. We can, however, easily understand how the structure arises and draw structural fragments to illustrate certain features.

Writing the general formula of a *normal spinel* as $A^{tet}B_2{}^{oct}O_4$, let us first consider the $B_2{}^{oct}O_4$ part. This is simply the rock salt structure with *ccp* O^{2-} ions but with only alternate octahedral sites occupied by B cations. Two such rock salt-derivative subcells are shown in Fig. 1.44(a) and (b) (each subcell forms one octant of the spinel unit cell); the bottom face of the subcell in (a) is the same as the top face of the subcell in (b). Oxygens are shown at corners and face centres (as in Fig. 1.24). Some octahedral positions are shown as occupied (solid circles) but alternate ones in any of the three unit cell directions are empty (small squares).

The base of the spinel unit cell is shown in (c); it is formed by the base of (b), dashed, and that of three adjacent subcells. The alternating, empty–occupied–empty sequence of octahedral sites is clearly seen and also occurs similarly in the third dimension (a, b).

To complete the structure, we need to locate the tetrahedral A cations. In an *fcc* anion array, Fig. 1.24, the eight tetrahedral sites, 5–12, are located as shown. These tetrahedral sites are equidistant from oxygen atoms but also from the octahedral cation sites (e.g. distances 9–A and 9–3 in Fig. 1.24 are equal). Cation–cation repulsions do not allow adjacent tetrahedral and octahedral sites to be occupied simultaneously (this would be equivalent to tetrahedra and octahedra sharing a common face). We therefore need to find those tetrahedral sites for which all four neighbouring octahedral sites are empty. In Fig. 1.44(a), *all* of the tetrahedral sites (5–12 in Fig. 1.20) have at least one neighbouring octahedral cation and so none of the tetrahedral sites in this octant is occupied. The octant immediately below (or above) that in (a) is shown in (b); note that its top face is the same as the bottom face in (a). In this octant, two of the tetrahedral sites (positions 8 and 9 in Fig. 1.24) have no octahedral cation neighbours and, therefore, both of these contain an A cation. Taking the spinel structure as a whole, the eight octants fall into two structural groups, as represented by (a) and (b). The A cations are located in the four octants of type (b).

The two types of octant are arranged such that they alternate in any of the three unit cell directions. They are, in fact, arranged in exactly the same way as anions and cations in the rock salt structure, Fig. 1.44(d). A projection of the unit cell onto one face, showing cation positions only, is shown in (e). The orientation is exactly the same as in (b) and (c) but now atom heights, above the plane of the paper, are given as multiples of $c/8$. Octahedral positions are solid circles; occupied tetrahedral positions are represented by A. You should be able to see that the octahedral atoms O′ in the base (height 0) are the same as in (b) and (c). An alternative view of the structure of $MgAl_2O_4$ spinel is shown in (f). It consists of a framework of edge-sharing AlO_6 octahedra with isolated MgO_4 tetrahedra inside channels formed by the octahedral array.

The above description of the spinel structure is somewhat idealised. In practice, the anions are displaced slightly from their corner and face centre positions. The degree of displacement varies slightly from compound to compound, which means that the bond lengths to A and B cations can change so as to fit best those required by the particular A and B ions. The origin of the spinel structure is usually shifted to coincide with a tetrahedral A cation. We need not pursue this any further as it may lead to confusion, but it does lead to alternative descriptions of the spinel structure (e.g. the A cations alone are arranged in the same way as the carbon atoms in the diamond structure!).

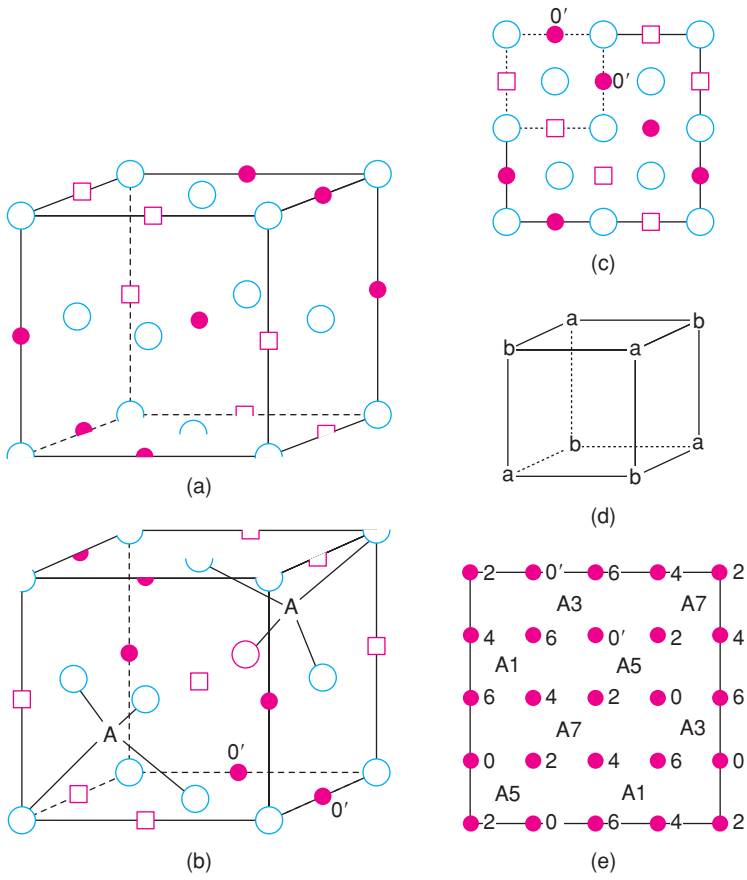


Figure 1.44 ***Representative parts of the spinel structure. (a) One octant of the unit cell showing oxygens at corner and face centres, empty (□) and occupied (●) octahedral sites. (b) A second octant, underneath the one in (a), showing in addition the occupation of two tetrahedral sites, A. (c) One face of the cubic unit cell of the spinel structure. The dashed part coincides with the base of the subcell shown in (b). (d) Alternating arrangement of the two types of octant (a) and (b). (e) Cation positions in spinel. Numbers refer to fractional heights, as multiples of c/8. Octahedral cation sites O′ are also shown in (b) and (c).***

A complicating factor in some spinel structures is that the cation distribution may vary. Two extreme types of behaviour may be distinguished. In *normal* spinels, the cations occupy sites given by the formula

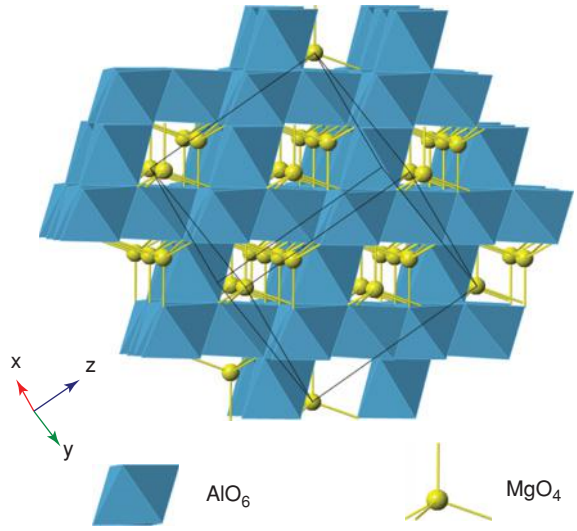
$$[A]^{tet}[B_2]^{oct}O_4$$

i.e. with A in tetrahedral sites and B in octahedral sites. Examples of normal spinels are $MgAl_2O_4$ and $MgTi_2O_4$. In *inverse* spinels, half of the B ions occupy tetrahedral sites, leaving the remaining B ions and all the A ions in octahedral sites, i.e.

$$[B]^{tet}[A,\ B]^{oct}O_4$$

Usually the A and B cations in octahedral sites are disordered. Examples of inverse spinels are $MgFe_2O_4$ and Mg_2TiO_4.

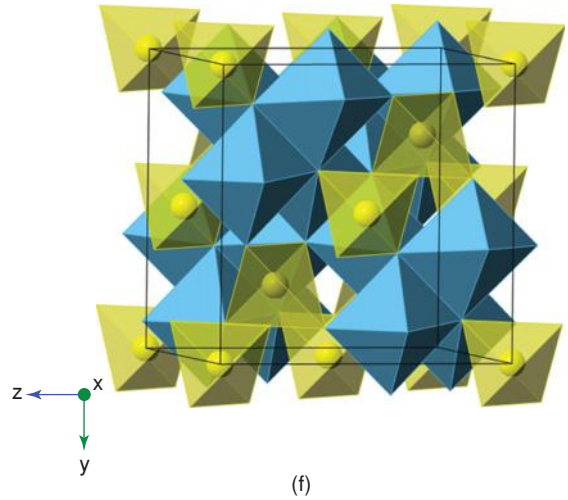
Spinel ($MgAl_2O_4$)



Atom	Wyckoff positions	x	y	z
Mg	8a	0	0	0
Al	16d	0.625	0.625	0.625
O	32e	0.3873(1)	0.3873(1)	0.3873(1)

Cubic : a = 8.0806 Å

Space Group : $Fd\bar{3}m$ (No. 227)



(f)

Figure 1.44 **(Continued)** ***(f) Two perspectives of the spinel structure of $MgAl_2O_4$ showing a framework of AlO_6 octahedra with MgO_4 tetrahedra in channel sites.***

Table 1.22 ***Some compounds with the spinel structure***

Compound	Type	a/Å	Structure	Compound	Type	a/Å	Structure
$MgAl_2O_4$	2, 3	8.0800	Normal	$MgIn_2O_4$	2, 3	8.81	Inverse
$CoAl_2O_4$	2, 3	8.1068	Normal	$MgIn_2S_4$	2, 3	10.708	Inverse
$CuCr_2S_4$	2, 3	9.629	Normal	Mg_2TiO_4	2, 4	8.44	Inverse
$CuCr_2Se_4$	2, 3	10.357	Normal	Zn_2SnO_4	2, 4	8.70	Inverse
$CuCr_2Te_4$	2, 3	11.051	Normal	Zn_2TiO_4	2, 4	8.467	Inverse
$MgTi_2O_4$	2, 3	8.474	Normal	$LiAlTiO_4$	1, 3, 4	8.34	Li in tet
Co_2GeO_4	2, 4	8.318	Normal	$LiMnTiO_4$	1, 3, 4	8.30	Li in tet
Fe_2GeO_4	2, 4	8.411	Normal	$LiZnSbO_4$	1, 2, 5	8.55	Li in tet
$MgFe_2O_4$	2, 3	8.389	Inverse	$LiCoSbO_4$	1, 2, 5	8.56	Li in tet
$NiFe_2O_4$	2, 3	8.3532	Inverse				

In addition to normal and inverse spinels, a complete range of intermediate cation distributions is possible and, in some cases, the distribution changes with temperature. The cation distribution may be quantified using a parameter, γ, which is the fraction of A ions on the octahedral sites:

- normal: $[A]^{tet}[B_2]^{oct}O_4$ $\gamma = 0$
- inverse: $[B]^{tet}[A, B]^{oct}O_4$ $\gamma = 1$
- random: $[B_{0.67}A_{0.33}]^{tet}[A_{0.67}B_{1.33}]^{oct}O_4$ $\gamma = 0.67$.

The cation distribution in spinels and the *degree of inversion*, γ, have been studied in considerable detail. Several factors influence γ, including the site preferences of ions in terms of size, covalent bonding effects and crystal field stabilisation energies (see Chapter 2). The γ value in any particular spinel is given by the net effect of these various parameters taken together. Some compounds with the spinel structure are given in Table 1.22.

1.17.10 Olivine

The olivine structure, typified by the minerals forsterite, Mg_2SiO_4, and triphylite, $LiFePO_4$, is the *hcp* analogue of the spinel structure. One-eighth of the tetrahedral sites are occupied by Si or P and half of the octahedral sites by Mg or Li, Fe within an *hcp* oxide array. There are two crystallographically distinct octahedral sites in olivine, which are occupied in ordered fashion by Li and Fe in $LiFePO_4$. The crystal structure is shown in Fig. 1.45 and some olivines are listed in Table 1.23. Olivines occur mainly with oxides but also with sulfides, selenides and some fluorides. Various cation charge combinations occur, such that in oxides the three cations have a net charge 8+.

Olivines (mainly forsterite and fayalite) are believed to be the main mineralogical constituent of the Earth's upper mantle. At high pressures, many olivines transform to the spinel structure and spinels are probably the main constituent of the Earth's lower mantle. Volume changes associated with the olivine to spinel phase transformation may have had fundamental geological consequences during the evolution of the Earth, involving the formation of mountain ranges and under-sea ridges. When spinel material from the lower mantle was pushed upwards to the Earth's surface, it transformed to olivine due to the reduction in pressure. The reverse transformation, olivine to spinel, with a volume contraction may be a contributing factor to earthquakes.

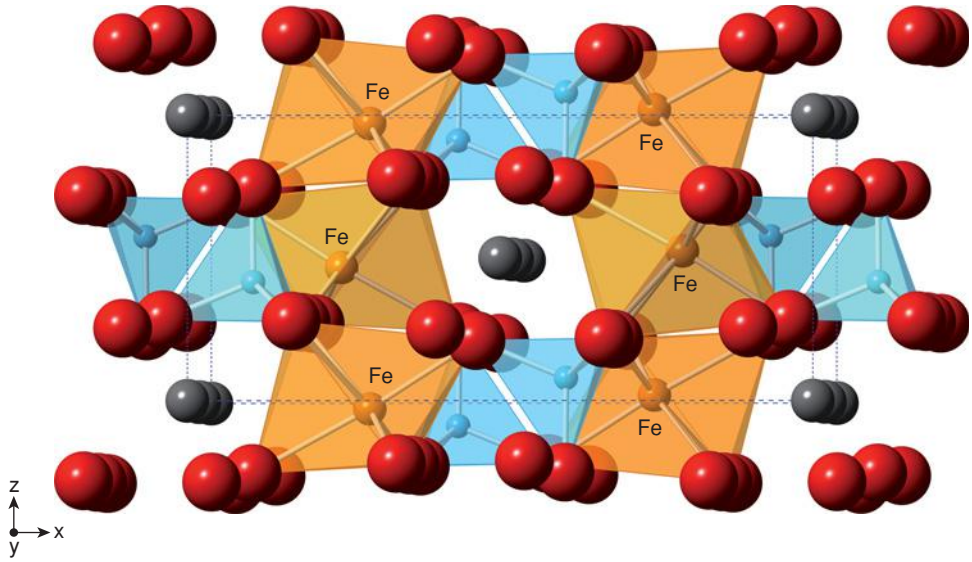


Figure 1.45 ***Olivine structure of $LiFePO_4$. Adapted from J. J. Biendicho and A. R. West,* Solid State Ionics*, 203, 33–36 © 2011, Elsevier. Reproduced under the terms of the STM Agreement.***

Table 1.23 ***Some compounds with the olivine structure***

General formula			
Octahedral site	Tetrahedral site	*hcp* anion	Examples
II_2	IV	O_4	Mg_2SiO_4 (forsterite) Fe_2SiO_4 (fayalite) Ca,Mg SiO_4 (monticellite) γ-Ca_2SiO_4 A_2GeO_4: A = Mg, Ca, Sr, Ba, Mn
III_2	II	O_4	Al_2BeO_4 (chrysoberyl) Cr_2BeO_4
II, III	III′	O_4	$MgAlBeO_4$
I, II	V	O_4	$LiFePO_4$ (triphylite) $LiMnPO_4$ (lithiophylite)
I, III	IV	O_4	$LiRESiO_4$: RE = Ho, . . . , Lu $NaREGeO_4$: RE = Sm, . . . , Lu $LiREGeO_4$: RE = Dy, . . . , Lu
II_2	IV	S_4	Mn_2SiS_4 Mg_2SnS_4 Ca_2GeS_4
I_2	II	F_4	γ-Na_2BeF_4

Unit cell: orthorhombic; for $LiFePO_4$, $a = 10.33$, $b = 6.01$, $c = 4.70$ Å; $Z = 4$.

There is much current interest in $LiFePO_4$ as the cathode material in rechargeable lithium batteries. On charging, Li deintercalates from the structure and a corresponding amount of Fe is oxidised from Fe^{2+} to Fe^{3+}, as follows:

$$LiFe^{2+}PO_4 - \underset{-xe^-}{\overset{-xLi^+}{\longrightarrow}} Li_{1-x}Fe^{2+}_{1-x}Fe^{3+}_{x}PO_4$$

Lithium ions occupy channels parallel to the *y* axis, Fig. 1.45, which allows them to leave and enter the structure readily during cell charge and discharge. This is an example of a solid-state redox reaction with a cell potential of about 3.08 V. $LiFePO_4$, and associated $LiMnPO_4$, are of interest because the redox reaction and the process of lithium removal and insertion are reversible over many cycles, giving a high cell capacity, and the materials are cheap, non-toxic and environmentally friendly.

1.17.11 Corundum, ilmenite and $LiNbO_3$

These three closely related structure types can be regarded, ideally, as *hcp* oxide ions with cations occupying two-thirds of the octahedral sites. Conceptually, they are related to the NiAs structure in which all the octahedral sites are occupied, and to the CdI_2 structure in which only half the octahedral sites are occupied, Table 1.4. The crystal structures are shown in Fig. 1.46 and some compounds adopting these structures are listed in Table 1.24. Corundum contains only one cation, Al^{3+}, whereas ilmenite contains two cations that are ordered over the octahedral sites that are occupied by Al in corundum. In $LiNbO_3$, the same octahedral sites are occupied but the cation ordering arrangement is different.

The unit cell of all three structures is hexagonal and has six *cp* oxygen layers parallel to the basal plane, shown in Fig. 1.46(a) at *c* heights 1/12, 3/12, 5/12, 7/12, 9/12 and 11/12. Cations are in octahedral sites mid-way between the oxygen layers; alternate layers of cation sites are occupied by Fe and Ti in ilmenite, Fig. 1.46(b). Pairs of octahedra share a common face in the *c* direction and cation repulsion between the cation pairs causes distortion from an ideal *hcp* structure. In all cases, the cation octahedra are distorted with three long and three short bonds. Repulsion between Nb^{5+} and Li^+ in $LiNbO_3$ causes displacement of Li to a position near the triangular face at the opposite side of the octahedron. $LiNbO_3$ and $LiTaO_3$ are ferroelectric materials and cation displacements within the face-sharing octahedra are responsible for the polar crystal structures and dipole reorientation in an applied electric field, which is a characteristic feature of ferroelectrics.

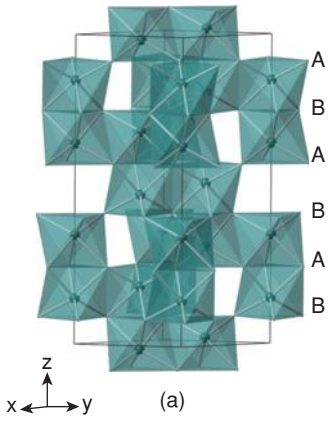
The cation ordering sequence in $LiNbO_3$ is different to that in ilmenite, Fig. 1.46(c) and (d). Li and Nb are both present, ordered, between any pair of close packed oxide ion layers whereas Fe and Ti occupy alternate sets of layers in ilmenite. An alternative view of the $LiNbO_3$ structure is given in (d), which illustrates that $LiNbO_3$ can also be regarded as a grossly distorted perovskite structure. Tilting and rotation of the NbO_6 octahedra (B sites) reduce the coordination of the A sites from 12 to distorted octahedral, and these are occupied by Li. If we regard $LiNbO_3$ as a distorted perovskite, its tolerance factor is 0.78, which, in practice, represents the lower limit for materials that can be regarded as distorted perovskites.

1.17.12 Fluorite-related structures and pyrochlore

The fluorite structure of CaF_2 can be described as eutactic *ccp* Ca^{2+} ions with F^- ions occupying all tetrahedral sites, Fig. 1.29. A number of more complex fluorite-related structures occur with an excess or deficiency of anions or with cation ordering.

A remarkable anion-excess fluorite is LaH_3; LaH_2 forms the basic fluorite structure and extra hydrogens occupy fully the octahedral sites. The structure therefore has all tetrahedral and octahedral sites occupied by H within a *ccp* La array, Fig. 1.24. A similar structure is observed in intermetallics such as Li_3Bi and Fe_3Al.

Corundum (Al_2O_3)

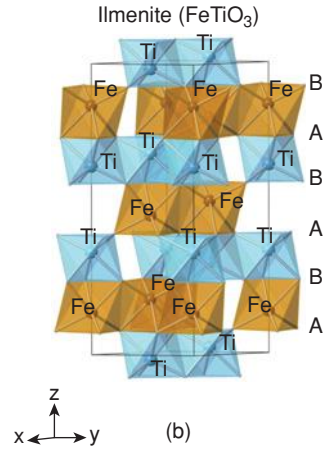


Atom	Wyckoff positions	x	y	z
Al	12c	0	0	0.3520(1)
O	18e	0.3060(6)	0	0.25

Trigonal : a = 4.7538, c = 12.9725 Å

Space Group : R$\bar{3}$c (No. 167)

Ilmenite ($FeTiO_3$)

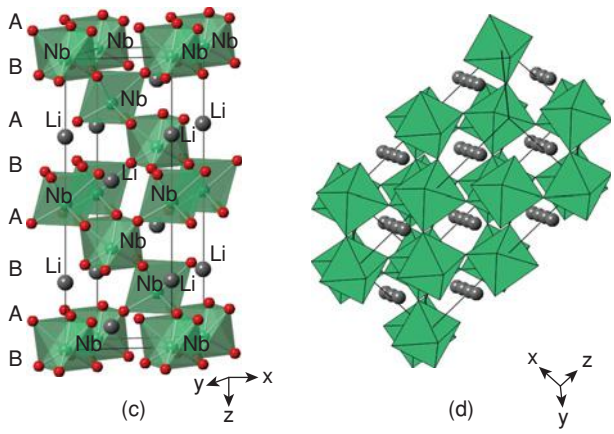


Atom	Wyckoff positions	x	y	z
Fe	6c	0	0	0.1446(1)
Ti	6c	0	0	0.3536(1)
O	18f	0.295(1)	–0.022(1)	0.2548(3)

Trigonal : a = 5.087, c = 14.042 Å

Space Group : R$\bar{3}$ (No. 148)

Lithium Niobate ($LiNbO_3$)



Atom	Wyckoff positions	x	y	z
Li	6a	0	0	0.268(11)
Nb	6a	0	0	0
O	18b	0.0.0437(69)	0.3309(82)	0.0636(12)

Trigonal : a = 5.1483, c = 13.8631 Å

Space Group : R$\bar{3}$c (No. 161)

Figure 1.46 ***Crystal structures of (a) corundum, (b) ilmenite and (c, d) $LiNbO_3$.***

Table 1.24 ***Some compounds with corundum-related structures***

Corundum (α-alumina)	M_2O_3: M = Al, Cr, Fe (hematite), Ti, V, Ga, Rh
	Al_2O_3: with Cr dopant (ruby)
	Al_2O_3: with Ti dopant (sapphire)
Ilmenite	$MTiO_3$: M = Mg, Mn, Fe, Co, Ni, Zn, Cd
	$MgSnO_3$, $CdSnO_3$
	$NiMnO_3$
	$NaSbO_3$
$LiNbO_3$, $LiTaO_3$	

These structures represent an extreme with full occupancy of all tetrahedral and octahedral sites. Partial tetrahedral site occupancies occur in the lanthanum hydrides which form a continuous solid solution between $LaH_{1.9}$ and LaH_3.

Oxygen-excess fluorites occur in UO_{2+x} (see Fig. 2.10); the structure is distorted locally and the extra oxide ions are displaced off cube body centres. This has a knock-on effect in which some of the corner oxide ions are displaced onto interstitial sites. The UO_{2+x} system has been studied in considerable detail because of its importance in the nuclear industry as a fuel in nuclear reactors.

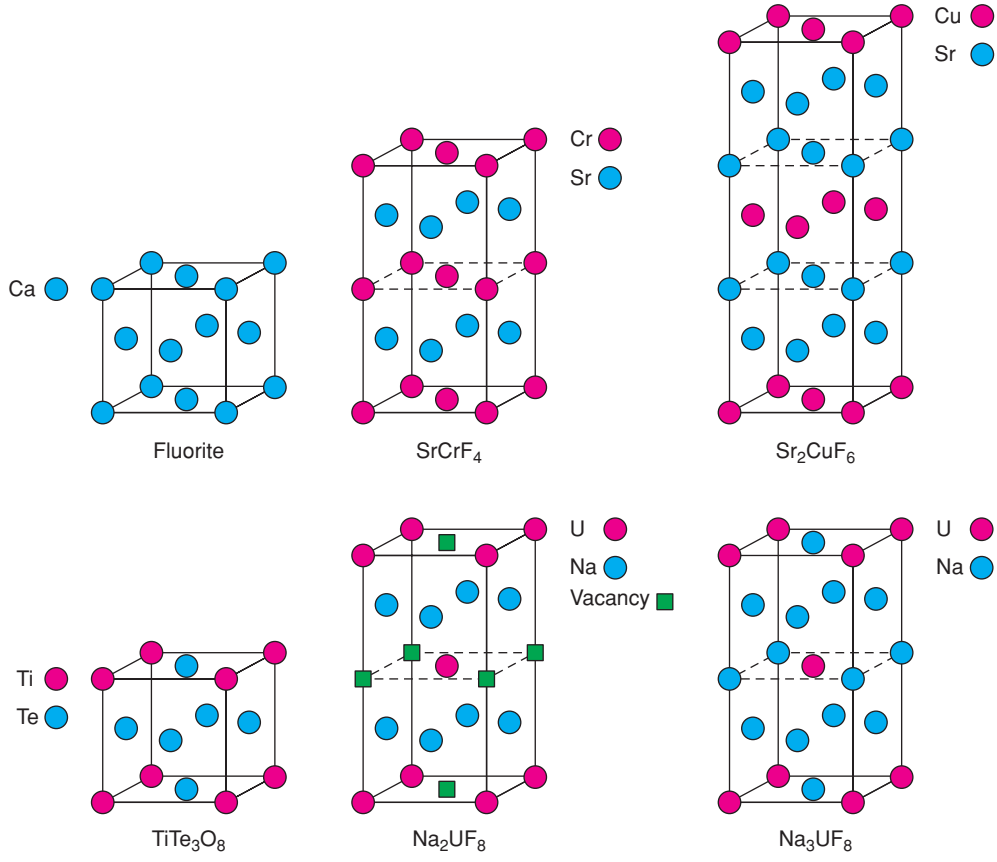


Figure 1.47 ***Some cation-ordered fluorites showing cation positions relative to those in*** **fcc** ***fluorite. Adapted with permission from A. F. Wells,*** **Structural Inorganic Chemistry,** ***© 1984 Oxford University Press.***

Mixed anion oxyfluorides such as LaOF and SmOF form the fluorite structure in which similarly sized O^{2-} and F^- ions are disordered over the tetrahedral sites. Various examples of mixed-cation fluorites are known in which two different cations are ordered, as shown for several examples in Fig. 1.47. These structures are rather idealised, however, since the anions are displaced off the centres of the tetrahedral sites in various ways to give, for instance, a distorted tetrahedral environment for Cr^{2+} in $SrCrF_4$ and distorted octahedral coordination for both Ti and Te in $TiTe_3O_8$.

Li_2O has the antifluorite structure and various Li-deficient antifluorites are good Li^+ ion conductors; for instance, $Li_9N_2Cl_3$ has 10% of the Li^+ sites vacant, giving rise to high Li^+ ion mobility.

The *pyrochlore* structure may be regarded as a distorted, anion-deficient fluorite with two different-sized cations A and B. Its formula is written as either $A_2B_2X_7$ or $A_2B_2X_6X'$. The unit cell is cubic with a ~11 Å and contains eight formula units. In principle, the structure is simple since, ideally, it is a fluorite with one-eighth of the tetrahedral anion sites empty. Also, there is only one variable positional parameter, the x fractional coordinate of the 48 X atoms in the position (x, 1/8, 1/8), etc., Fig. 1.48. Various compounds form the pyrochlore structure and their differences depend on the value of x, which is usually in the range 0.31 to 0.36. In the extreme case that x = 0.375, the structure is derived from an undistorted fluorite with 8-coordinate A cations, as in fluorite, but grossly distorted BX_6 octahedra. As x decreases, the 48 X atoms move off their regular tetrahedral sites and the cubic coordination of A becomes distorted: six X neighbours form a puckered hexagon and two X′ atoms are at a different distance, in apical positions, on either side of the puckered hexagon.

At x = 0.3125, the B coordination becomes undistorted octahedral and the BX_6 octahedra link by sharing corners to form a 3D network. The A coordination may be described as 2X′ + 6X, with short A-X′ bonds. These A-X′ bonds form a 3D network, A_2X', that interpenetrates the network of BX_6 octahedra, of stoichiometry B_2X_6. The A_2X' net, with linear X′-A-X units and $X'A_4$ tetrahedra, is similar to that in cuprite, Cu_2O. The B_2X_6 and A_2X' networks, that together form the pyrochlore structure, are shown separately in Fig. 1.48.

Pyrochlore-based oxides have a range of interesting properties and applications. $La_2Zr_2O_7$ is an electronic insulator whereas $Bi_2Ru_2O_{7-\delta}$ is metallic and $Cd_2Re_2O_7$ is a low temperature superconductor. $(Gd_{1.9}Ca_{0.1})Ti_2O_{6.9}$ is an oxide ion conductor and $Y_2Mo_2O_7$ is a spin glass. As indicated in some of the above examples, the anion content is not always 7 and indeed the amount of X′ in the general formula can have values between 0 and 1 in different pyrochlores.

1.17.13 Garnet

The garnets are a large family of complex oxides, including both minerals and synthetic materials, some of which are important ferromagnetic materials; yttrium aluminium garnet (YAG) is used as a synthetic gemstone and when doped with neodymium is the key component in YAG lasers; more recently, garnets with high Li^+ ion conductivity have been synthesised; garnets have hardness 6.5–7.5 on the Mho scale and are used industrially as an abrasive on sandpaper. They have the general formula $A_3B_2X_3O_{12}$: A = Ca, Mg, Fe, etc.; B = Al, Cr, Fe, etc.; X = Si, Ge, As, V, etc. A is a large ion with a radius of ~1 Å and has a coordination number of eight in a distorted cubic environment. B and X are smaller ions which occupy octahedral and tetrahedral sites, respectively. Garnets with A = Y or a rare earth: Sm, Gd, Tb, Dy, Ho, Er, Tm, Yb or Lu and B, X = Fe^{3+} have interesting magnetic properties. One of the most important is yttrium iron garnet (YIG), $Y_3Fe_5O_{12}$, the structure of which is shown in Fig. 1.49. Many other A, B, X combinations are possible, such as those in Table 1.25.

The unit cell of garnet is body centred cubic, $a \approx 12.4$ Å, and contains eight formula units. The structure may be regarded as a framework built of corner-sharing BO_6 octahedra and XO_4 tetrahedra. The larger A ions occupy eight-coordinate cavities within this framework. In YIG and the rare earth garnets, the B and X ions are the same, Fe^{3+}.

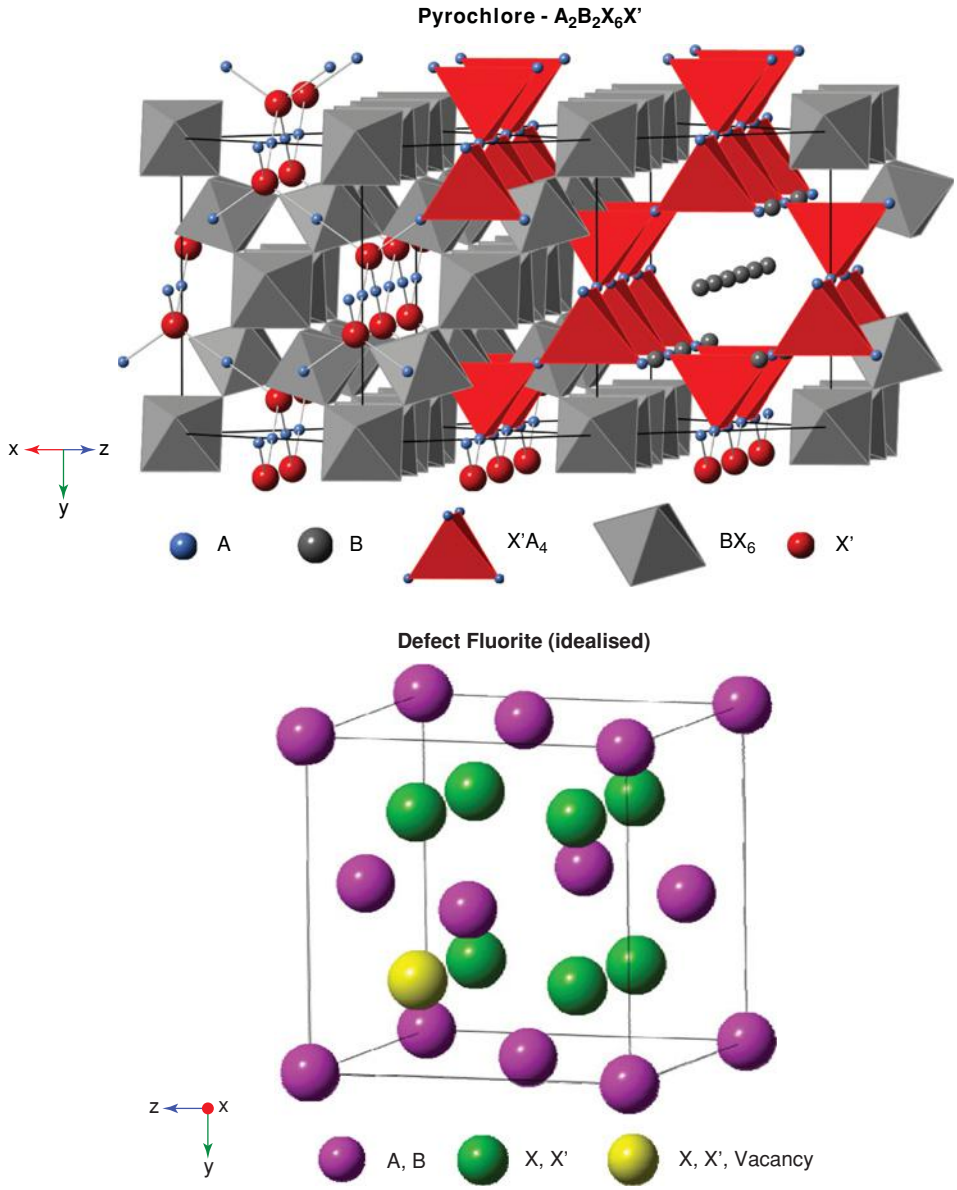
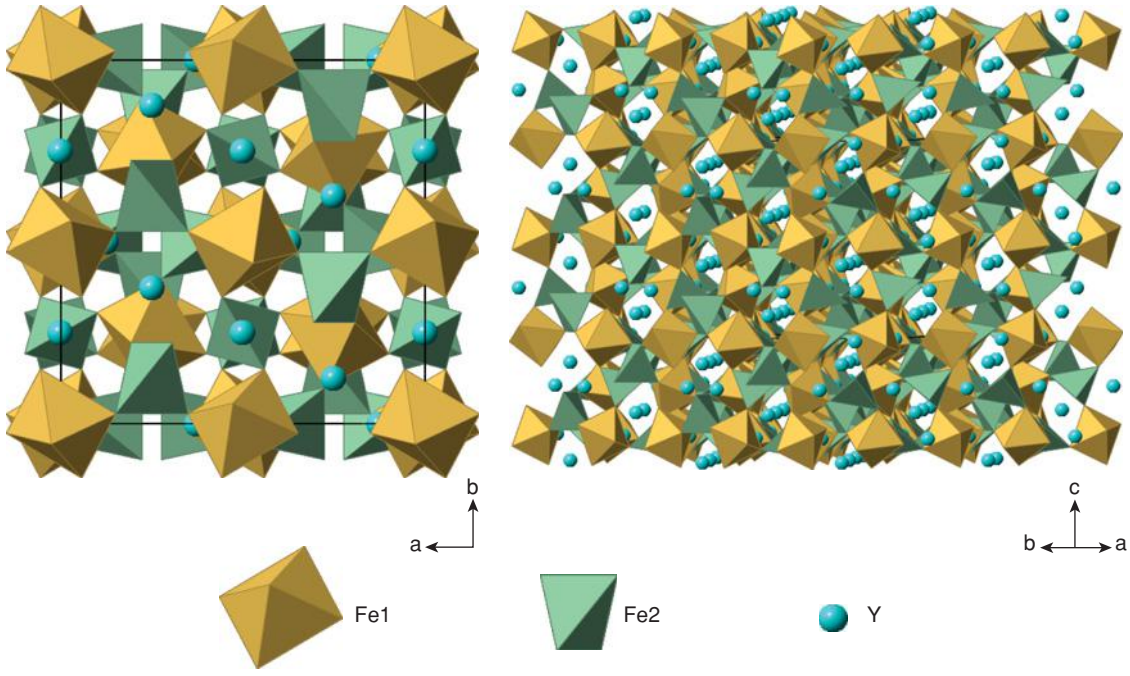


Figure 1.48 ***The pyrochlore structure, which may be regarded as a distorted, 2 × 2 × 2 superstructure of a cation-ordered, anion-deficient fluorite. Anions in eightfold positions, e.g. ($^1/_4$, $^1/_4$, $^1/_4$) in fluorite split into two groups in pyrochlore, with one group, X, in 48-fold positions at e.g. (x, 1/8, 1/8).***

1.17.14 Perovskite-rock salt intergrowth structures: K_2NiF_4, Ruddlesden–Popper phases and layered cuprate superconductors

The K_2NiF_4 structure is the simplest structure of a large family of related materials that has attracted much attention in recent years because several of them are superconductors. The structure may be regarded as alternating layers of perovskite and rock salt structures, as shown in Fig. 1.50. The formula K_2NiF_4 could be

Garnet ($Y_3Fe_5O_{12}$)



Atom	Wyckoff positions	x	y	z
Y	24c	0.1250	0	0.25
Fe1	16a	0	0	0
Fe2	24d	0.375	0	0.25
O	96h	–0.0271(1)	0.0567(1)	0.1504(1)

Cubic : a = 12.376 Å

Space Group : $Ia\bar{3}d$ (No. 230)

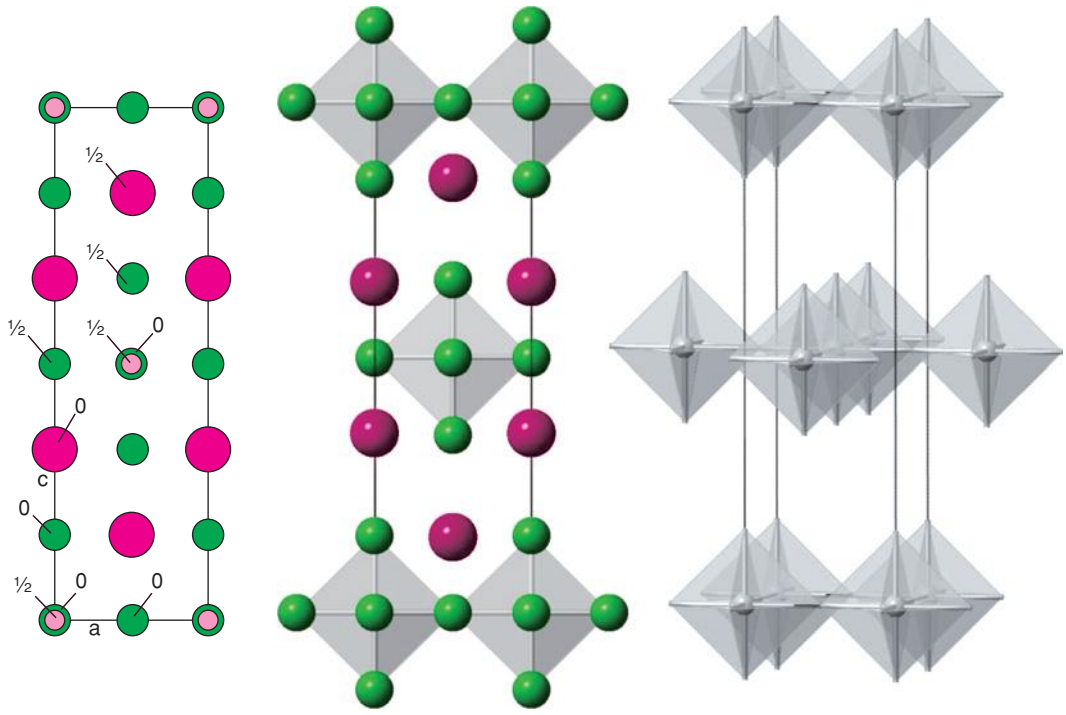
Figure 1.49 ***The garnet crystal structure.***

written in expanded form as $KNiF_3.KF$ to indicate the perovskite and rock salt components. The structure is body centred tetragonal with perovskite-like layers of octahedra centred at $c = 0$ and $c = {}^1/_2$. K^+ ions lie at the interface between rock salt and perovskite blocks and have a coordination number of 9. The coordination number of K in rock salt would be 12 but the rock salt blocks are less than one unit cell thick and the K coordination in rock salt is interrupted by the atomic displacements that generate perovskite blocks. The K_2NiF_4 structure occurs with a range of fluorides and chlorides containing small divalent cations, Table 1.26, and with a range of oxides with different charge combinations such as $A_2{}^+B^{6+}O_4$, $A_2{}^{2+}B^{4+}O_4$, $A_2{}^{3+}B^{2+}O_4$, $A^{2+}A'^{3+}B^{3+}O_4$ and $A_2{}^{3+}B_{1/2}{}^+B'_{1/2}{}^{3+}O_4$ containing large A and small B cations; probably other cation combinations are also possible.

A-site cation ordering occurs in $A^+A'^{3+}B^{4+}O_4$: A = Na, A′ = La → Lu, B = Ti; Na and the lanthanides occupy different layers in the c direction. Anion-deficient and anion-excess K_2NiF_4 structures exist. For

Table 1.25 ***Examples of A, B, X combinations in garnets***

Garnet	A	B	X	O
Grossular	Ca_3	Al_2	Si_3	O_{12}
Uvarovite	Ca_3	Cr_2	Si_3	O_{12}
Andradite	Ca_3	Fe_2	Si_3	O_{12}
Pyrope	Mg_3	Al_2	Si_3	O_{12}
Almandine	Fe_3	Al_2	Si_3	O_{12}
Spessartine	Mn_3	Al_2	Si_3	O_{12}
	Ca_3	CaZr	Ge_3	O_{12}
	Ca_3	Te_2	Zn_3	O_{12}
	Na_2Ca	Ti_2	Ge_3	O_{12}
	$NaCa_2$	Zn_2	V_3	O_{12}



Ni: 0 0 0; ½ ½ ½

K: ± 0 0 z; ± ½ ½ z + ½: z = 0.35

F(1): ½ 0 0; 0 ½ 0; 0 ½ ½, ½ 0 ½
F(2): ± 0 0 z; ± ½ ½ z + ½: z = 0.15

Tetragonal: a = 4.006 Å, c=13.076 Å.

Space group: 14 / mmm (139)

Figure 1.50 ***The K_2NiF_4 structure.***

Table 1.26 ***Some compounds with the K_2NiF_4 structure***

Compound[a]	*a*/Å	*c*/Å	*z* (M^+ ion)	*z* (anion)
K_2NiF_4	4.006	13.076	0.352	0.151
K_2CuF_4	4.155	12.74	0.356	0.153
Ba_2SnO_4	4.140	13.295	0.355	0.155
Ba_2PbO_4	4.305	13.273	0.355	0.155
Sr_2SnO_4	4.037	12.53	0.353	0.153
Sr_2TiO_4	3.884	12.60	0.355	0.152
La_2NiO_4	3.855	12.652	0.360	0.170
K_2MgF_4	3.955	13.706	0.35	0.15

Other examples[b]

$M_2Y^{6+}O_4$: M = K, Rb, Cs; Y = U, Np
Ln_2YO_4: Ln = La→Nd; Y = Ni, Cu
$CaLnAlO_4$: Ln = La → Er, Y
$SrLnFeO_4$: Ln = La → Tb
$SrLnCrO_4$: Ln = La → Dy
$BaLnFeO_4$: Ln = La → Eu
$La_2Li_{1/2}Y_{1/2}O_4$: Y = Ni, Co
A_2BF_4: A = K, Rb, Tl; B = Mg, Ni, Zn, Co, Fe
A_2BCl_4: A = Rb, Cs; B = Cr, Mn, Cd
Sr_2BO_4: B = Ti, Sn, Zr, Hf, Mo, Tc, Ir, Ru, Rh,Mn

[a]Data reproduced from R. G. Wyckoff, *Crystal Structures*, Vol. 3, © 1965 Wiley Interscience. Reproduced under the terms of the STM Agreement.
[b]Data taken from O. Muller and R. Roy, *The Major Ternary Structural Families*, © 1974 Springer-Verlag.

example, in Sr_2XO_3: X = Cu, Pd, one-quarter of the oxygens are removed, giving square planar coordination for X and seven-coordinate Sr, whereas in La_2NiO_{4+x}, interstitial oxygens are present.

The *Ruddlesden–Popper* series of compounds have the K_2NiF_4 structure as their simplest family member; other family members exist with more layers in the perovskite blocks. Thus, with strontium titanates, a family of phases exists with general formula $Sr_{n+1}Ti_nO_{3n+1}$ and consist of $n SrTiO_3$ perovskite blocks interleaved with SrO rock salt layers. Two members of this family are $Sr_3Ti_2O_7$ ($n = 2$) and $Sr_4Ti_3O_{10}$ ($n = 3$). The $n = \infty$ member corresponds to $SrTiO_3$ perovskite. These phases retain a tetragonal unit cell with similar *a* dimensions but have much longer *c* axes. A similar family of phases forms for $La_{n+1}Ni_nO_{3n+1}$: $n = 1, 2, 3, \infty$. In most cases, only one or two members of the Ruddlesden–Popper series are known, with examples such as

$Sr_3B_2O_7$: B = Ir, Ti, Zr, Cr, Fe

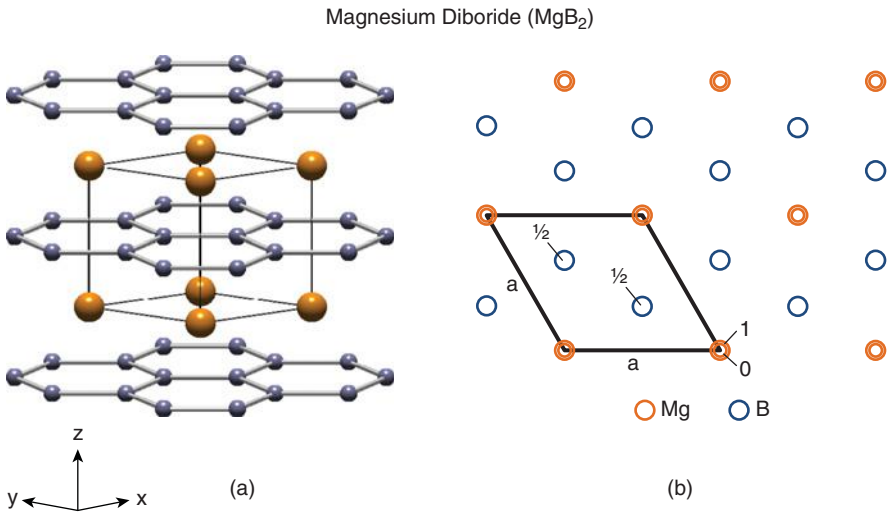
$Ca_3B_2O_7$: B = Mn, Ti

$Sr_4B_3O_{10}$: B = Ir, Ti, Zn

$Ca_4B_3O_{10}$: B = Mn, Ti

$K_3Zn_2F_7$, $K_3Fe_2F_7$, $Rb_3Cd_2F_7$, $K_3Mn_2Cl_7$, $Rb_3Mn_2Cl_7$, $Cs_3Ca_2Cl_7$

$Rb_4Cd_3Cl_{10}$.



Atom	Wyckoff positions	x	y	z
Mg B	1a 2d	0 0.3333 0.6667	0 0.6667 0.3333	0 0.5 0.5

Trigonal : a = 3.08489, c = 3.52107 Å

Space Group : P6 / mmm (No. 191)

Figure 1.51 ***The crystal structure of MgB_2 as (a) an oblique projection showing hexagonal rings of B atoms with 12-coordinate Mg situated between pairs of rings and (b) [001] projection of the crystal structure.***

1.17.15 The aluminium diboride structure (AlB_2)

The aluminium diboride family of crystal structures shot to prominence in 2000 when isostructural MgB_2 was discovered to be a superconductor with $T_c = 39$ K. It has a relatively simple crystal structure, shown in Fig. 1.51, in which Mg atoms form close packed layers stacked in an AAA sequence, which may be referred to as primitive hexagonal packing, *hp*. The *cp* Mg layers are separated by B layers arranged as in graphite; hence Mg is 12-coordinate with hexagonal rings of B atoms above and below. Each B has three B nearest neighbours in a trigonal planar arrangement and six Mg next nearest neighbours arranged in a trigonal prism.

Numerous borides and silicides have the AlB_2 structure, including MB_2: M = Ti, Zr, Nb, Ta, V, Cr, Mo, Mg, U, and MSi_2: M = U, Pu, Th. The crystal structure of some metal hydroxides such as $Cd(OH)_2$ is also closely related.

The AlB_2 structure is closely related to the NiAs structure. In both structures, the metal atoms form a primitive hexagonal array but only half the trigonal prismatic sites are occupied by As in NiAs whereas all trigonal prismatic sites are occupied by B in AlB_2. Consequently, the coordination of Ni is octahedral in NiAs instead of 12-coordinate Al in AlB_2.

1.17.16 Silicate structures – some tips to understanding them

Silicates, especially many minerals, often have complex formulae and structures. The purpose of this section is not to give a review of their structures but simply to show that a considerable amount of structural information may be obtained from their chemical formulae. Using certain guidelines, one can appreciate, without the necessity of remembering a large number of complex formulae, whether a particular silicate is a 3D framework structure, sheet-like, chain-like, etc.

It is common practice to regard silicate structures as composed of cations and silicate anions. Various silicate anions are possible, ranging from the extremes of isolated SiO_4^{4-} tetrahedra in orthosilicates such as olivine (Mg_2SiO_4), to infinite 3D frameworks, as in silica itself (SiO_2). The structures of silicate anions are based on certain principles:

1. Almost all silicate structures are built of SiO_4 tetrahedra.
2. The tetrahedra may link by sharing corners to form larger polymeric units.
3. No more than two SiO_4 tetrahedra may share a common corner (i.e. oxygen).
4. SiO_4 tetrahedra never share edges or faces with each other.

Exceptions to (1) are structures in which Si is octahedrally coordinated to O as in one polymorph of SiP_2O_7 and in high-pressure polymorphs of SiO_2 (coesite, stishovite). The number of these exceptions is very small, however, and we can regard SiO_4 tetrahedra as the normal building block in silicate structures. Guidelines (3) and (4) are concerned, respectively, with maintaining local electroneutrality and with ensuring that highly charged cations, such as Si^{4+}, are not too close together.

The important factor in relating formula to structure type is the Si:O ratio. This ratio is variable since two types of O may be distinguished in the silicate anions: *bridging oxygens* and *non-bridging oxygens*. Bridging oxygens are those that link two tetrahedra, Fig. 1.52. Effectively, they belong half to one Si and half to another Si. In evaluating the net Si:O ratio, bridging oxygens count as $^1/_2$. Non-bridging oxygens are linked to only one Si or silicate tetrahedron as in (b). They are also called *terminal oxygens*. In order to maintain charge balance, non-bridging oxygens must of course also be linked to other cations in the crystal structure. In evaluating the overall Si:O ratio, non-bridging oxygens count as 1.

The overall Si:O ratio in a silicate structure depends on the relative number of bridging and non-bridging oxygens. Some examples are given in Table 1.27; they are all straightforward and one may deduce the type of silicate anion directly from the chemical formula.

Many more complex examples could be given. In these, although the detailed structure cannot be deduced from the formula, one can at least get an approximate idea of the type of silicate anion. For example, in $Na_2Si_3O_7$, the Si:O ratio is 1:2.33. This corresponds to a structure in which, on average, two-thirds of an O per SiO_4 is non-bridging. Clearly, therefore, some SiO_4 tetrahedra must be composed entirely of bridging oxygens whereas others contain one non-bridging oxygen. The structure of the silicate anion would therefore

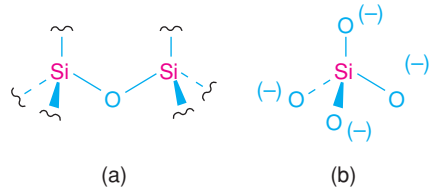


Figure 1.52 ***Silicate anions with (a) bridging and (b) non-bridging oxygens.***

Table 1.27 ***Relation between chemical formula and silicate anion structure***

Si:O ratio	Number of oxygens per Si		Type of silicate anion	Examples
	Bridging	Non-bridging		
1:4	0	4	Isolated SiO_4^{4-}	Mg_2SiO_4 olivine, Li_4SiO_4
1:3.5	1	3	Dimer $Si_2O_7^{6-}$	$Ca_3Si_2O_7$ rankinite, $Sc_2Si_2O_7$ thortveite
1:3	2	2	Chains $(SiO_3)_n^{2n-}$	Na_2SiO_3, $MgSiO_3$ pyroxene
			Rings, e.g. $Si_3O_9^{6-}$	$CaSiO_3$[a], $BaTiSi_3O_9$ benitoite
			$Si_6O_{18}^{12-}$	$Be_3Al_2Si_6O_{18}$ beryl
1:2.5	3	1	Sheets $(Si_2O_5)_n^{2n-}$	$Na_2Si_2O_5$
1:2	4	0	3D framework	SiO_2[b]

[a] $CaSiO_3$ is dimorphic. One polymorph has $Si_3O_9^{6-}$ rings and the other has infinite $(SiO_3)_n^{2n-}$ chains.
[b] The three main polymorphs of silica, quartz, tridymite and cristobalite, each have a different kind of 3D framework structure.

be expected to be something between an infinite sheet and a 3D framework. In fact, it is an infinite, double-sheet silicate anion in which two-thirds of the silicate tetrahedra have one non-bridging O.

The relation between formula and anion structure is more complex when Al is present. In some cases, Al substitutes for Si, in the tetrahedra; in others, it occupies octahedral sites. In the plagioclase feldspars typified by albite, $NaAlSi_3O_8$, and anorthite, $CaAl_2Si_2O_8$, Al partly replaces Si in the silicate anion. It is therefore appropriate to consider the overall ratio (Si + Al):O. In both cases, this ratio is 1:2 and, therefore, a 3D framework structure is expected. Framework structures also occur in orthoclase, $KAlSi_3O_8$, kalsilite, $KAlSiO_4$, eucryptite, $LiAlSiO_4$, and spodumene, $LiAlSi_2O_6$.

Substitution of Al for Si occurs in many sheet structures such as micas and clay minerals. Talc has the formula $Mg_3(OH)_2Si_4O_{10}$ and, as expected for an Si:O ratio of 1:2.5, the structure contains infinite silicate sheets. In the mica phlogopite, one-quarter of the Si in talc is effectively replaced by Al and extra K is added to preserve electroneutrality. Hence phlogopite has the formula $KMg_3(OH)_2(Si_3Al)O_{10}$. In talc and phlogopite, Mg occupies octahedral sites between silicate sheets; K occupies 12-coordinate sites.

The mica muscovite, $KAl_2(OH)_2(Si_3Al)O_{10}$, is more complex; it is structurally similar to phlogopite, with infinite sheets, $(Si_3Al)O_{10}$. However, two other Al^{3+} ions replace the three Mg^{2+} ions of phlogopite and occupy octahedral sites. By convention, only ions that replace Si in tetrahedral sites are included as part of the complex anion. Hence octahedral Al^{3+} ions are formally regarded as cations in much the same way as alkali and alkaline earth cations, Table 1.27.

With a few exceptions, silicate structures cannot be described as *cp*. However, this disadvantage is offset by the clear identification of the silicate anion component which facilitates classification and description of a very wide range of structures. In addition, the Si–O bond is strong and partially covalent and the consequent stability of the silicate anion is responsible for many of the properties of silicates.

Many examples of silicate crystal structures are given in the minerals section of the CrystalViewer Companion website. To facilitate viewing of the silicate anion, the remaining cations in the structures can be hidden; you can check out the numbers of bridging and non-bridging oxygens for consistency with the Si-O ratio in the mineral formulae.

2

Crystal Defects, Non-Stoichiometry and Solid Solutions

2.1 Perfect and Imperfect Crystals

In a perfect crystal, all the atoms are at rest on their correct lattice positions. Such a perfect crystal can be obtained, hypothetically, only at absolute zero; at all real temperatures, crystals are imperfect. Atoms vibrate, which may be regarded as a form of defect, but also a number of atoms are inevitably misplaced. In some crystals, the number of defects may be very small, $\ll 1\%$, as in, e.g., high-purity diamond or quartz. In others, high defect concentrations may be present. In highly defective crystals, the question often arises as to whether the defects themselves should be regarded as forming a fundamental part of the structure rather than as some imperfection in an otherwise ideal structure.

Crystals are invariably imperfect because the presence of defects up to a certain concentration leads to a reduction of free energy, Fig. 2.1. Let us consider the effect on the free energy of a perfect crystal of creating a single defect, say a vacant cation site. This requires a certain amount of energy, ΔH, but causes a considerable increase in entropy, ΔS, because of the large number of positions which this defect can occupy. Thus, if the crystal contains 1 mol of cations, there are $\sim 10^{23}$ possible positions for the vacancy. The entropy gained is called *configurational entropy* and is given by the Boltzmann equation:

$$S = k \ln W \tag{2.1}$$

where the probability, W, is proportional to 10^{23}; other, smaller, entropy changes are also present due to the disturbance of the crystal structure in the neighbourhood of the defect. As a result of this increase in entropy, the enthalpy required to form the defect initially is more than offset by the gain in entropy. Consequently, the free energy, given by

$$\Delta G = \Delta H - T\Delta S \tag{2.2}$$

decreases.

If we go now to the other extreme where, say, 10% of the cation sites are vacant, the change in entropy on introducing yet more defects is small because the crystal is already very disordered in terms of occupied

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

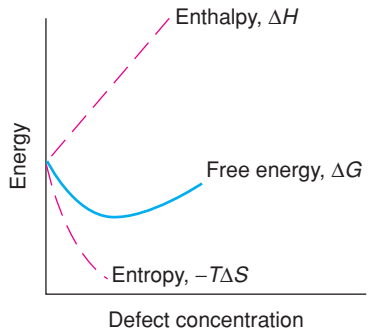


Figure 2.1 ***Energy changes on introducing defects into a perfect crystal.***

and vacant cation sites. The energy required to create more defects may be larger than any subsequent gain in entropy and hence such a high defect concentration may not be stable. In between these two extremes lie most real materials. A minimum in free energy exists which represents the number of defects present under conditions of thermodynamic equilibrium, Fig. 2.1.

Although this is a simplified explanation, it does illustrate why crystals are imperfect. It also follows that the equilibrium number of defects increases with temperature; assuming that ΔH and ΔS are independent of temperature, the $-T\Delta S$ term becomes larger and the free energy minimum is displaced to higher defect concentrations with increase in temperature.

For a given crystal, curves such as shown in Fig. 2.1 can be drawn for every possible type of defect and the main difference between them is the position of their free energy minimum. The defect that predominates is the one which is easiest to form, i.e. with the smallest ΔH and for which the free energy minimum occurs at the highest defect concentration. Thus, for NaCl it is easiest to form vacancies (the Schottky defect), whereas in AgCl the reverse is true and interstitial (Frenkel) defects predominate. In Table 2.1, the defects which predominate in a variety of inorganic solids are summarised.

2.2 Types of Defect: Point Defects

Various classification schemes have been proposed for defects, none of which is entirely satisfactory. Thus defects can be divided into two groups: *stoichiometric defects*, in which the crystal composition is unchanged on introducing the defects, and *non-stoichiometric defects*, which are a consequence of a change in composition. They are also referred to as *intrinsic defects* and *extrinsic defects*, respectively.

Table 2.1 ***Predominant point defects in various ionic crystals***

Crystal	Crystal structure	Predominant intrinsic defect
Alkali halides (not Cs)	Rock salt, NaCl	Schottky
Alkaline earth oxides	Rock salt	Schottky
AgCl, AgBr	Rock salt	Cation Frenkel
Cs halides, TlCl	CsCl	Schottky
BeO	Wurtzite, ZnS	Schottky
Alkaline earth fluorides, CeO_2, ThO_2	Fluorite, CaF_2	Anion Frenkel

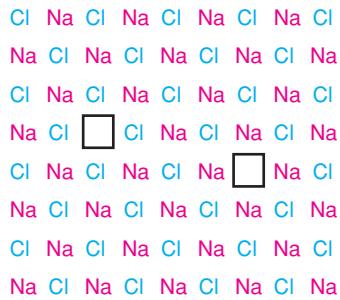


Figure 2.2 ***2D representation of a Schottky defect with cation and anion vacancies.***

Alternatively, the size and shape of the defect can be used for classification: *point defects* involve only one atom or site, e.g. vacancies or interstitials, although the atoms immediately surrounding the defect are also somewhat perturbed; *line defects*, i.e. *dislocations*, are effectively point defects in two dimensions but in the third dimension the defect is very extensive or infinite; in *plane defects*, whole layers in a crystal structure can be defective. Sometimes the name *extended defects* is used to include all those which are not point defects.

It is convenient here to begin with point defects. These are the 'classical' defects, proposed in the 1930s following the work of Schottky, Frenkel, Wagner and others; it was several decades before direct experimental evidence for their existence was forthcoming.

2.2.1 Schottky defect

In ionic solids such as halides or oxides, the Schottky defect, a stoichiometric defect, is a pair of vacant sites, an anion vacancy and a cation vacancy. To compensate for the vacancies, there should be two extra atoms at the surface of the crystal for each Schottky defect. The Schottky defect is the principal point defect in the alkali halides and is shown for NaCl in Fig. 2.2. There must be equal numbers of anion and cation vacancies in order to preserve local electroneutrality.

The vacancies may be distributed at random in the crystal or may associate into pairs or larger clusters. They tend to associate because they carry an effective charge, and oppositely charged vacancies attract each other. Thus an anion vacancy in NaCl has a net positive charge of +1 because the vacancy is surrounded by six Na^+ ions, each with partially unsatisfied positive charge. Put another way, the anion vacancy has charge +1 because, on placing an anion of charge –1 in the vacancy, local electroneutrality is restored. Similarly, the cation vacancy has a net charge of –1. In order to dissociate vacancy pairs, energy equivalent to the enthalpy of association, 1.30 eV for NaCl ($\sim$120 kJ mol^{-1}) must be provided.

The number of Schottky defects in a crystal of NaCl is either very small or very large, depending on one's point of view. At room temperature, typically one in 10^{15} of the possible anion and cation sites is vacant, an insignificant number in terms of the average crystal structure of NaCl as determined by X-ray diffraction. On the other hand, a grain of salt weighing 1 mg (and containing $\sim 10^{19}$ atoms) contains $\sim 10^4$ Schottky defects, hardly an insignificant number! The presence of defects, even in small concentrations, often influences properties. Thus, Schottky defects are responsible for the optical and electrical properties of NaCl.

2.2.2 Frenkel defect

This stoichiometric defect involves an atom displaced off its lattice site into an interstitial site that is normally empty. AgCl (which also has the NaCl crystal structure) has predominantly this defect, with Ag

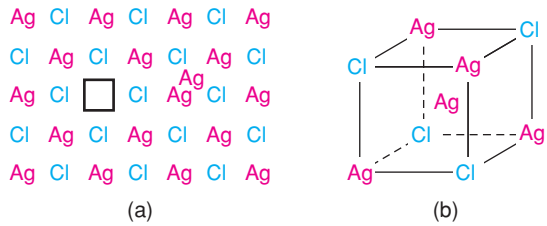


Figure 2.3 ***(a) 2D representation of a Frenkel defect in AgCl; (b) interstitial site showing tetrahedral coordination by both Ag and Cl.***

as the interstitial atom, Fig. 2.3. The nature of the interstitial site is shown in Fig. 2.3(b). It is surrounded tetrahedrally by four Cl^- ions but also, and at the same distance, by four Ag^+ ions. The interstitial Ag^+ ion is in an eight-coordinate site, and therefore with four Ag^+ and four Cl^- nearest neighbours. There is probably some covalent interaction between the interstitial Ag^+ ion and its four Cl^- neighbours which acts to stabilise the defect and give Frenkel defects, in preference to Schottky defects, in AgCl. On the other hand, Na^+, with its 'harder', more cationic character, would not find much comfort in a site which was tetrahedrally surrounded by four other Na^+ ions. Frenkel defects therefore do not occur to any significant extent in NaCl.

Calcium fluoride, CaF_2, has predominantly *anion* Frenkel defects in which F^- occupies interstitial sites. These interstitial sites (empty cubes) can be seen in Fig. 1.34. Other materials with fluorite and antifluorite structures have similar defects, e.g. Na_2O (Na^+ interstitial).

As with Schottky defects, the vacancy and interstitial are oppositely charged and may attract each other to form a pair. These pairs are electrically neutral but are dipolar and attract other dipoles to form larger aggregates or clusters. Clusters similar to these may act as nuclei for the precipitation of phases of different composition in non-stoichiometric crystals.

2.2.2.1 The Kroger–Vink notation for crystal defects

In discussing point defects, the sites involved and the effective charges on the defects, the notation developed by Kroger and Vink is extremely useful. Each defect is represented by a combination of three symbols:

- the symbol for the atom involved, or *V* if the site is vacant;
- a superscript indicating the net charge:
  - $\bullet$ for charge $+1$,
  - x for zero net charge,
  - $'$ for charge -1;
- a subscript to indicate the nature of the site in an ideal crystal, with i for an interstitial site and s for a surface site.

Some examples are as follows:

V'_{Na}	a sodium ion vacancy (in e.g. NaCl), net charge -1
$V^{\bullet}_{Cl}$	a chloride ion vacancy (in NaCl), net charge $+1$
Na^{x}_{Na}, Cl^{x}_{Cl}	Na and Cl ions on regular lattice sites, no net charge
$Cd^{\bullet}_{Na}$	a substitutional Cd ion on a Na site, charge $+1$
$Ag^{\bullet}_{i}$	an interstitial Ag^+ ion in AgCl, charge $+1$
F'_{i}	an interstitial F^- ion in CaF_2, charge -1.

2.2.2.2 *Thermodynamics of Schottky and Frenkel defect formation*

Schottky and Frenkel defects are *intrinsic defects*, i.e. they are present in pure material and a certain minimum number must be present from thermodynamic considerations. It is very common for crystals to have more defects than the thermodynamic equilibrium concentration. Thus, crystals are usually prepared at high temperatures where, intrinsically, more defects are present due to the increasing importance of the $T\Delta S$ term, Fig. 2.1. On cooling, some defects may be eliminated, but in general, and unless the cooling rate is extremely slow, the defects present are in excess of the equilibrium concentration.

An excess of defects may also be generated, deliberately, by bombarding a crystal with high-energy radiation. Atoms may be knocked out of their normal lattice sites but the reverse reaction, involving elimination or recombination of defects, usually takes place only slowly.

Two approaches are used to study point defect equilibria. Using statistical thermodynamics, the complete partition function for a model of the defective crystal is constructed; the free energy is expressed in terms of the partition function and then minimised to obtain the equilibrium condition. This method can also be applied to non-stoichiometric equilibria. Alternatively, the law of mass action can be applied and the concentration of defects expressed as an exponential function of temperature. Because of its simplicity and ease of application, the latter method is used here. Most research has been carried out on pure and doped alkali halides for which it has been relatively easy to grow large single crystals. The results obtained have greatly helped our understanding of defect equilibria and its application to ceramic oxide materials.

Schottky defects

Schottky equilibria in a crystal of NaCl can be treated by assuming that a pair of ions is removed from the crystal interior, leaving vacancies, and relocated at the crystal surface, as follows:

$$\mathrm{Na_{Na}} + \mathrm{Cl_{Cl}} + V_{\mathrm{Na},S} + V_{\mathrm{Cl,S}} \rightleftharpoons V'_{\mathrm{Na}} + V^{\bullet}_{\mathrm{Cl}} + \mathrm{Na_{Na,S}} + \mathrm{Cl_{Cl,S}} \tag{2.3}$$

where V'_{Na}, $V^{\bullet}_{\mathrm{Cl}}$, $V_{\mathrm{Na,S}}$, $V_{\mathrm{Cl,S}}$, $\mathrm{Na_{Na}}$, $\mathrm{Cl_{Cl}}$, $\mathrm{Na_{Na,S}}$ and $\mathrm{Cl_{Cl,S}}$ represent cation and anion vacancies, vacant cation and anion surface sites, occupied cation and anion sites and occupied cation and anion surface sites, respectively. The equilibrium constant for formation of Schottky defects is given by

$$K = \frac{[V'_{\mathrm{Na}}][V^{\bullet}_{\mathrm{Cl}}][\mathrm{Na_{Na,S}}][\mathrm{Cl_{Cl,S}}]}{[\mathrm{Na_{Na}}][\mathrm{Cl_{Cl}}][V_{\mathrm{Na,S}}][V_{\mathrm{Cl,S}}]} \tag{2.4}$$

where square brackets represent concentrations of the species involved. The number of surface sites is constant in a crystal of constant total surface area. Hence the number of Na^+ and Cl^- ions that occupy surface sites is also constant. On formation of Schottky defects, Na^+ and Cl^- ions move to surface sites but, at the same time, an equal number of fresh surface sites is created. (The total surface area of a crystal must increase slightly as Schottky defects are created, but this effect can be ignored.) Therefore, $[\mathrm{Na_{Na,s}}] = [V_{\mathrm{Na,s}}]$ and $[\mathrm{Cl_{Cl,S}}] = [V_{\mathrm{Cl,S}}]$, so equation (2.4) reduces to

$$K = \frac{[V'_{\mathrm{Na}}][V^{\bullet}_{\mathrm{Cl}}]}{[\mathrm{Na_{Na}}][\mathrm{Cl_{Cl}}]} \tag{2.5}$$

Let N be the total number of sites of each kind. N_V is the number of vacancies of each kind and hence the number of Schottky defects. The number of occupied sites of each kind is $N - N_V$. Substituting into equation (2.5) gives

$$K = \frac{(N_V)^2}{(N - N_V)^2} \tag{2.6}$$

For small concentrations of defects:

$$N \simeq N - N_V \tag{2.7}$$

and therefore

$$N_V \simeq N\sqrt{K} \tag{2.8}$$

The equilibrium constant, K, is an exponential function of temperature:

$$K \propto \exp(-\Delta G/RT) \tag{2.9}$$

$$\propto \exp(-\Delta H/RT)\exp(\Delta S/R) \tag{2.10}$$

$$= \text{constant} \times \exp(-\Delta H/RT) \tag{2.11}$$

Therefore,

$$N_V = N \times \text{constant} \times \exp(-\Delta H/2RT) \tag{2.12}$$

where ΔG, ΔH and ΔS are the free energy, enthalpy and entropy of formation of 1 mol of defects. This shows that the equilibrium concentration of Schottky defects increases exponentially with temperature. For a given temperature, the smaller the enthalpy of formation, H, the higher is the Schottky defect concentration.

A similar expression can be derived for the concentration of vacancies in a monoatomic crystal of, e.g., a metal. The difference is that, because only one type of vacancy is present, equations (2.5)–(2.7) simplify to give

$$N_V = NK \tag{2.13}$$

The factor of 2 therefore drops out of the exponential term in equation (2.12).

Frenkel defects

The Frenkel equilibria in AgCl involve the displacement of Ag^+ ions from their lattice sites into empty interstitial sites and can be represented by

$$Ag_{Ag} + V_i \rightleftharpoons Ag_i^{\bullet} + V'_{Ag} \tag{2.14}$$

where V_i and $Ag_i^\bullet$ represent empty and occupied interstitial sites, respectively. The equilibrium constant is given by

$$K = \frac{[Ag_i^\bullet][V'_{Ag}]}{[Ag_{Ag}][V_i]} \tag{2.15}$$

Let N be the number of lattice sites in a perfect crystal and N_i the number of occupied interstitial sites, i.e.

$$[V'_{Ag}] = [Ag_i^\bullet] = N_i \tag{2.16}$$

and

$$[Ag_{Ag}] = N - N_i \tag{2.17}$$

For most structures,

$$[V_i] = \alpha N \tag{2.18}$$

i.e. the number of available interstitial sites is simply related to the number of lattice sites. For AgCl, $\alpha = 2$ because there are two tetrahedral interstitial sites for every octahedral site that is occupied by Ag^+ (in the *ccp* rock salt structure of AgCl, there are twice as many tetrahedral sites as octahedral sites). Substituting into equation (2.15) gives

$$K = \frac{N_i^2}{(N - N_i)(\alpha N)} \simeq \frac{N_i^2}{\alpha N^2} \tag{2.19}$$

Combining equations (2.9) and (2.19) gives

$$[V_{Ag}] = [Ag_i^+] = N_i = N\sqrt{\alpha}\exp(-\Delta G/2RT) \tag{2.20}$$

$$= \text{constant} \times N\exp(-\Delta H/2RT) \tag{2.21}$$

For both Frenkel and Schottky defects, a factor of 2 appears in the exponential expression [equations (2.21) and (2.12)], because there are two defective sites per defect, i.e. two vacancies per Schottky defect and one vacancy, one interstitial site per Frenkel defect. In each case, therefore, the overall ΔH for defect formation can be regarded as the sum of two component enthalpies.

Experimental values for the number of Frenkel defects in AgCl are given in Fig. 2.4. The method of presentation is based on the Arrhenius (or Boltzmann) equation, i.e. taking logarithms in equation (2.21) gives

$$\log_{10}(N_i/N) = \log_{10}(\text{constant}) - (\Delta H/2RT)\log_{10}\mathrm{e} \tag{2.22}$$

A plot of $\log_{10}(N_i/N)$ against $1/T$ should be a straight line of slope $-(\Delta H \log_{10}\mathrm{e})/2R$. The experimental data for AgCl fit the Arrhenius equation reasonably well, although there is a slight but definite upward departure from linearity at high temperatures. From Fig. 2.4, and ignoring for the moment the slight curvature at high temperatures, the numbers of vacancies and interstitials increase rapidly with increase in temperature: just below the melting point of AgCl, 456 °C, the equilibrium concentration of Frenkel defects is estimated as

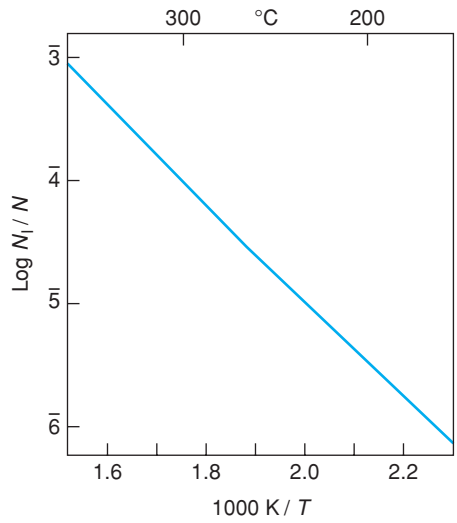


Figure 2.4 ***Fraction of Frenkel defects in AgCl as a function of temperature. Reproduced with permission from Abbink and Martin,* J. Phys. Chem. Solids*, 27, 205, © 1966 Elsevier.***

0.6%, i.e. approximately 1 in 200 of the Ag^+ move off their octahedral lattice sites to occupy interstitial tetrahedral sites. This defect concentration is one to two orders of magnitude higher than for most ionic crystals, with either Frenkel or Schottky disorder, just below their melting points. The enthalpy of formation of Frenkel defects in AgCl is ~1.35 eV (130 kJ mol^{-1}) and of Schottky defects in NaCl is 2.3 eV (~ 220 kJ mol^{-1}); these values are fairly typical for ionic crystals.

The departure from linearity at high temperatures in Fig. 2.4 is attributed to the existence of a long-range Debye–Hückel attractive force between defects of opposite charge, e.g. between vacancies and interstitials in AgCl. This attraction reduces somewhat the energy of formation of the defects and so the number of defects increases, especially at high temperatures.

2.2.3 Colour centres

The best known example of a colour centre is the F-centre (from the German *Farbenzentre*), shown in Fig. 2.5; it is an electron trapped on an anion vacancy. F-centres can be prepared by heating an alkali halide in vapour of an alkali metal. NaCl heated in Na vapour becomes slightly non-stoichiometric due to the uptake of Na to give $Na_{1+\delta}Cl$: $\delta \ll 1$, which has a greenish yellow colour. The process must involve the absorption of Na atoms, which ionise on the crystal surface. The resulting Na^+ ions stay at the surface but the ionised electrons

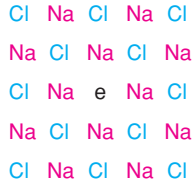


Figure 2.5 ***The F-centre, an electron trapped on an anion vacancy.***

Figure 2.6 ***(a) H-centre and (b) V-centre in NaCl.***

diffuse into the crystal where they encounter and occupy vacant anion sites. To preserve charge balance, an equal number of Cl^- ions must find their way out to the surface. The trapped electron provides a classic example of an 'electron in a box'. A series of energy levels are available for the electron within this box and the energy required to transfer from one level to another falls in the visible part of the electromagnetic spectrum; hence the colour of the F-centre. The magnitude of the energy levels and the colour observed depend on the host crystal and not on the source of the electron. Thus, NaCl heated in K vapour has the same yellowish colour as NaCl heated in Na vapour, whereas KCl heated in K vapour is violet.

Another means of producing F-centres is by irradiation. Using one of the normal methods of recording an X-ray diffraction (XRD) pattern (see Chapter 5), powdered NaCl turns a greenish yellow colour after bombardment with X-rays. The cause of the colour is again trapped electrons, but in this case they cannot arise from a non-stoichiometric excess of Na. They probably arise from ionisation of some Cl^- ions.

The F-centre is a single trapped electron which has an unpaired spin and, therefore, an electron paramagnetic moment. A powerful method for studying such colour centres is electron spin resonance (ESR) spectroscopy (see Chapter 6), which detects unpaired electrons.

Many other colour centres have been characterised in alkali halide crystals; two of these, the H-centre and V-centre, are shown in Fig. 2.6. Both contain the chloride molecule ion, Cl_2^-, but this occupies one site in the H-centre and two sites in the V-centre; in both cases the axis of the Cl_2^- ion is parallel to a $<101>$ direction. The V-centre occurs on irradiation of NaCl with X-rays. The mechanism of formation presumably involves ionisation of a Cl^- ion to give a neutral Cl atom, which then covalently bonds with a neighbouring Cl^- ion. Other defect centres which have been identified in the alkali halides include:

(a) the F′-centre, which is two electrons trapped on an anion vacancy;
(b) the F_A-centre, which is an F-centre, one of whose six cationic neighbours is a foreign monovalent cation, e.g. K^+ in NaCl;
(c) the M-centre, which is a pair of nearest neighbour F-centres;
(d) the R-centre, which is three nearest neighbour F-centres located on a (111) plane;
(e) ionised or charged cluster centres, such as M^+, R^+ and R^-.

2.2.4 Vacancies and interstitials in non-stoichiometric crystals: extrinsic and intrinsic defects

Schottky and Frenkel defects are intrinsic defects which occur in stoichiometric crystals. Defects can also occur as a consequence of variation in composition and these are then extrinsic defects. Extrinsic defects may arise on doping pure crystals with *aliovalent impurities*, i.e. impurity atoms which have a different valency to those in the host crystal. For instance, NaCl may be doped with $CaCl_2$ to give crystals of formula $Na_{1-2x}Ca_xV_{Na,x}Cl$. In these, the *ccp* Cl^- arrangement is retained but the Na^+, Ca^{2+} ions and V_{Na} are distributed over the octahedral cation sites. The overall effect of doping NaCl with Ca^{2+} ions is to increase the number

of cation vacancies. Those vacancies that are controlled by the impurity level are extrinsic defects, in contrast to the thermally created intrinsic defects such as Schottky pairs.

For crystals with dilute ($\ll 1\%$) defect concentrations, the law of mass action can be applied. From equation (2.5), the equilibrium constant K for the formation of Schottky defects is proportional to the product of anion and cation vacancy concentrations, i.e.

$$K \propto [V_{Na}][V_{Cl}]$$

It is assumed that the addition of small amounts of impurities such as Ca^{2+} does not affect the value of K. Since the cation vacancy concentration increases with increase in $[Ca^{2+}]$ and K is constant, the anion vacancy concentration must decrease correspondingly.

The practice of doping crystals with aliovalent impurities, coupled with the study of mass transport or electrical conductivity, is a powerful method for studying point defect equilibria. Mass transport in NaCl occurs by migration of vacancies. In reality, an ion adjacent to a vacancy moves into that vacancy, leaving its own site vacant, but this process can effectively be regarded as vacancy migration. From an analysis of the dependence of conductivity on temperature and defect concentration, thermodynamic parameters such as the enthalpies of creation and migration of defects may be determined. Further details are given in Chapter 8.

2.2.5 Defect clusters or aggregates

As defects are studied in more detail, using high-resolution electron microscopy and other techniques, coupled with computer-assisted modelling of defect structures, it is clear that the apparently simple point defects such as vacancies and interstitials are often more complex; instead, larger defect clusters tend to form. Take the example of an interstitial metal atom in an *fcc* metal. If the assumption is made that defects, such as interstitials, do not perturb the host structure then there are two possible sites for the interstitial atom, tetrahedral and octahedral. Recent research shows, however, that interstitial atoms *do* perturb the host structure in the immediate vicinity of the interstitial atom. An example is shown in Fig. 2.7 for Pt metal containing an interstitial Pt atom. Instead of the interstitial Pt occupying the octahedral site, it is displaced by about 1 Å off the centre of this site and in the direction of one of the face centre atoms. The Pt on this face centre also suffers a corresponding displacement in the same [100] direction. Thus, the defect involves *two* atoms, both of which are on distorted interstitial sites. This defect complex is known as a *split interstitial* or *dumb-bell-shaped interstitial*.

A similar split interstitial defect is present in *bcc* metals such as α-Fe. The 'ideal' site for the interstitial would be in the centre of a cube face, but instead it is displaced off the centre of this site in the direction of one of the corners. At the same time, an adjacent corner atom is displaced in the same [110] direction, i.e. along a face diagonal, Fig. 2.8.

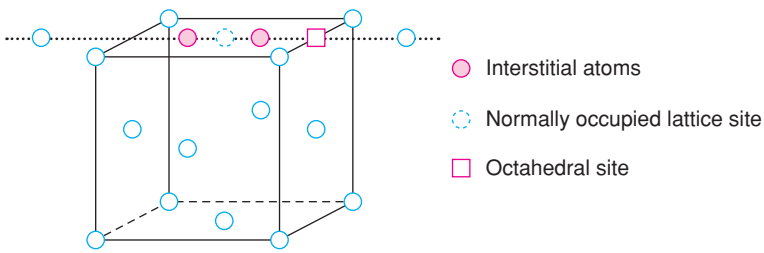


Figure 2.7 ***Split interstitial defect in an fcc metal.***

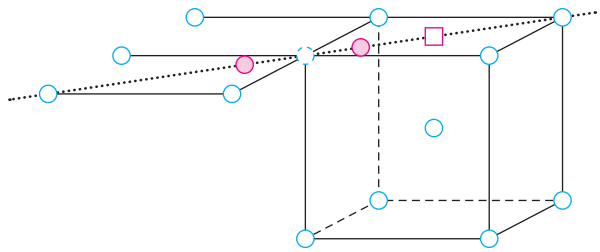


Figure 2.8 ***Split interstitial in a* bcc *metal, e.g. α-Fe. Symbols as in Fig. 2.7.***

The exact structure of interstitial defects in alkali halides is not clear. Although Schottky defects predominate, interstitials are also present but in much smaller quantities. Calculations indicate that, in some materials, occupation of undistorted interstitial sites is favoured whereas in others a split interstitial is preferred.

The presence of vacancies in both metals and ionic crystals causes a relaxation of the structure in the immediate environment of the vacancy. In metals, the atoms surrounding the vacancy relax *inwards* by a few percent, i.e. the vacancy becomes smaller, whereas in ionic crystals the reverse occurs and as a result of an imbalance in electrostatic forces the atoms relax *outwards*.

Vacancies in ionic crystals are charged and, therefore, vacancies of opposite charge can attract each other to form clusters. The smallest cluster is either an anion vacancy/cation vacancy pair or an aliovalent impurity (e.g. Cd^{2+})/cation vacancy pair. These pairs are dipolar, although overall electrically neutral, and so can attract other pairs to form larger clusters.

One of the most studied and best understood defect systems is wüstite, $Fe_{1-x}O$: $0 \leq x \leq 0.1$. Stoichiometric FeO has the rock salt structure with Fe^{2+} ions on octahedral sites. Density measurements show that non-stoichiometric $Fe_{1-x}O$ contains a deficiency of iron rather than an excess of oxygen, relative to stoichiometric FeO. Using point defect considerations one would anticipate that non-stoichiometric $Fe_{1-x}O$ would have the structural formula $Fe^{2+}_{1-3x}Fe^{3+}_{2x}V_xO$ in which Fe^{2+}, Fe^{3+} and cation vacancies were distributed at random over the octahedral sites in the *ccp* oxide ion array. The defect structure is different, however, since neutron and X-ray diffraction studies have shown that Fe^{3+} ions are in *tetrahedral* sites and that defect clusters appear to form.

The structure of one possibility, the so-called Koch cluster, is shown in Fig. 2.9. The oxygen ions are *ccp* throughout and the cluster involves all the cation sites in a cube the size of the normal *fcc* rock salt unit cell. The 12 edge centre and one body centre octahedral sites are all empty; four of the eight possible tetrahedral sites contain Fe^{3+} ions (dividing the cube into eight smaller cubes, these tetrahedral sites occupy the body centres of the small cubes). This cluster has a net charge of 14^- since there are 13 vacant M^{2+} sites (26^-) and only four interstitial Fe^{3+} ions (12^+). Extra Fe^{3+} ions are distributed over octahedral sites around the clusters to preserve electroneutrality. The concentration of these Koch clusters increases with increase in x and hence the average separation between clusters decreases. Evidence from diffuse neutron scattering indicates that the clusters order into a regularly repeating pattern resulting in a superstructure for wüstite.

Another well-studied defect system is oxygen-rich uranium dioxide. XRD is virtually useless for studying this material because information about the O positions is lost in the heavy scattering from U. Instead, neutron diffraction has been used. Stoichiometric UO_2 has the fluorite structure and non-stoichiometric UO_{2+x} contains interstitial oxide ions, which form part of the cluster shown in Fig. 2.10. In UO_2, the interstitial positions are present at the centres of alternate cubes containing oxygen ions at the corners. In UO_{2+x}, an interstitial oxygen is displaced off a cube centre interstitial position in the [110] direction, i.e. towards one of the cube edges. At the same time, the two nearest corner oxygens are displaced into adjacent empty cubes along [111]

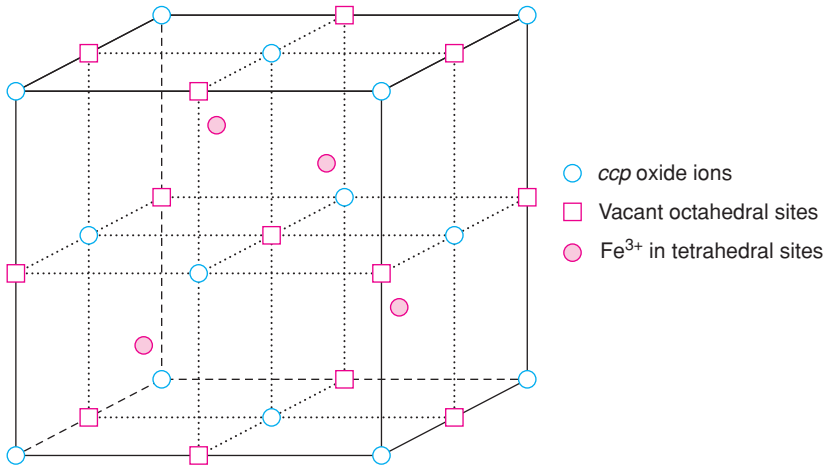


Figure 2.9 ***Koch cluster postulated to exist in wüstite, $Fe_{1-x}O$.***

directions, Fig. 2.10. Thus, in place of a single interstitial atom, the cluster contains three interstitial oxygens and two vacancies.

At this point, it is worth emphasising why it is difficult to determine the exact structures of defects in crystals. Diffraction methods (X-ray, neutron, electron), as they are normally used, yield *average* structures for crystals. For crystals that are pure and relatively free from defects, the average structure is a close representation of the true structure. For non-stoichiometric and defective crystals, however, the average structure may give a

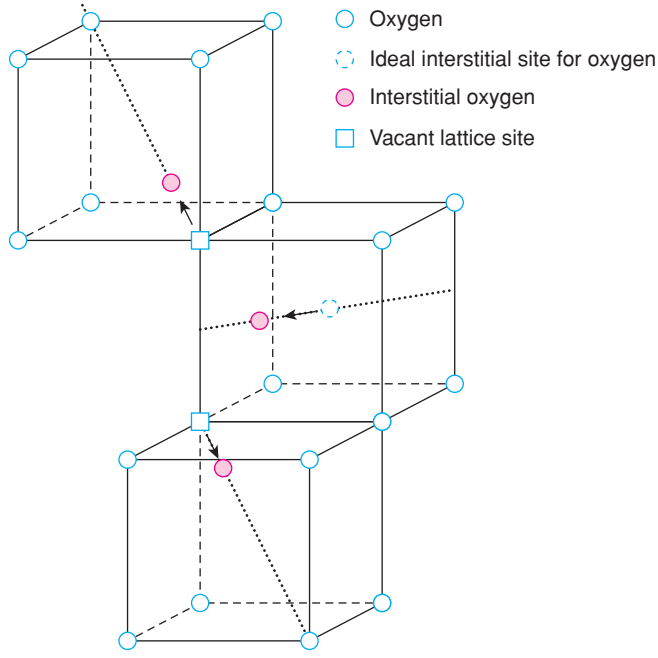


Figure 2.10 ***Interstitial defect cluster in UO_{2+x}. Uranium positions (not shown) are in the centre of every alternate cube.***

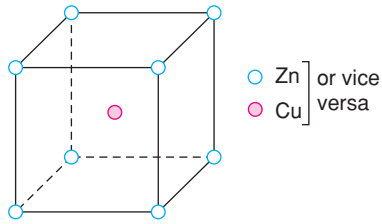


Figure 2.11 ***Ordered, primitive cubic unit cell of β′-brass, CuZn.***

very poor and even an incorrect representation of the actual structure in the region of the defect. In order to determine defect structures, techniques are needed that are sensitive to local structure. Study of local structure is a principal application of spectroscopic techniques (see Chapter 6); the sites occupied by impurity atoms can be determined if the impurity is spectroscopically active. However, the spectra are not usually informative about structure extending beyond the immediate coordination environment of a particular atom. What is needed, therefore, are techniques that give information on local structure extending over a distance of, say, 5–20 Å. Such techniques are very scarce, although there is hope that improved EXAFS methods and in particular pair distribution functions, PDF, may be suitable.

2.2.6 Interchanged atoms: order–disorder phenomena

In certain crystalline materials, pairs of atoms or ions may swap places. This frequently occurs in alloys that contain two or more different elements, each arranged on a specific set of sites. It also occurs in certain ionic structures that contain two or more types of cation, with again, each being located on a specific set of sites. If the number of interchanged atoms is large, and especially if it increases significantly with increase in temperature, this takes us into the realm of *order–disorder phenomena*. The limit is reached when sufficient pairs have swapped places that atoms no longer show preference for particular sites. The structure is then *disordered* as regards the particular sets of atoms involved.

Alloys, by their very nature, involve a distribution of atoms of two or more different metals over one (sometimes more) set of crystallographic sites. Such alloys are examples of *substitutional solid solutions*. Alloys can either be disordered, with the atoms distributed at random over the available sites, or ordered, with the different atoms occupying distinct sets of sites. Ordering is generally accompanied by the formation of a *supercell* which is detected by XRD from the presence of extra reflections. The ordered superstructure present in β'-brass, CuZn, below ~450 °C is shown in Fig. 2.11. Cu atoms occupy the body centre positions in a cube with Zn atoms at the corners, as in the CsCl structure; the lattice type is therefore primitive. In the disordered alloy of the same composition, β-brass, the Cu and Zn atoms are distributed at random over the corner and body centre positions and therefore the lattice is body centred, as in the structure of α-Fe. The main point of interest to this chapter is that it is possible, within the ordered, β'-brass structure, to mix up *some* of the Cu and Zn atoms while still retaining the long-range order and superstructure. This disorder can be regarded as the introduction of defects into an otherwise perfect structure.

Good examples of interchanged cations in non-metallic oxide structures occur with the spinel structure. More details are given in Chapters 1 and 9.

2.3 Solid Solutions

Thus far, we have seen how defects can be *intrinsic* if they are associated with stoichiometric crystals or *extrinsic* if they are associated with dopants or impurities. In the latter case, the dopants either occupy

interstitial sites or substitute for atoms or ions in the parent lattice. As the dopant concentration rises, above ~0.1–1%, it is common practice to refer to the materials as *solid solutions* rather than as doped materials, but these two terminologies are interchangeable.

A solid solution is basically a crystalline phase that can have variable composition. As with doped crystals, simple solid solutions are one of two types: in *substitutional* solid solutions, the atom or ion that is being introduced directly replaces an atom or ion of the same charge in the parent structure; in *interstitial* solid solutions, the introduced species occupies a site that is normally empty and no ions or atoms are left out. Starting with these two basic types, a considerable variety of more complex solid solution mechanisms may be derived, by having both substitution and interstitial formation occurring together, by introducing ions of different charge to those in the host structure or by creating vacancies. Let us first look in some detail at simple substitutional and interstitial solid solutions.

2.3.1 Substitutional solid solutions

An example of a substitutional solid solution is the series of oxides formed on reacting together Al_2O_3 and Cr_2O_3 at high temperatures. Both end members have the corundum structure (approximately *hcp* oxide ions with Al^{3+}, Cr^{3+} ions occupying two-thirds of the available octahedral sites) and the solid solution may be formulated as $(Al_{2-x}Cr_x)O_3$: $0 \leq x \leq 2$. At intermediate x, Al^{3+} and Cr^{3+} ions are distributed at random over those octahedral sites that are occupied by Al^{3+} in Al_2O_3. Thus, while any particular site must contain either Cr^{3+} or Al^{3+}, the probability that it is one or the other is related to the composition x. When the structure is considered as a whole and the occupancy of all the sites is averaged, it is useful to think of each site as being occupied by an 'average cation' whose properties, atomic number, size, etc., are intermediate between those of Al^{3+} and Cr^{3+}.

In addition to providing a simple example of a substitutional solid solution, the Al_2O_3–Cr_2O_3 system provides a dramatic example of how properties may change on doping or solid solution formation. Al_2O_3 is a white, insulating solid. With small amounts (<1%) of Cr^{3+} substituting for Al^{3+}, the colour becomes bright red. This is the composition of the ruby gemstone and was the material in which the principle of laser action was first demonstrated. What a dramatic effect such a small amount of impurity has! This is compounded by the fact that the ruby red colour (and laser action) is lost if too much Cr^{3+} is added; the solid solutions then revert to the green colour of Cr_2O_3.

For a range of simple substitutional solid solutions to form, there are certain requirements that must be met. First, the ions that replace each other must have the same charge. If this were not the case, then other structural changes, involving vacancies or interstitials, would be required to maintain electroneutrality; such cases are discussed later.

Second, the ions that replace each other must be fairly similar in size. From a review of experimental results on metal alloy formation, it has been suggested that a difference of 15% in the radii of the metal atoms that replace each other is the most that can be tolerated if a substantial range of substitutional solid solutions is to form. For solid solutions in non-metallic systems, the limiting acceptable size difference appears to be larger than 15%, although it is difficult to quantify this, partly because it is difficult to quantify the sizes of the ions (see Chapter 3 for a discussion of ionic radii) and partly because solid solution formation is often very temperature dependent. Thus, extensive solid solutions may form at high temperatures whereas at lower temperatures, these may be more restricted or practically non-existent.

There is a simple thermodynamic explanation for an increase in solid solution formation with increase in temperature. For a solid solution to be stable, it must have lower free energy than the corresponding mixture of phases without solid solution formation. Free energy may be separated into enthalpy and entropy components according to equation (2.2); let us consider the effect on each of solid solution formation. The entropy term always favours solid solution formation because the entropy associated with two types of cation distributed

at random over a set of sites in the crystal structure is much greater than the entropy of the corresponding mixture of two stoichiometric end-member phases. The enthalpy term may, or may not, favour solid solution formation. If the enthalpy of formation is negative, then solid solution formation is favoured by both enthalpy and entropy terms and the solid solutions are stable at all temperatures. If the enthalpy of formation is positive, then the enthalpy and entropy effects are opposed. The enthalpy dominates the free energy at low temperatures and extensive solid solutions do not form. With increase in temperature, the $T\Delta S$ term increases in magnitude until such a temperature that it more than compensates for the positive enthalpy. Consequently, solid solutions are favoured by high temperatures.

Let us now see examples of ions which can and cannot replace each other in solid solution formation. We need a consistent set of radii in order to make comparisons; using Shannon and Prewitt radii for octahedral coordination, based on an oxide ion radius of 1.26 Å, monovalent alkali metal and silver ions have radii: Li 0.88, Na 1.16, Ag 1.29, K 1.52, Rb 1.63 and Cs 1.84 Å. The radii of K/Rb and Rb/Cs pairs are both within 15% of each other and it is common to obtain solid solutions between say, pairs of corresponding Rb/Cs salts. Na/K salts sometimes form solid solutions with each other, especially at high temperatures (e.g. KCl and NaCl at 600 °C) and the K^+ ion is ∼30% larger than Na^+. The difference in size of Li^+ and K^+ is too large, however, and these ions do not generally replace each other in solid solutions. The Ag^+ ion is similar in size to Na^+ and solid solutions between pairs of corresponding Na/Ag salts are common.

Some divalent ions, with octahedral radii, are: Mg 0.86, Ca 1.14, Sr 1.30, Ba 1.50, Mn 0.96 (high spin), Fe 0.91 (high spin), Co 0.88 (high spin), Ni 0.84, Cu 0.87, Zn 0.89 and Cd 1.09 Å. It is common to find solid solutions in which the divalent transition metal ions, Mn . . . Zn, substitute for each other and we can see that their radii are not too different. Mg also commonly forms solid solutions with them and is of similar size; Ca does not, however, because it is 20–30% larger. Similar considerations apply to trivalent ions; similar-sized Al, Ga, Fe and Cr (0.67–0.76 Å) commonly substitute for each other, as do many of the trivalent lanthanides (0.99–1.20 Å).

In summary, ions of similar size (e.g. Zr 0.86, Hf 0.85 Å) substitute for each other easily and extensive solid solutions form which are stable at all temperatures; the enthalpy of mixing of such similar-sized ions is likely to be small and the driving force for solid solution formation is the increased entropy. With ions that differ in size by 15–20%, solid solutions may form at high temperatures where the entropy term is able to offset the positive enthalpy term. With ions that differ in size by more than ∼30%, solid solutions are not expected to form.

In considering whether solid solutions form, an important factor is the crystal structure of the two end members. In systems that exhibit complete solid solution, it is essential that the end members be isostructural. The reverse is not necessarily true, however, and just because two phases are isostructural it does not follow that they will form solid solutions with each other; e.g. LiF and CaO both have the rock salt structure but they do not form solid solutions.

While complete ranges of solid solution form in favourable cases, as between Al_2O_3 and Cr_2O_3, it is far more common to have partial or limited ranges of solid solution. In such cases, the end members need not be isostructural. For example, Mg_2SiO_4 (forsterite, an olivine) and Zn_2SiO_4 (willemite) can dissolve about 20% of each other in solid solution formation, even though their crystal structures are different. In the forsterite solid solutions, of formula

$$Mg_{2-x}Zn_xSiO_4$$

some of the octahedrally coordinated Mg^{2+} ions are replaced by Zn^{2+}. In the willemite solid solutions, on the other hand, of formula

$$Zn_{2-x}Mg_xSiO_4$$

Mg^{2+} ions substitute partially for the tetrahedrally coordinated Zn^{2+} ions. Such solid solutions are possible because Mg and Zn are of similar size and are happy in both tetrahedral and octahedral coordination.

Some ions, especially transition metal ions, have strong preference for a particular site symmetry or coordination. Thus, Cr^{3+} is almost always found in octahedral sites, whereas similar-sized Al^{3+} can occupy either tetrahedral or octahedral sites. Such site preferences preclude the formation of certain solid solutions. For example, $LiCrO_2$ contains octahedral Cr^{3+} and forms solid solutions of formula

$$LiCr_{1-x}Al_xO_2$$

in which Al partially replaces Cr on the octahedral sites. The converse solid solution does not form, however, since in $LiAlO_2$ the coordination of Al is tetrahedral and Cr has a strong dislike for tetrahedral coordination.

Silicates and germanates are often isostructural and form solid solutions by $Si^{4+} \rightleftharpoons Ge^{4+}$ replacement. The lanthanides, because of their similarity in size, are notoriously good at forming solid solutions in, say, their oxides. Indeed, the great difficulty experienced by early chemists in separating the lanthanides was due to this very easy solid solution formation.

Anions may also replace each other in substitutional solid solutions, e.g. AgCl–AgBr solid solutions, but these are not nearly as common as those formed by cation substitution, probably because there are not many pairs of anions with similar size and coordination/bonding requirements. Many alloys are nothing more than substitutional solid solutions, e.g. in brass, Cu and Zn replace each other over a wide range of compositions, with general formula $Cu_{1-x}Zn_x$.

2.3.2 Interstitial solid solutions

Many metals form interstitial solid solutions in which small atoms, especially H, C, B and N, enter empty interstitial sites within the host metal structure. Pd metal is well known for its ability to 'occlude' enormous volumes of H_2 gas and the product hydride is an interstitial solid solution, PdH_x: $0 \leq x \leq 0.7$, in which H atoms occupy interstitial sites within the *fcc* Pd metal structure.

Possibly the technologically most important interstitial solid solution is that of C in the octahedral sites of *fcc* γ-Fe. This solid solution is the starting point for steel manufacture. It is useful to consider why C dissolves in γ-Fe but not in *bcc* α-Fe, since this illustrates structural aspects of interstitial solid solution formation.

Iron exists in three polymorphic forms: *bcc* α, stable below 910 °C; γ, stable between 910 and 1400 °C; and *bcc* (again!) δ, stable between 1400 °C and the melting point 1534 °C. γ-Fe can dissolve appreciable amounts of C, up to 2.06 wt%, whereas the α and δ forms dissolve very much less C, up to a maximum of 0.02 and 0.1 wt%, respectively.

Although *fcc* γ-Fe is more densely packed than the *bcc* α structure, the interstitial holes are larger but less numerous in γ-Fe. Unit cells of the two forms are shown in Fig. 2.12 together with the octahedral interstitial sites at cube face centres in α-Fe and at the body centre in γ-Fe. The Fe–C distances and, hence, the sizes of the sites are larger in γ-Fe than in α-Fe. In α-Fe, these sites are also grossly distorted, as shown by the following calculation. The cubic unit cell edge in α-Fe is 2.866 Å. Consequently, two Fe–C distances, involving Fe atoms 1 and 2 in the centres of adjacent cells, Fig. 2.12(a), are $a/2 = 1.433$ Å. The other four Fe–C distances, involving Fe atoms 3–6 at the corners of the cell face, are $a\sqrt{2}/2 = 2.03$ Å. In γ-Fe, Fig. 2.12(b), the octahedral sites are undistorted and the Fe–C distance is $a/2 = 1.796$ Å. This is the value at room temperature. The bond length at 900 °C is likely to be a few percent greater, which is then comparable to expected Fe–C bond distances, as judged by values in the range 1.89–2.15 Å in, for instance, Fe_3C. The two shortest distances of 1.433 Å in α-Fe are prohibitively short and make this an unattractive site for an interstitial C atom.

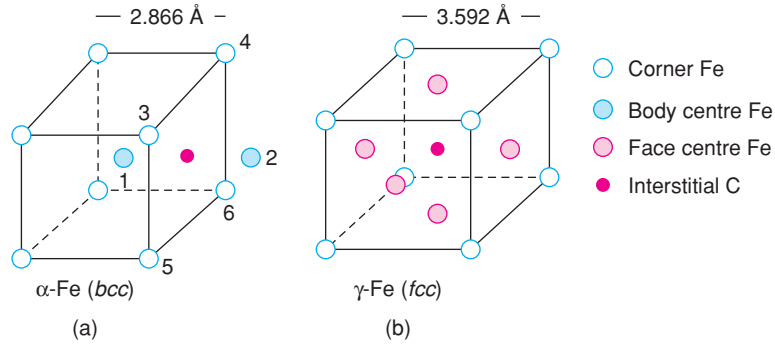


Figure 2.12 ***Interstitial sites for carbon in (a) α-Fe and (b) γ-Fe.***

2.3.3 More complex solid solution mechanisms: aliovalent substitution

So far, we have considered simple solid solutions associated with either site substitution or occupancy of interstitial sites; examples of the latter were taken from metallic systems but, of course, changes in bonding must accompany the incorporation of C into Fe or H into Pd. Formation of interstitial solid solutions is, therefore, more complex than straightforward substitution of one ion by another ion of the same charge. Let us now look at solid solutions from a different and wider perspective.

Substitutional solid solutions can be of two broad types, *homovalent* and *heterovalent* or *aliovalent.* In homovalent substitution, ions are substituted by other ions of the same charge; no additional changes are needed to retain charge balance, as in, for example, $Al_{2-x}Cr_xO_3$: $0 < x < 2$.

In aliovalent solid solutions, ions are substituted by other ions of different charge. Additional changes involving vacancies or interstitials (*ionic compensation*) or electrons or holes (*electronic compensation*) are needed. For ionic compensation, there are four possibilities for cation substitution, summarised in Fig. 2.13(a). A similar scheme is possible for anion substitution but is less common because there are only two other anions of comparable size to O^{2-}, i.e. F^- and N^{3-}.

2.3.3.1 Ionic compensation mechanisms

1. Creating cation vacancies

If the replaceable cation of the host structure has lower charge than that of the cation which replaces it, cation vacancies may form by leaving out additional cations of the host structure. For example, NaCl can dissolve a small amount of $CaCl_2$ in which two Na^+ ions are replaced by one Ca^{2+} ion leaving one Na^+ site, *V,* vacant. The formula is

$$Na_{1-2x}Ca_x V_x Cl$$

Experimentally, at e.g. 600 °C, $0 < x < 0.15$. The Ca^{2+} ions, the cation vacancies and the host Na^+ ions are disordered over the octahedral sites of the rock salt structure. In practice, there may be local ordering of the defects since the substituted Ca^{2+} ions, with an effective excess site charge of $+1$, are attracted to Na^+ vacancies with their effective site charges of -1, in the same way that Schottky defects associate into pairs.

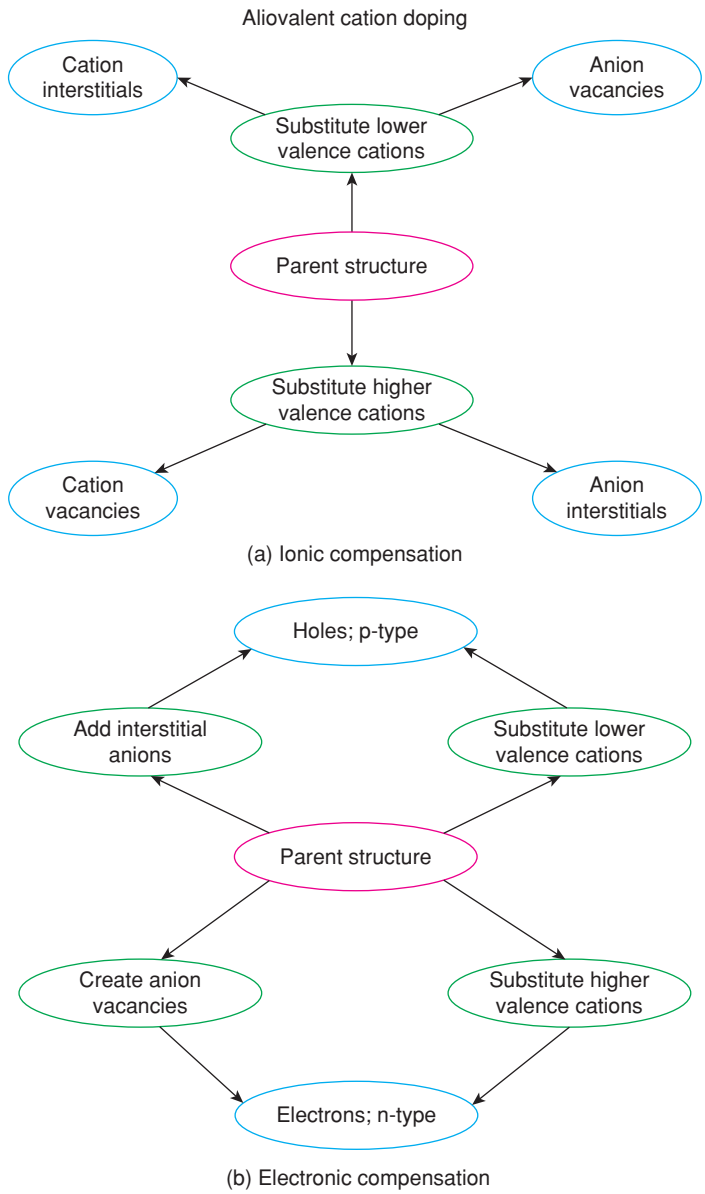


Figure 2.13 ***Solid solution mechanisms involving substitution of aliovalent cations: (a) ionic compensation; (b) electronic compensation.***

There are countless other examples of similar solid solutions in which various charge combinations occur. Thus, divalent Mg^{2+} ions on tetrahedral sites may be replaced by Al^{3+} ions, in the number ratio 3:2 in spinel $MgAl_2O_4$:

$$Mg_{1-3x}Al_{2+2x}O_4$$

in which x tetrahedral site vacancies are also created.

Li_2TiO_3 at high temperatures has a rock salt structure with Li^+ and Ti^{4+} ions disordered over the octahedral sites in a *ccp* oxide ion array. It can form solid solutions with excess TiO_2:

$$Li_{2-4x}Ti_{1+x}O_3:\ 0 < x <\sim 0.19$$

This involves interchange of mono- and tetravalent ions with the creation of Li^+ vacancies to maintain electroneutrality. The large difference in charge of Li^+ and Ti^{4+} obviously does not prevent solid solution formation. Part of the reason why solid solutions form is that Li^+ and Ti^{4+} are able to occupy similar-sized octahedral sites with metal–oxygen distances in the range 1.9–2.2 Å.

Other examples include the corundum-like phase $LiNbO_3$, an important optoelectronic material, which forms a solid solution by $5Li^+ \rightleftharpoons Nb^{5+}$ substitution on octahedral sites. Similarly, Na^+ is of a similar size to Zr^{4+} and these ions replace each other on octahedral sites in the solid solutions $Na_{5-4x}Zr_{1+x}P_3O_{12}$: $0.04 < x < 0.15$.

2. Creating interstitial anions

The other mechanism by which a cation of higher charge may substitute for one of lower charge is to create, at the same time, interstitial anions. This mechanism is not common because most structures do not have interstitial sites large enough to accommodate extra anions. It is favoured by the fluorite structure in certain cases. For example, CaF_2 can dissolve small amounts of YF_3. The total number of cations remains constant and Ca^{2+} and Y^{3+} ions are disordered over the Ca sites, but fluoride interstitials are created to give

$$Ca_{1-x}Y_xF_{2+x}$$

These extra F^- ions occupy large sites surrounded by eight other F^- ions at the corners of a cube, Fig. 1.30.

Another fluorite-related structure capable of containing extra interstitial anions is UO_2. On oxidation it forms a solid solution UO_{2+x}, containing interstitial oxide ions and mixed valency U, i.e.

$$U^{4+}_{1-x}U^{6+}_xO_{2+x}$$

Defect clusters, Fig. 2.10, form, rather than randomly distributed, interstitial oxide ions.

3. Creating anion vacancies

If the replaceable ion of the host structure has higher charge than that of the replacing cation, charge balance may be maintained by creating either anion vacancies or interstitial cations. The best-known examples of anion vacancies occur again with the fluorite structure in oxides such as zirconia, ZrO_2, as in lime-stabilised zirconia:

$$Zr_{1-x}Ca_xO_{2-x}:\ 0.1 < x < 0.2$$

These materials are important in modern technology, both as engineering ceramics and as oxide ion-conducting solid electrolytes (see Chapter 8).

4. Creating interstitial cations

This is a common mechanism provided the host structure has suitably sized interstitial sites to accommodate the extra cations. Good examples are the various 'stuffed silica' phases. These are aluminosilicates in which the structure of one of the three polymorphs of silica – quartz, tridymite or cristobalite – is modified by partial replacement of Si^{4+} by Al^{3+}; at the same time, alkali metal cations enter interstitial holes in the silica framework.

Stuffed quartz structures include $Li_x(Si_{1-x}Al_x)O_2$: $0 < x \leq 0.5$; special compositions exist at $x = 0.5$ ($LiAlSiO_4$, eucryptite) and $x = 0.33$ ($LiAlSi_2O_6$, spodumene). β-Spodumene has the unusual property of a very small coefficient of thermal expansion; β-spodumene ceramics are dimensionally stable and resistant to thermal shock and find high-temperature applications. The interstitial holes in the quartz structure are too small to accommodate cations larger than Li^+. Tridymite and cristobalite have lower densities than quartz with larger interstices. Stuffed tridymite and stuffed cristobalite solid solutions form in which the interstitial or stuffing cations are Na^+ and K^+.

5. Double substitution

Two substitutions take place simultaneously in, e.g., synthetic olivines: Mg^{2+} may be replaced by Fe^{2+} at the same time as Si^{4+} is replaced by Ge^{4+} to give

$$(Mg_{2-x}Fe_x)(Si_{1-y}Ge_y)O_4$$

AgBr and NaCl form solid solutions in which both anions and cations replace each other:

$$(Ag_{1-x}Na_x)(Br_{1-y}Cl_y):\ 0 < x, y < 1$$

The substituting ions may be of different charge, if overall electroneutrality prevails. In the plagioclase feldspars, solid solutions form between anorthite, $CaAl_2Si_2O_8$, and albite, $NaAlSi_3O_8$:

$$(Ca_{1-x}Na_x)(Al_{2-x}Si_{2+x})O_8 : 0 < x < 1$$

The two substitutions, $Na \cup Ca$ and $Si \rightleftharpoons Al$, occur simultaneously and to the same extent.

Double substitution occurs in *sialons*, which are solid solutions in the system Si–Al–O–N, based on the Si_3N_4 parent structure. β-Silicon nitride is built of SiN_4 tetrahedra linked at their corners to form a 3D network. Each N is in planar coordination and forms the corner of three SiN_4 tetrahedra. In the sialon solid solutions, Si^{4+} is partly replaced by Al^{3+} and N^{3-} is partly replaced by O^{2-}. In this way, charge balance is retained. The structural units in the solid solutions are $(Si, Al)(O, N)_4$ tetrahedra and the general formula is

$$(Si_{3-x}Al_x)(N_{4-x}O_x)$$

Silicon nitride and sialons are important high-temperature ceramics.

2.3.3.2 Electronic compensation: metals, semi- and superconductors

With aliovalent substitution, charge compensation may be ionic, as discussed above, to give materials that are electrically insulating or show ionic conductivity associated with either vacancies or interstitials. These examples show no electronic conductivity. In many transition metal-containing materials, however, especially

those in which mixed valency occurs, solid solutions may form which are semiconducting, metallic or even, at low temperatures, superconducting. To discuss these, we use the scheme shown in Fig. 2.13(b), in which higher or lower valence cations *of the same element* are involved, as follows:

1. Cation vacancies may be generated by deintercalation, as in removal of Li^+ (and e^-) from materials such as $LiCoO_2$ and $LiMn_2O_4$. To preserve charge balance, positive holes are created, which are usually located on transition metal cations in the structural framework, as in

$$Li_{1-x}Co^{3+}_{1-x}Co^{4+}_{x}O_2$$
$$Li_{1-x}Mn^{3+}_{1-x}Mn^{4+}_{1+x}O_4$$

   These solid solutions may also be thought of as arising from substitution of higher valence cations (Co^{4+}, Mn^{4+}) for lower valence ones (Co^{3+}, Mn^{3+}), together with creation of cation (Li) vacancies for charge balance, Fig. 2.13(a). These materials are extremely important in commercial secondary lithium batteries: $LiCoO_2$ is used currently as the cathode and $LiMn_2O_4$ is a possible contender to take its place in the next generation of Li batteries (see Section 8.5.4.11).

   Mixed valence solid solutions form with many oxides, which absorb oxygen on heating; the O_2 molecules dissociate, O atoms gain electrons from low oxidation state transition metal ions and, because the structures now have an excess of O^{2-} ions, cation vacancies are created. A good example is pale green, insulating NiO, which oxidises and becomes black and semiconducting on heating:

$$NiO + O_2 \xrightarrow{\text{Heat}} Ni^{2+}_{1-3x}Ni^{3+}_{2x}V_{Ni,x}O$$

   The product has the same rock salt structure as NiO, but with a mixture of Ni^{2+} and Ni^{3+} ions and cation vacancies distributed over the octahedral sites. In reality, the crystals become slightly larger because the extra oxide ions are located at the sample surface and not in interstitial sites. Another good example is iron oxide, $Fe_{1-x}O$, wüstite, which at first sight has a simple face centred cubic rock salt structure with Fe^{2+} and Fe^{3+} ions and cation vacancies distributed over the octahedral sites. However, as discussed in Section 2.2.5, the defect structure is rather more complex.
2. Creation of mixed valency cations may also be associated with insertion of interstitial anions. Several of the new families of high-T_c superconductors are solid solutions in this category. The best known is YBCO or Y123, $YBa_2Cu_3O_\delta$. Depending on the oxygen content δ, Cu has either a mixed +I, +II valence (e.g. for $\delta = 6$), is entirely +II (for $\delta = 6.5$) or is a mixture of +II and +III (for $\delta = 7$). Starting with $\delta = 6$, extra oxygen may be introduced into the structure (by heating in air or O_2 at 350 °C), oxidation of Cu ions occurs and the material gradually transforms from a semiconductor, for $\delta = 6$, to a superconductor with a critical temperature, T_c, of 90 K for $\delta = 7$ (Section 8.3.7).
3. In the case of $YBa_2Cu_3O_\delta$, the extra interstitial sites that are occupied by O for $6 < \delta < 7$ are completely empty for $\delta = 6$ and full for $\delta = 7$. Hence this may be regarded either as an interstitial solid solution, beginning at $\delta = 6$ or a vacancy solid solution starting from $\delta = 7$.

   A number of oxides lose oxygen with creation of anion vacancies and associated reduction at high temperatures, especially if heated in a reducing atmosphere. The process may be regarded as

$$2O^{2-} \rightarrow O_2 + 4e^-$$

   The liberated electrons enter the structure and form a mixed valence state with any transition metal cations present. The resulting materials are often semiconducting or metallic. Examples are TiO_{2-x} (see later for its defect structure which involves crystallographic shear planes), WO_{3-x} and $BaTiO_{3-x}$ (above 1400 °C).

4. Interstitial cations of one element and mixed valency cations of another element arise on intercalation of, e.g., Li into MnO_2; this is simply the reverse of the process described in point 1 for $LiCoO_2$ and $LiMn_2O_4$. Other well-known examples are the tungsten bronzes formed by reacting WO_3 with alkali metals, *n*-butyllithium or by electrochemical insertion. These have formulae such as

$$Li_xW^{6+}_{1-x}W^{5+}_xO_3$$

The mixed valence of W gives rise to metallic conductivity and a bronze-like appearance. Since Li can be rapidly and reversibly added to or removed from the structure, these materials find applications in thin-film electrochromic devices and glass coatings since the colour changes dramatically with x in certain cases (see Section 8.5.4.11).

Mixed valence can also arise when one element forms half of a double substitution mechanism. A good example is Ba-doped La_2CuO_4, the material discovered by Bednorz and Muller in 1986 which triggered the scientific revolution in high-T_c superconductivity. The formula is

$$La_{2-x}Ba_xCu^{2+}_{1-x}Cu^{3+}_xO_4$$

A second example is semiconducting $BaBiO_3$, containing Bi^{3+} and Bi^{5+} as discrete species. Superconducting materials are formed by substitution of Ba^{2+} for K^+; $BaBiO_3$ has a distorted perovskite structure, with equal numbers of Bi^{3+} and Bi^{5+}. On K doping, the proportion of Bi^{5+} increases to give:

$$Ba_{1-2x}K_{2x}Bi^{3+}_{0.5-x}Bi^{5+}_{0.5+x}O_3$$

In reality, the structure is more complex since the crystallographic distinction between Bi^{3+} and Bi^{5+} is lost and the K-doped structure reverts to a simple cubic perovskite. All we can say, then, is that the average valence state of Bi is $>4+$ on K substitution. These examples do not fit within the simple schemes shown in Fig. 2.13 but are further examples of double substitution.

2.3.4 Thermodynamically stable and metastable solid solutions

Solid solutions may be *thermodynamically stable* or *metastable*. Thermodynamically stable solid solutions appear on the appropriate phase diagram (see Chapter 7), and represent the structures that must exist in equilibrium at a given composition and temperature (and oxygen partial pressure if important). As the phase diagrams often show, solubility limits may change with temperature and attention to cooling rate may be important if it is desired to quench high-temperature solid solutions to lower temperatures; otherwise precipitation of a secondary phase may occur.

Under non-equilibrium conditions, however, and especially by using '*chimie douce*' or other preparative techniques, it is often possible to prepare solid solutions that are much more extensive than those existing under equilibrium conditions. A simple example is provided by the β-aluminas, $Na_2O \sim 8Al_2O_3$. Part or all of the Na^+ ions may be ion exchanged for a variety of other monovalent ions, including Li^+, K^+, Ag^+ and Cs^+, even though most of these products are not thermodynamically stable (see Chapter 8).

2.3.5 Experimental methods for studying solid solutions

2.3.5.1 X-ray powder diffraction, XRD

There are three main ways in which XRD (see Chapter 5) may be used to study solid solutions. The first is as a simple fingerprint method in which the objective is to determine the crystalline phases that are present without

necessarily measuring the patterns accurately. Although powder XRD is the most important technique by far for identifying crystalline materials, the detection limit of phases in a mixture is usually of the order of 2–3 wt%, which means that if the sample is thought to be phase-pure by XRD, it is always possible that there is a small amount of secondary phases present that is below the detection limit of the technique.

The second is to measure the XRD pattern of solid solutions accurately and obtain their unit cell dimensions, which may undergo a small contraction or expansion as composition varies. If a calibration graph of unit cell dimensions against composition is available or can be constructed, this can be used to determine the composition of solid solutions in a particular sample. Usually, a unit cell expands if a small ion is replaced by a larger ion and vice versa, contracts if a smaller ion is substituted into the structure. From Bragg's law and the *d*-spacing formulae (see Chapter 5), the whole pattern shifts to lower values of 2θ with increasing unit cell parameters, although all the XRD lines do not necessarily move by the same amount in non-cubic crystals, if the expansion is not isotropic.

According to *Vegard's law*, unit cell parameters should change linearly with solid solution composition. In practice, Vegard's law is often obeyed only approximately and accurate measurements may reveal departures from linearity. Vegard's law is not really a law but rather a generalisation that applies to solid solutions formed by random substitution or distribution of ions. It assumes implicitly that the changes in unit cell parameters with composition are governed purely by the relative sizes of the atoms or ions that are 'active' in the solid solution mechanism, e.g. the ions that replace each other in a simple substitutional mechanism.

Third, using Rietveld refinement of the powder XRD patterns of solid solutions, and in particular the intensities of the XRD reflections, it is possible to gain detailed crystallographic information such as the sites occupied by atoms and the location of vacancies and interstitials. Again, however, this is not sensitive to small compositional changes and could not be used to locate small dopant concentrations.

2.3.5.2 Density measurements

The mechanism of solid solution formation may sometimes be inferred by a combination of density and unit cell volume measurements for a range of compositions. The key parameter is the mass of the average unit cell contents and whether it increases or decreases on solid solution formation.

As an example, consider the stabilised zirconia solid solutions formed between ZrO_2 and CaO over the approximate range 10–25% CaO. Two simple mechanisms could be postulated: (a) the total number of oxide ions remains constant and, therefore, interstitial Ca^{2+} ions are created according to the formula $(Zr_{1-x}Ca_{2x})O_2$; (b) the total number of cations remains constant and O^{2-} vacancies are created according to the formula $(Zr_{1-x}Ca_x)O_{2-x}$. In (a), two Ca^{2+} ions replace one Zr^{4+} and the formula unit decreases in mass by 11 g as x varies, hypothetically, from 0 to 1. In (b), one Zr and one O is replaced by one Ca with a decrease in mass of the formula unit by 67 g as x varies from 0 to 1. Assuming that the cell volume does not change with composition (not strictly true), mechanism (b) would lead to a larger decrease in density with x than mechanism (a).

Experimental results, Fig. 2.14, confirm that mechanism (b) is operative, at least for samples heated at 1600 °C. It is possible, in theory at least, to propose alternative and more complex mechanisms than (a) and (b); e.g. the total number of Zr^{4+} ions remains constant, in which case both interstitial Ca^{2+} and O^{2-} ions are needed. Usually, however, simple mechanisms operate and there is no need to consider more far-fetched possibilities.

Density data for the $CaF_2 - YF_3$ solid solutions described earlier are given in Fig. 2.15. These clearly show that a model based on interstitial fluoride ions fits the data rather than one based on cation vacancies. Density measurements do not, of course, give atomistic details of the vacancies or interstitials involved, but only an overall average. Other techniques, such as diffuse neutron scattering, are needed to probe the defect structure.

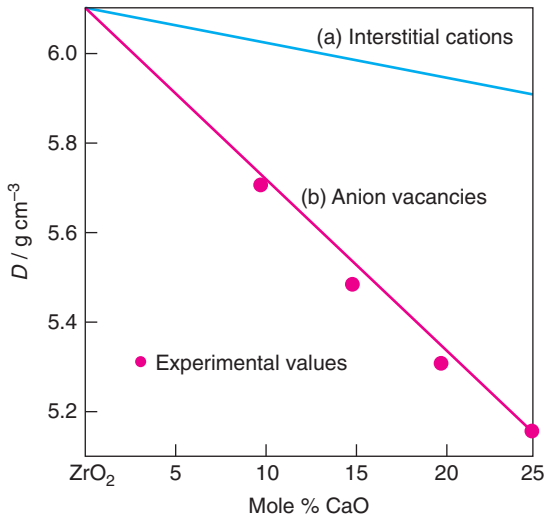


Figure 2.14 ***Density data for cubic CaO-stabilised zirconia solid solutions for samples quenched from 1600 °C. Data reproduced from Diness and Roy,* Solid State Commun.*, 3, 123, © 1965 Elsevier.***

Densities can be measured by several techniques. The volume of a few grams of material may be measured by displacement of liquid from a specific gravity bottle whose volume is known accurately. From the difference in weight of the bottle filled with displacement liquid, e.g. CCl_4, and the bottle containing the solid topped up with liquid, the volume of the solid may be calculated if the density of the displacement liquid is known. In the float–sink method, a few crystals of the material are suspended in liquids of a range of densities until a liquid is found in which the crystals neither sink nor float. The density of the crystals then equals that of the liquid. A variant is to use a density gradient column, which is a column of liquid of gradually increasing density.

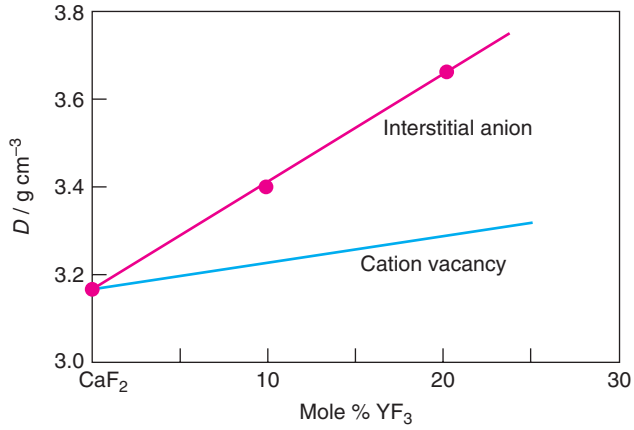


Figure 2.15 ***Density data for solid solutions of YF_3 in CaF_2. Reproduced from Kingery, Bowen and Uhlmann,* Introduction to Ceramics*, Wiley, New York, © 1976 John Wiley & Sons, Inc. Reproduced under the terms of the STM Agreement.***

Crystals are dropped in at the top and sink until their density equals that of the liquid. The crystal density is obtained from a calibration curve of height in the column against density. In all of the above methods, it is important that air bubbles are not trapped on the surface of the crystals, otherwise anomalously low density values are obtained.

A good method for measuring the density of larger samples (10–100 g) is by gas displacement pycnometry. In this, the sample is compressed in a gas-filled chamber by a piston until a certain pressure, e.g. 2 atm, is reached. The sample volume is obtained from the difference in the position of the piston when the sample is in the chamber compared with that when the chamber is empty and contains only gas compressed to the same pressure.

2.3.5.3 Changes in other properties – thermal activity and DTA/DSC

Many materials undergo abrupt changes in structure or property on heating and, if the material forms a solid solution, the temperature of the change usually varies with composition. The changes can often be studied by differential thermal analysis/differential scanning calorimetry (DTA/DSC) since most phase transitions have an appreciable enthalpy of transition. This provides a very sensitive way of studying solid solutions because transition temperatures vary over tens or hundreds of degrees, as shown by the following examples:

- Addition of carbon to iron causes the temperature of the $\alpha \rightleftharpoons \gamma$ transition to drop rapidly from 910 to 723 °C with addition of only 0.02 wt% carbon.
- The ferroelectric Curie temperature of $BaTiO_3$, ~125 °C, is very sensitive to doping; most dopants on either A or B sites reduce T_c but substitution of Pb for Ba increases T_c. The effect of Ca as a dopant is interesting since isovalent substitution for Ba causes very little change in T_c whereas aliovalent substitution for Ti, together with creation of oxygen vacancies, causes a rapid reduction in T_c, Fig. 2.16.

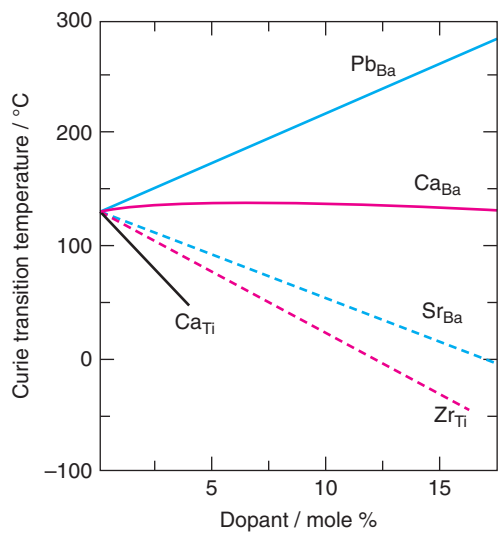


Figure 2.16 ***Effect of dopants on the ferroelectric Curie temperature of $BaTiO_3$ showing the effect of isovalent substitution of Sr, Ca and Pb for Ba, isovalent substitution of Zr for Ti and aliovalent substitution of Ca for Ti.***

- Pure zirconia, ZrO_2, cannot be used as a ceramic even though it is inert at high temperatures and has a very high melting point, because, during cooling, it undergoes a phase transition from a tetragonal to a monoclinic polymorph and the large volume change associated with the transition causes ceramics made from zirconia to disintegrate. This problem is solved by partially replacing Zr^{4+} ions by e.g. Ca^{2+} or Y^{3+} ions, which has the effect of depressing entirely the phase transition temperature and allowing either the cubic or the tetragonal polymorph to be stabilised to room temperature.

2.4 Extended Defects

2.4.1 Crystallographic shear structures

For a long time, it was known that certain transition metal oxides could be prepared with an apparent wide range of non-stoichiometry, e.g. WO_{3-x}, MoO_{3-x}, TiO_{2-x}. Following the work of Magneli, it was recognised that in these systems, instead of continuous solid solution formation, a series of closely related phases with very similar formulae and structures existed. In oxygen-deficient rutile, a homologous series of phases occurs of formula Ti_nO_{2n-1} with $n = 3$–10. Thus, $Ti_8O_{15}(TiO_{1.875})$ and $Ti_9O_{17}(TiO_{1.889})$ are each homogeneous, physically separate phases.

The structural principle underlying these phases was worked out by Magneli, Wadsley and others and the term *crystallographic shear* was coined for the type of 'defect' involved. In the oxygen-deficient rutiles, regions of unreduced rutile structure are separated from each other by crystallographic shear planes (CS planes), which are thin lamellae of rather different structure and composition. All of the oxygen deficiency is concentrated within these CS planes. With increased reduction, the variation in stoichiometry is accommodated by increasing the number of CS planes and decreasing the thickness of the rutile blocks between adjacent CS planes.

For the purpose of understanding CS structures, it is useful to consider schematically how they might form on reduction of WO_3, Fig. 2.17; the CS planes in WO_3, which has the ReO_3 structure, Section 1.17.8, are easier to draw than those in rutile. WO_3 has a 3D framework structure of corner-sharing octahedra (a). The first step in the reduction involves formation of some vacant oxygen sites and reduction of W^{6+} to W^{5+}. The vacant oxygen sites are not distributed at random but are located on certain planes within the crystal (b). Such a structure would, of course, be unstable and so a partial collapse occurs to eliminate the layer of vacancies and form the CS planes (c). As a result of this condensation, octahedra within the CS planes share some edges, whereas in unreduced WO_3, the linkages are by corner-sharing only. The CS plane is outlined in heavy lines in (c) and runs obliquely through the structure.

CS planes may occur at random or be spaced at regular intervals. If they are at random, they may be regarded as planar defects and are known as *Wadsley defects*. If they are regularly spaced, then separate, ordered phases may be identified. With increased reduction, the number of CS planes increases and the average spacing between adjacent CS planes decreases. In the homologous series of reduced WO_3 phases, the integrity and composition of each member are given by the distance between adjacent CS planes. Thus, as n decreases, so the spacing between adjacent CS planes decreases.

So far, we have considered homologous series of phases in which the structure and orientation of the CS planes are the same in each member of the series. CS planes may occur in different orientations, however, in which case their detailed structure also changes. This is shown in Fig. 2.17(d); the CS plane is in a different orientation to that in (c) and the linkages of the octahedra are different. The example in (d) is actually for Mo_nO_{3n-1}: $n = 8$, which is also derived from the cubic WO_3 structure.

In reduced rutiles, the CS planes contain face-sharing octahedra, whereas edge-sharing occurs in unreduced rutile. Two series of homologous phases form, both with the general formula Ti_nO_{2n-1} but the orientations of

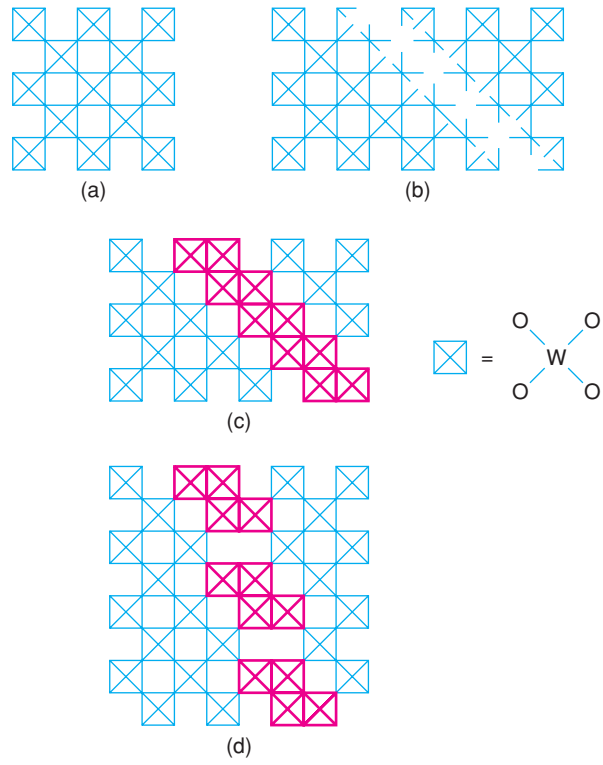


Figure 2.17 ***Formation of CS planes in WO_3 and related structures. Each crossed square represents a chain of WO_6 octahedra perpendicular to the plane of the paper linked by sharing corners. The parent, unreduced part of the structure is the same as that of cubic ReO_3.***

the CS planes are separated by an angular rotation of 11.53°. One series has $3 < n < 10$ and the other has $16 < n < 36$. Compositions with $10 < n < 14$ are particularly interesting because the orientation of the CS planes gradually changes. Each composition has its own value for the angular orientation of the CS planes, known as 'swinging shear planes,' which are perfectly well ordered. Since each composition in the range $n = 10$–14 has a unique structure, this raises interesting questions as to exactly what is meant by the term 'a phase'. Thus, these compositions cannot be described as forming a solid solution since every structure is different and cannot therefore be described in terms of an average structure that represents the whole composition range.

Reduced vanadium dioxides, V_nO_{2n-1}: $4 \leq n \leq 8$, and mixed oxides of chromium and titanium, $(Cr, Ti)O_{2-5}$, also have structures containing CS planes similar to reduced rutiles.

In the examples considered so far, the crystals contain one set of parallel and regularly spaced CS planes. The regions of unreduced structure are limited to thin slabs or sheets. In reduced Nb_2O_5 and mixed oxides of Nb, Ti and Nb, W, CS planes occur in *two* orthogonal sets (i.e. at 90° to each other) and the regions of perfect (unreduced) structure are reduced in size from infinite sheets to infinite columns or blocks. These 'block' or 'double shear' structures are characterised by the length, width and manner of connectivity of the blocks of unreduced ReO_3 structure. In addition to having phases which are built of blocks of only one size, the complexity can be much increased by having blocks of two or three different sizes arranged in an ordered fashion. Examples are $Nb_{25}O_{62}$, $Nb_{47}O_{116}$, $W_4Nb_{26}O_{77}$ and $Nb_{65}O_{161}F_3$. Their formulae can be written as

members of homologous series, but are rather clumsy, involving several variables, and are not reproduced here.

In principle, it should be possible to have structures with *three* sets of mutually orthogonal CS planes, in which the regions of unreduced material diminish in size to small blocks of finite length. As yet, there appear to be no known members of this group.

Structures that contain CS planes are normally studied by XRD and high-resolution electron microscopy (HREM). Single-crystal XRD is, of course, the most powerful method for solving crystal structures and in this sense, electron microscopy usually plays a minor role; its use is limited to the determination of unit cells and space groups for very small crystals and to studying defects such as stacking faults and dislocations. However, the technique of direct *lattice imaging* (see Section 6.2.2.5), by electron microscopy is very useful for structural studies of CS phases.

An image of the projected structure at about 3 Å resolution is obtained which usually takes the form of fringes or lines which correspond to the more strongly diffracting heavy metal atoms. The separation of the fringes in a 'perfect' crystal should be absolutely regular. Whenever a CS plane is imaged, irregularity in the fringe spacing occurs because, as a result of the condensation to form a shear plane, the metal atom separation across the CS plane is reduced. A pair of fringes that are more closely spaced than normal may indicate a CS plane. By counting the number of normal fringes in each block, the n value of the phase in the homologous series can usually be determined. If the structure of one member of the series has been worked out in detail using X-ray methods, it is a relatively simple matter to deduce the structures of the remainder from electron microscopy.

In this way Wadsley defects (random CS planes) can be recognised immediately, as can heterogeneities within supposedly single crystals, e.g. if a crystal is zoned and has a slight variation in composition with position or if a crystal is composed of intergrowths of two or more phases.

2.4.2 Stacking faults

Stacking faults are common in materials with layered structures, especially those that exhibit *polytypism*, and are also examples of 2D or plane defects. Co metal exhibits both polytypism and stacking faults. It occurs in two main polytypes, either *ccp* (ABC) or *hcp* (AB). In these, the structures are the same in two dimensions, i.e. within the layers, and differ only in the third dimension, i.e. the sequence of layers. *Stacking disorder* occurs when the normal stacking sequence is interrupted by the presence of 'wrong' layers, e.g. . . . ABABABA*BCA*BABA . . . The letters in italics correspond to layers that either are completely wrong (*C*) or do not have their normal neighbouring layers (*A* and *B*) on either side. Graphite also exhibits polytypism (usually *hcp* but sometimes *ccp* of carbon atoms) and stacking disorder.

2.4.3 Subgrain boundaries and antiphase domains (boundaries)

One type of imperfection in many so-called single crystals is the presence of a domain or mosaic texture. Within the domains, typically ~10 000 Å in size, the structure is relatively perfect, but at the interface between domains there is a structural mismatch, Fig. 2.18. This mismatch may be very small and involve a difference in angular orientation between the domains that is several orders less than 1°. The interfaces between grains are called *subgrain boundaries* and can be treated in terms of dislocation theory (Section 2.5).

Another type of boundary, an *antiphase boundary*, involves a relative lateral displacement of two parts of the same crystal, shown schematically for a 2D crystal, AB, in Fig. 2.19. Across the antiphase boundary, like atoms face each other and the . . . ABAB . . . sequence (in the horizontal rows) is reversed. The term antiphase arises because, if the A and B atoms are regarded as the positive and negative parts of a wave, then a phase change of π occurs at the boundary.

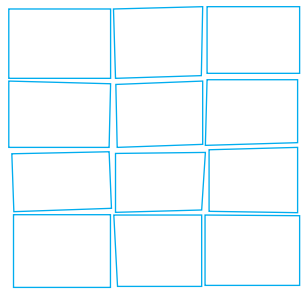


Figure 2.18 ***Domain texture in a single crystal.***

The occurrence of antiphase domains can be seen by dark field imaging in the electron microscope; the boundaries are seen as fringes. If the boundaries are regularly spaced, unusual diffraction effects occur; a superlattice is associated with the ordering of domains but because, at the boundaries, the structure expands due to repulsions between like atoms, satellite spots appear on either side of the supercell positions in the reciprocal lattice. This has been observed in alloys such as CuAu and in minerals such as plagioclase feldspars.

2.5 Dislocations and Mechanical Properties of Solids

Dislocations are an extremely important class of crystal defect:

- They are responsible for the relative weakness of pure metals and in certain cases (after work-hardening) for just the opposite effect of extra hardness.
- The mechanism of crystal growth from either solution or vapour appears to involve dislocations.
- Reactions of solids often occur at active surface sites where dislocations emerge from the crystal.

Dislocations are stoichiometric line defects. Their existence was postulated long before direct experimental evidence for their occurrence was obtained. There were several types of observation which indicated that defects other than point defects must be present in crystals:

(a) Metals are generally much softer than expected. Calculations of the shearing stress of metals gave values of $\sim 10^6$ psi whereas experimental values for many metals are as low as $\sim 10^2$ psi. This indicated that there must be some kind of weak link in their structures which allows metals to cleave so easily.

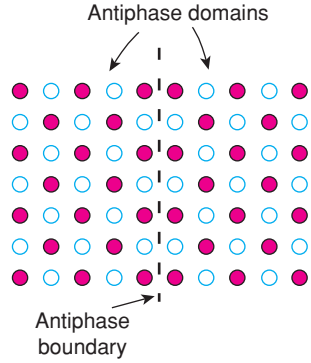


Figure 2.19 ***Antiphase domains and boundaries in an ordered crystal AB: A, open circles; B, closed circles.***

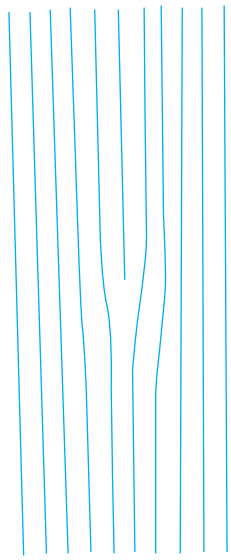


Figure 2.20 ***Edge dislocation in projection.***

(b) Many well-formed crystals were seen under the microscope, or even with the naked eye, to have spirals on their surfaces which clearly provided a mechanism for crystal growth. Such spirals could not occur in perfect crystals, however.
(c) The malleable and ductile properties of metals were difficult to explain without invoking dislocations. Thus, ribbons of magnesium metal can be stretched out to several times their original length, almost like chewing gum, without rupture.
(d) The process of work-hardening of metals was difficult to explain without invoking dislocations.

Dislocations can be one of two extreme types, edge or screw dislocations, or can have any degree of intermediate character.

2.5.1 Edge dislocations

A simple *edge dislocation*, Fig. 2.20, is an extra half-plane of atoms, i.e. a plane of atoms that goes only part of the way through a crystal structure. Planes of atoms within the structure are shown in projection as lines. These lines are parallel except in the region where the extra half-plane terminates. The centre of the distorted region is a line that passes right through the crystal, perpendicular to the paper, and approximates to the end of the extra half-plane; this is the *line* of the dislocation. Outside this stressed region, the crystal is essentially normal; the top half in the drawing must be slightly wider than the bottom half in order to accommodate the extra half-plane.

In order to understand the effect of dislocations on the mechanical properties of crystals, consider the effect of applying a shearing stress to a crystal that contains an edge dislocation, Fig. 2.21. The top half of the crystal is pushed to the right and the bottom half to the left. Comparing (a) and (b), the extra half-plane which terminates at 2 in (a) can effectively move simply by breaking the bond 3–6 and forming a new linkage 2–6. Thus, with a minimum of effort, the half-plane has moved one unit of distance in the direction of the applied

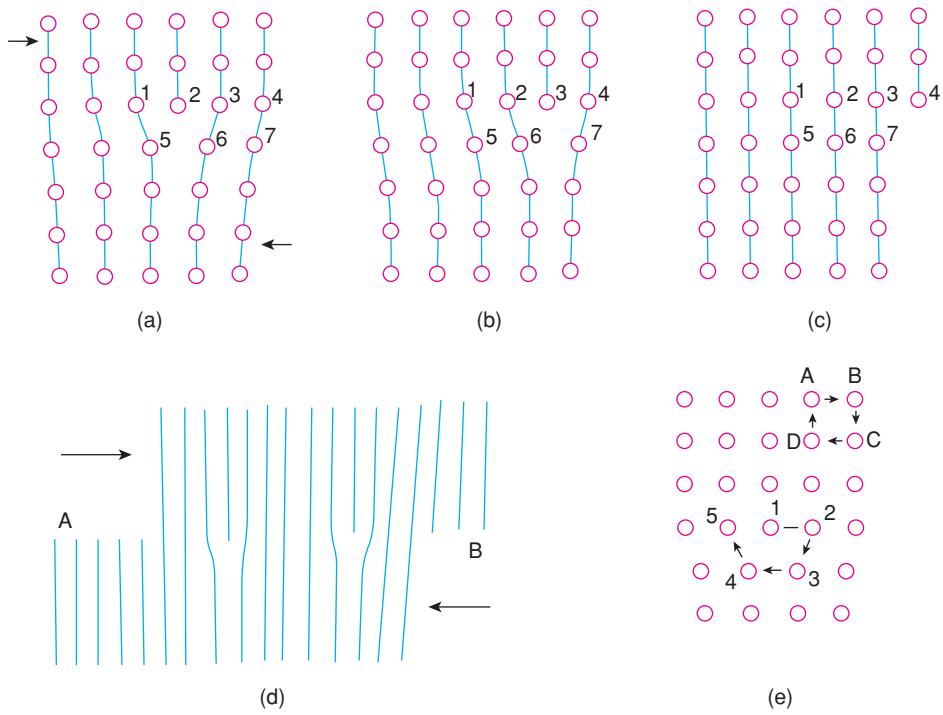


Figure 2.21 ***Migration of an edge dislocation under the action of a shearing stress.***

stress. If this process continues, the extra half-plane eventually arrives at the surface of the crystal as in (c). All that is now needed is an easy means of generating half-planes and, by a process of repetition, the crystal will eventually shear completely. In (d), it is assumed that half-planes are generated at the left-hand end of the crystal (mechanisms of generation are discussed later). With each one that is generated, another one, equal and opposite in orientation and sign, is left behind. At the left-hand side, five half-planes in the bottom part of the crystal have accumulated at the surface. Balancing these in the top part of the crystal, three half-planes have arrived at the right-hand end and two more (*positive dislocations*) are still inside the crystal and in the process of moving out.

The easy motion of dislocations has an analogy in one's experience of trying to rearrange a carpet, preferably a large one. The method of lifting up one end and tugging needs a great deal more effort than making a ruck in one end of the carpet and gliding this through to the other end.

The process of movement of dislocations is called *slip* and the pile-up of half-planes at opposite ends gives ledges or *slip steps*. The line AB in (d) represents the projection of the plane over which the dislocation moves and is called the *slip plane*.

Dislocations are characterised by a vector, the *Burgers vector*, **b**. To find the magnitude and direction of **b**, it is necessary to make an imaginary atom-to-atom circuit around the dislocation, Fig. 2.21(e). In normal regions of the crystal, a circuit such as ABCDA, involving one unit of translation in each direction, is a closed loop and the starting point and finishing point are the same, A. However, the circuit 12345 which passes round the dislocation is not a closed circuit because 1 and 5 do not coincide. The magnitude of the Burgers vector is given by the distance 1–5 and its direction by the direction 1–5 (or 5–1). For an edge dislocation, **b** is

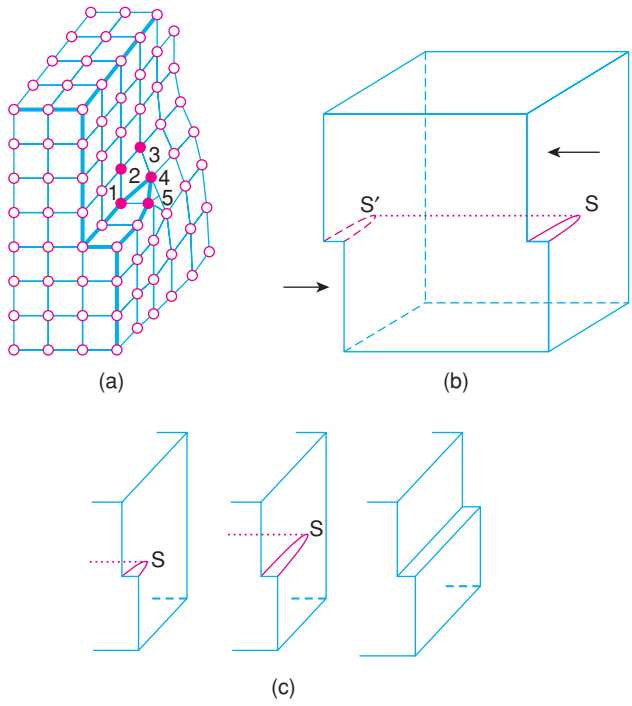


Figure 2.22 ***Screw dislocation.***

perpendicular to the line of the dislocation and parallel to the direction of motion of the line of the dislocation under the action of an applied stress. It is also parallel to the direction of shear.

2.5.2 Screw dislocations

The *screw dislocation* is a little more difficult to visualise and is shown in Fig. 2.22. In (b), the line SS′ represents the line of the screw dislocation. In front of this line the crystal has undergone slip but behind the line it has not. The effect of continued application of a shearing stress, arrowed in (b), is such that the slip step gradually extends across the whole face (side faces) of the crystal at the same time as the line SS′ moves towards the back face (c). To find the Burgers vector of a screw dislocation, consider the circuit 12345 (a) which passes around the dislocation. The magnitude and direction of the distance 1–5 defines **b**. For a screw dislocation, the Burgers vector is parallel to the line of the dislocation (SS′) and perpendicular to the direction of motion of this line. This contrasts with the edge dislocation and it can be seen that edge and screw dislocations are effectively at 90° to each other. For both edge and screw dislocations, **b** is parallel to the direction of shear or slip. The origin of the term ‘screw’ is easy to see by considering the atoms shaded in (a). They lie on a spiral which passes right through the crystal and emerges with opposite hand or sign at S′. As in the case of an edge dislocation, it is necessary to break only a few bonds in order for a screw dislocation to move. Thus in (a), the bond between atoms 2 and 5 has just broken and 2 has joined up with 1; the bond between 3 and 4 will be the next to break. Note that it is useful to consider bonds as breaking and forming although in practice it is not nearly as clear-cut as this; in metals and ionic crystals the bonds are certainly not covalent.

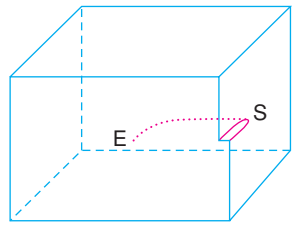


Figure 2.23 ***A quarter dislocation loop.***

2.5.3 Dislocation loops

The generation of dislocations is complicated but seems always to involve the formation of *dislocation loops*. Consider the crystal shown in Fig. 2.23. On looking at the right-hand face, it is obvious that a screw dislocation emerges near to point S, where the slip step terminates. However, this dislocation does not reach the left-hand face because there is no corresponding slip step on it. Looking now at the front face, in isolation, it is clear that there is an extra half-plane and a positive edge dislocation in the top part of the crystal which emerges, say, at E. However, once again this dislocation does not extend to the appropriate point on the opposite face, i.e. the back face. If a dislocation enters a crystal, it must reappear somewhere because, by their nature, dislocations cannot just terminate inside a crystal. What happens in this case is that the two dislocations change direction inside the crystal and meet up to form a *quarter dislocation loop*. Thus, the same dislocation is pure edge at one end, pure screw at the other, and in between has a whole range of intermediate properties. It is difficult without the aid of a 3D model to picture the structural distortions which must occur around that part of the dislocation which has a mixed edge and screw character.

The origin of dislocation loops is by no means well understood. One source, discussed later, is by means of clustering of vacancies on to a plane, followed by an inward collapse of the structure. This generates loops which are entirely inside the crystal. For the purpose of illustration, an alternative source is shown in Fig. 2.24. To start a loop, all that is needed is a nick in one edge of the crystal; the displacement of a few atoms is then sufficient to create a small quarter loop (a). Once created, the loops can expand very easily (b). Usually loops do not expand symmetrically because edge dislocations move more easily and more rapidly than corresponding screw dislocations. The result is shown in (c); the edge component has sped across to the

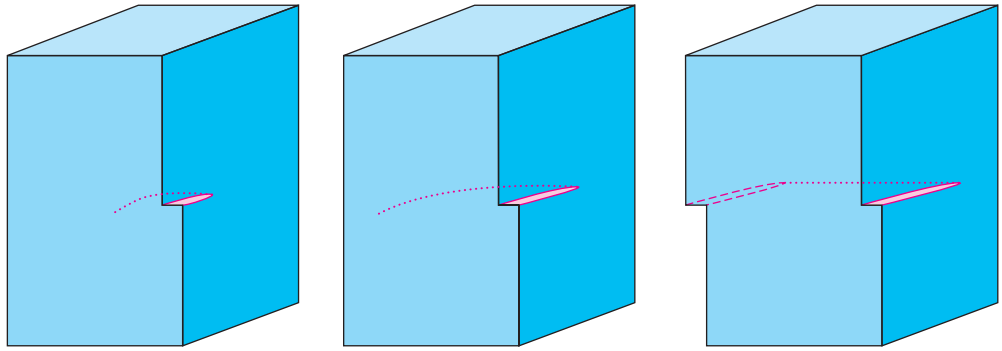


Figure 2.24 ***Generation and motion of a dislocation loop.***

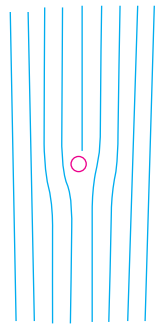


Figure 2.25 ***Locking of an edge dislocation at an impurity atom.***

opposite, left-hand face and the slip process for the front few layers of the crystal is complete. The dislocation which remains is mainly or entirely screw in character and this continues to move slowly to the back of the crystal, thus terminating the slip process for this dislocation.

As a consequence of the very easy generation of dislocation loops, the mechanical strength of metals is greatly reduced. This is obviously very serious for metals which are used in construction. The impact resistance of metals to high stresses is easy to measure because the experiments can be carried out rapidly. However, much more difficult to assess is the resistance over a long period of time to much smaller stresses. Small stresses may be sufficient to generate and move dislocations only very slowly but the results are cumulative and, generally, non-reversible, so that one day catastrophic failure of the metal may occur, perhaps for no apparent reason.

In addition to dislocations being a great source of weakness of materials, they can also have the reverse effect and greatly increase the strength or hardness. One mechanism involves the 'locking' of dislocations at certain impurity atoms, Fig. 2.25. The dislocation moves freely until it encounters the impurity atom and then effectively becomes trapped by it. Hardening of steels with interstitial carbon atoms involves precipitation of iron carbide particles which effectively act as a trap for dislocations.

A very important strengthening process in metals is *work-hardening*. On hammering a metal, an enormous number of dislocations are generated, which in a polycrystalline material are in a large number of orientations. These start to move through the crystal but, sooner or later, depending on the metal and its structure, movement stops. Grain boundaries provide an effective means of stopping dislocations inside the grains because, as a dislocation passes out of a grain, the surface of the grain becomes deformed and the resulting stress imposed upon neighbouring grains may act to prevent the dislocation from ever reaching the surface. Because the area around a dislocation is stressed, two dislocations may repel each other if they get too close. Thus, once the leading dislocations become trapped at grain boundaries or against dislocations arriving from other directions, the following dislocations pile up. A log-jam of dislocations then forms rapidly and the individual dislocations cannot move either forwards or backwards, so that a considerable increase in the strength of the metal results. This process is called *strain-hardening*.

Strain-hardened metals can be rendered malleable and ductile once again by high-temperature annealing. At high temperatures, atoms have considerable thermal energy which enables them to move. Dislocations may therefore be able to reorganise themselves or annihilate each other. If a positive and a negative edge dislocation meet on the same slip plane, they cancel each other out, leaving behind a strain-free area in the crystal. The process can be very rapid, e.g. platinum crucibles used in laboratory experiments may be softened in a few minutes by placing them at, say, 1200 °C.

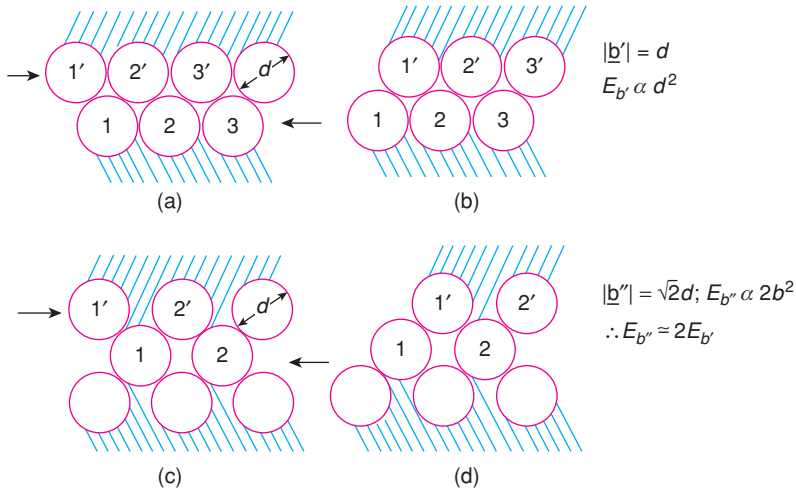


Figure 2.26 ***Burgers vector in (a), (b)* cp *and (c), (d) non-*cp *directions.***

2.5.4 Dislocations and crystal structure

For a given structure, there is usually a preferred plane or set of planes on which dislocations can occur and also preferred directions for dislocation motion. The energy, E, that is required to move a dislocation by one unit of translation is proportional to the square of the magnitude of the Burgers vector, **b**, i.e.

$$\boldsymbol{E} \propto |\mathbf{b}|^2 \tag{2.23}$$

Thus, the dislocations that are most important are those with the smallest **b** value.

In metals, the direction of motion of dislocations is usually parallel to a direction of close packing in the structure. In Fig. 2.26(a), two rows of spheres are shown, each representing a layer of atoms in projection; the rows are close packed, because the atoms of each row touch neighbouring atoms in the same row. Under the action of a shearing stress and with the aid of a dislocation (not shown), the top row shears over the bottom row by one unit. After shear, (b), atom $1'$ occupies the position that was previously occupied by atom $2'$ in (a), and so on. The distance $1'$–$2'$ corresponds to the unit of translation for slip and for dislocation motion. It is equivalent to the magnitude of the Burgers vector, i.e. $|\mathbf{b}'| = d$, where d is the diameter of the spheres. Therefore $E_b \propto d^2$.

An example of dislocation motion in a non-close packed direction is shown in Fig. 2.26(c) and (d). The distance $1'$–$2'$ is longer than in (a) and (b); by a simple geometric construction, it can be shown that $\left|\mathbf{b}''\right| = d\sqrt{2}$. Therefore, $E_b'' \propto 2d^2$. Consequently, motion of dislocations in this particular non-*cp* direction requires double the energy for motion in a *cp* direction.

The preferred direction of dislocation motion is, therefore, parallel to a *cp* direction. The preferred plane of shear or slip is usually a close packed plane. This is because the distance of separation between two *cp* planes, d_1, is greater than the corresponding distance between two non-*cp* planes, d_2, Fig. 2.27. During shear, the atoms to either side of the slip plane move in opposite directions and the distortion of the structure for intermediate or saddle positions is much less in (a) than in (b). Alternatively, the energy barrier to be surmounted by an atom in moving from one position to an equivalent one (represented by the curved arrows) is much greater for (b).

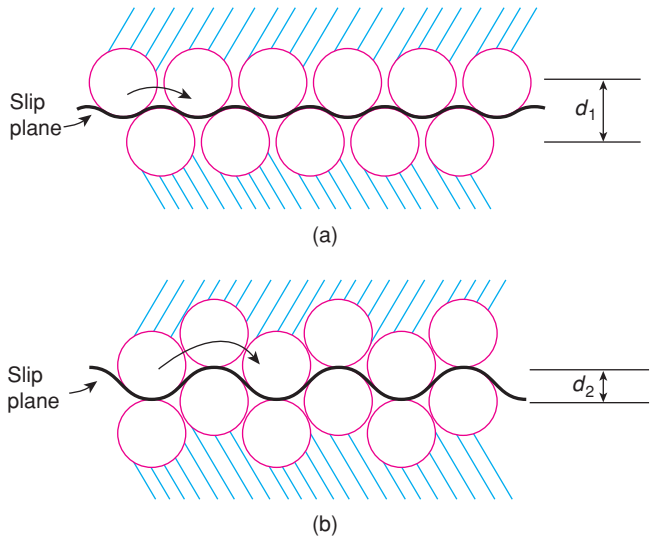


Figure 2.27 ***Slip occurs more easily if the slip plane is a* cp *plane, as in (a).***

2.5.5 Mechanical properties of metals

Metals which are *fcc* – Cu, Ag, Au, Pt, Pb, Ni, Al, etc. – are generally more malleable and ductile than either *hcp* – Ti, Zr, Be – or *bcc* metals – W, V, Mo, Cr, Fe – although there are notable exceptions much as Mg (*hcp*) and Nb (*bcc*) which are malleable and ductile. Many factors influence malleability and ductility. In part, they depend on the numbers of *cp* planes and directions possessed by a structure.

fcc metals have four sets of *cp* planes which are perpendicular to the body diagonals of the cubic unit cell, Fig. 1.20. Each *cp* layer possesses three *cp* directions, x–x', y–y', z–z', Fig. 1.16. These *cp* directions are the face diagonals of the cube and have indices <110>; there is a total of six *cp* directions.

An *hcp* structure contains only one set of *cp* layers, which is parallel to the basal plane of the unit cell, Fig. 1.21. There are also only three *cp* directions, which are in the plane of the *cp* layers. A *bcc* unit cell does not contain *cp* layers. The atomic coordination number in *bcc* is eight, Fig. 1.11(c), whereas *cp* structures have a coordination number of 12. The *bcc* structure does, however, contain four *cp* directions; these correspond to the four body diagonals of the cube, i.e. <111>.

The behaviour of a metal under stress depends on the direction of the applied stress relative to the direction and orientation of the slip directions and slip planes. The effect of tensile stress on a single crystal rod of Mg metal, oriented such that the basal plane of the hexagonal unit cell is at 45° to the rod axis, is shown in Fig. 2.28. For stresses below ~100 psi, the crystal undergoes *elastic deformation*, i.e. no permanent elongation occurs. Above the *yield point*, ~100 psi, *plastic flow* occurs and the crystal suffers an irreversible elongation. Mg metal is remarkable in that it can be stretched to several times its original length. How this occurs is shown in Fig. 2.28(c). Slip occurs on a massive scale and the resulting slip steps are so large that they can be observed directly with an optical microscope. If the passage of a dislocation gives rise to a slip step that is, say, 2 Å high (or wide), then in order that such features are visible in the microscope the slip steps must have dimensions of at least 2 μm. At least 10 000 dislocations must therefore pass on the slip plane for every slip step that can be seen.

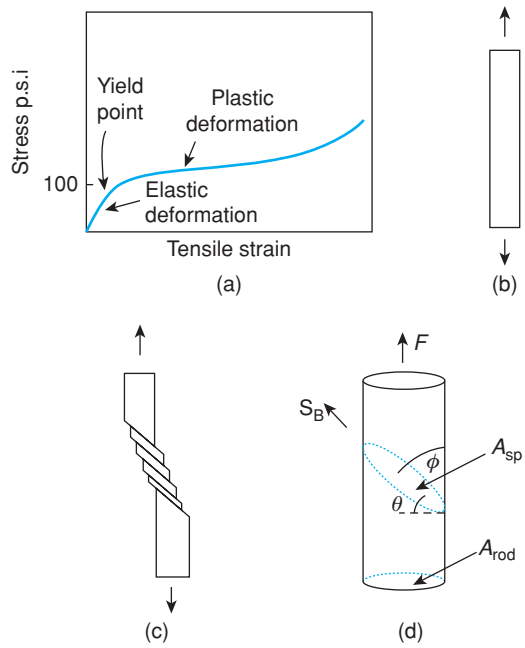


Figure 2.28 ***(a) Tensile stress–strain curve for single-crystal Mg; (b), (c) tensile stress resulting in slip and elongation; (d) definition of the resolved shear stress.***

If the basal planes of Mg are oriented at other than 45° to the rod axis, a larger applied stress is needed for plastic flow to occur. This is because the important factor is the magnitude of the applied shear stress when resolved parallel to the basal plane. For an applied force, F, the stress S_A, on the slip plane (d) is given by

$$S_A = F/A_{sp} = F(\cos\theta/A_{rod}) \tag{2.24}$$

where A_{rod} is the cross-sectional area of the rod, A_{sp} the area of the slip plane and θ the angle between A_{rod} and A_{sp}. The stress S_A on the slip plane has a component S_B parallel to the direction of slip given by

$$S_B = S_A \cos\phi = (F/A_{rod})\cos\theta\cos\phi \tag{2.25}$$

where ϕ is the angle between the slip direction and the stress axis. S_B is the resolved shear stress and the value which is needed to cause plastic flow is the *critical resolved shear stress*. From the interrelation between θ and φ, the maximum value of S_B occurs when $\theta = \phi = 45°$, i.e.

$$S_B = \tfrac{1}{2}(F/A_{rod}) = \tfrac{1}{2}S \quad \text{for } \theta = \phi = 45° \tag{2.26}$$

where S is the stress applied to the crystal. It follows from equation (2.25) that when the slip plane is either perpendicular or parallel to the direction of applied stress, the resolved shear stress is zero and therefore slip cannot occur.

The ability of *fcc* metals to undergo severe plastic deformation in contrast to *hcp* and especially *bcc* metals can now be explained qualitatively. *Fcc* metals have four *cp* planes and six *cp* directions which are suitable

for slip. For any given applied stress, at least one and usually more of these slip planes and directions is suitably oriented, relative to the direction of applied stress, such that slip can occur. In contrast, hexagonal metals slip easily only when in a certain range of orientations relative to the applied stress. The difference between these groups of metals shows up markedly in the mechanical properties of polycrystalline pieces of metal. Polycrystalline hexagonal metals almost certainly contain grains whose slip planes are either parallel or perpendicular to the direction of applied stress. The amount of plastic deformation that is possible in these is therefore more limited. On the other hand, *fcc* metals are malleable and ductile whether as single crystals or polycrystalline pieces.

The critical resolved shear stresses for *fcc* metals are usually small, e.g. Cu 92, Ag 54, Au 132 and Al 148 psi. Values for the hexagonal metals fall into two groups. One group has similar values to *fcc* metals, e.g. Zn 26, Cd 82, Mg 63 psi. The other group has much higher values, e.g. Be 5700, Ti 16 000 psi (slip in Ti occurs more easily on a non-basal plane, but one which has a *cp* direction; it has a critical resolved shear stress of 7100 psi). The reason for the much higher values for Be, Ti (and Zr) is that their unit cell dimensions show the structures to be somewhat compressed in the *c* direction (perpendicular to the basal plane). An ideal hexagonal unit cell has a *c*:*a* ratio of 1.632; values of *c*:*a* for Zn, Cd and Mg are 1.856, 1.886 and 1.624, respectively, and are similar to or greater than the ideal value. Values for Be, Ti and Zr are 1.586, 1.588 and 1.590, respectively, and are *less* than the ideal ratio. Hence it seems that this decrease in distance between adjacent basal planes makes it much more difficult for basal slip to occur, Fig. 2.27(a).

The critical resolved shear stress for *bcc* metals is also high, e.g. α-Fe $\simeq$ 4000 psi, because, although they possess four *cp* directions, they do not have any *cp* planes.

2.5.6 Dislocations, vacancies and stacking faults

There is a close relation between dislocations, point defects (vacancies) and plane defects (stacking faults), as shown by the following examples.

The process of *climb* is a special mechanism of dislocation motion which involves vacancies, Fig. 2.29. Suppose that near the line of an edge dislocation, which is an extra half-plane of atoms, there is a vacancy. If an atom can move from the dislocation line into this vacancy, then effectively, the extra half-plane has become one atom shorter at that point. If this process is repeated along the length of the dislocation line, then the half-plane effectively starts to climb out of the crystal.

An intimate structural relation exists between dislocations and stacking faults. Two adjacent layers of an *fcc* metal are shown in projection in Fig. 2.30(a). The upper layer (white circles) has a zig-zag row of atoms

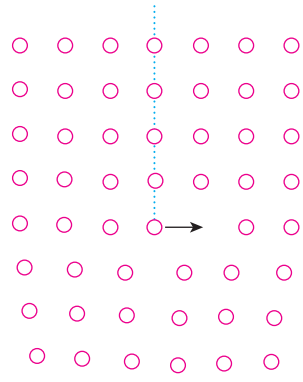


Figure 2.29 ***Climb of an edge dislocation by vacancy migration.***

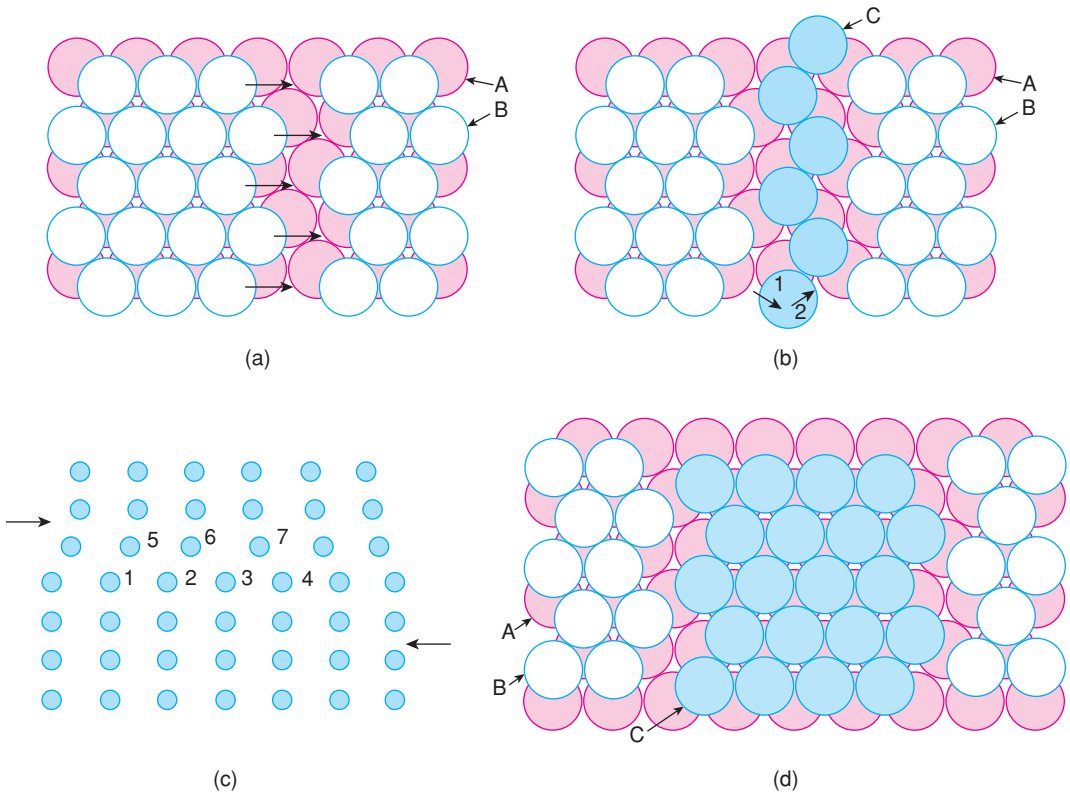


Figure 2.30 ***(a) An edge dislocation and (b) a partial dislocation in an fcc metal; (c) the edge dislocation in projection at atom 3. (d) Separation of two partial dislocations to give a stacking fault.***

missing. This may alternatively be regarded as an extra row of atoms or a negative edge dislocation in the lower layer (shaded circles), immediately below the row of missing atoms in the upper layer. Looked at from the latter point of view, the presence of the edge dislocation in the lower layer effectively forces apart the atoms in the upper layer in the vicinity of the dislocation line, to the extent that a row of atoms in it appears to be missing. This same effect is shown from a different angle in (c). The row containing atoms 1–4 corresponds to the lower layer in (a) and that containing atoms 5–7 to the upper layer. The negative edge dislocation is represented by atom 3 in (c). Motion of the edge dislocation is shown arrowed in (a) and (c) and involves one row of white circles in the upper layer moving to occupy the row of missing atoms. This movement to the right by one unit effectively causes the negative edge dislocation to move to the left by one unit. The layer stacking sequence is unchanged by passage of the dislocation, since atoms in the upper layer occupy B sites before and after slip. The Burgers vector is given by the direction of the arrows and has magnitude equal to one atomic diameter.

The atomic motions arrowed in Fig. 2.30(a) are relatively difficult because atoms in the upper layer virtually have to climb over the top of atoms in the layer below in order to follow the direction indicated by the arrows. A much easier alternative route is shown in (b), in which the migration is divided into two smaller steps (arrows 1 and 2). For each of these, the B atoms traverse a low pass between two A atoms; after the first pass (arrow 1), the atoms find themselves in a C position; a single row of atoms in C positions is shown in (b). This process is repeated (arrow 2) and after the second pass the atoms enter the adjacent set of B positions.

In this manner, a dislocation is divided into two *partial dislocations*. The formation of partial dislocations can be regarded from two viewpoints. In terms of the energy barriers to be surmounted, two small passes are easier than one big hump, even though the combined distance is longer. In terms of the Burgers vector for the two alternatives, the longer, divided pathway is favoured, provided that the distance is not too much longer. Thus, if the direct jump is across a distance b, the energy E_1 is given by $E_1 \propto |b|^2$. For each partial jump, the distance involved is $b/\sqrt{3}$, because the direction is at 30° to the horizontal. Thus for the two partial jumps the combined energy E_2 is proportional to $(b/\sqrt{3})^2 + (b/\sqrt{3})^2$, i.e.

$$E_2 \propto \tfrac{2}{3}|b|^2$$

Therefore,

$$E_2 < E_1$$

Separation of two partial dislocations may occur as in Fig. 2.30d, due to their mutual repulsion. In between the two partial dislocations, the stacking sequence is 'wrong' because all the white atoms occupy C positions; we now have the makings of a *stacking fault*. If the two partial dislocations can be encouraged to go to the edges of the crystal, the whole layer will be faulty. Whether or not this happens in practice, a clear relationship exists between dislocations, partial dislocations and stacking faults.

One suggested way by which dislocations are generated involves the clustering of vacancies onto a certain plane in a structure followed by an inward collapse of the structure, thereby generating a dislocation loop. This mechanism would be important at high temperatures where vacancies form and are relatively mobile. If, in a pure metal, vacancies cluster onto a certain plane in the structure such that every atomic site within, say, a certain radius is empty, then a disc-shaped hole would be present in the structure. The two sides of the disc would cave in, as shown in cross-section in Fig. 2.31, and it can be seen that the two opposite ends have the appearance of edge dislocations of opposite sign. This therefore gives a means of generating dislocations that has nothing to do with mechanical stresses.

2.5.7 Dislocations and grain boundaries

An elegant relationship exists between dislocations and grain boundaries such that the interface between two grains in a polycrystalline material may be regarded as a dislocation network, provided that the angular difference in orientation of the grains is not large. In Fig. 2.32 are shown six positive edge dislocations at different heights in a crystal. With each one, the top half of the crystal becomes slightly wider than the bottom half and the effect of having several such dislocations in a similar orientation and position is to introduce an angular misorientation between the two halves of the crystal. Thus the surface perpendicular to the paper on which these dislocations terminate can be regarded as an interface or boundary between the left-hand and right-hand grains, although the continuity in the structures of the crystals across this boundary is generally

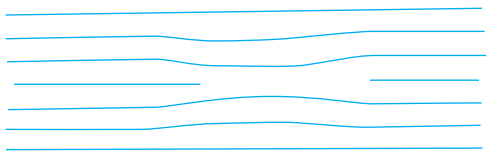


Figure 2.31 ***Collapse of the structure around a cluster of vacancies, thereby generating a dislocation loop.***

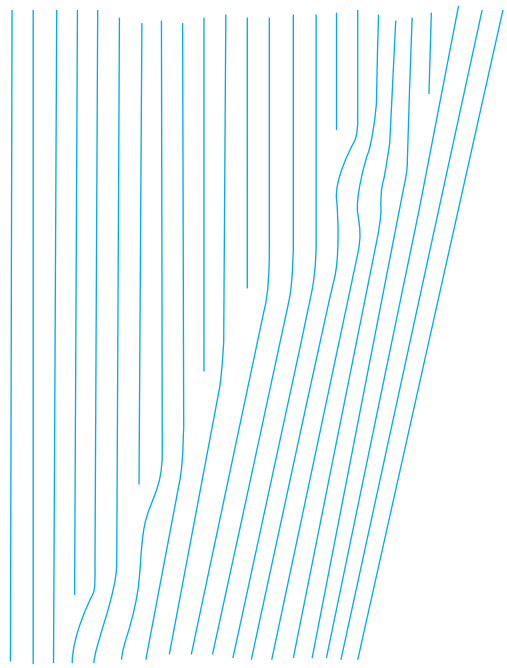


Figure 2.32 ***Array of edge dislocations at a low angle grain boundary.***

excellent. This type of interface is known as a *low angle grain boundary*. The mosaic or domain texture of single crystals, Fig. 2.18, can often be described in terms of low angle grain boundaries.

The angular misorientation of two grains can be calculated readily. Suppose a piece of crystal 1 μm (i.e. 10 000 Å) wide has five positive edge dislocations forming a low angle grain boundary; one end of the crystal will be five atomic planes, say 10 Å, wider than the other end. The angular misorientation of the grains, $\theta = \tan^{-1}(10/10\,000) = 0.057^{\circ}$. Low angle grain boundaries can sometimes be revealed by appropriate etching of the surface, although for the pits to be seen in isolation in the optical microscope they must be at least 1–2 μm in size and separated by the same distance. The dislocation model for grain boundaries is useful for angular misorientations of up to several degrees. For misorientations > 10–20°, however, the model is probably no longer meaningful because the edge dislocations would have to be spaced too closely and would lose their separate identity.

3

Bonding in Solids

3.1 Overview: Ionic, Covalent, Metallic, van der Waals and Hydrogen Bonding in Solids

Crystalline materials exhibit the complete spectrum of bond types, ranging from ionic to covalent, van der Waals and metallic. Sometimes more than one type of bonding is present, as in salts of complex anions, e.g. Li_2SO_4, which have both ionic bonds between Li^+ and SO_4^{2-} ions and covalent bonds linking sulfur and oxygen atoms within SO_4^{2-} ions. Commonly, the bonds are a blend of different types, as in TiO, which is ionic/metallic, and CdI_2, which is ionic/covalent/van der Waals. In discussing structures, it is often convenient to ignore temporarily the complexities of mixed bond types and to treat bonds as though they were purely ionic.

Ionic bonding leads to structures with high symmetry in which the coordination numbers are as high as possible. In this way, the net electrostatic attractive force which holds crystals together (and hence the lattice energy) is maximised. *Covalent bonding*, by contrast, gives highly directional bonds in which atoms prefer a certain coordination environment, irrespective of other atoms that are present. The coordination numbers in covalent structures are usually small and may be less than those in ionic structures which contain similar-sized atoms.

The bonding in a particular compound correlates fairly well with the position of the component atoms in the periodic table and, especially, with their electronegativity. Alkali and alkaline earth elements usually form ionic structures (Be is sometimes an exception), especially in combination with small electronegative anions such as O^{2-} and F^-. Covalent structures occur especially with (a) small atoms of high valency which, in the cationic state, would be highly polarising, e.g. B^{3+}, Si^{4+}, P^{5+} and S^{6+} and to a lesser extent with (b) large atoms which in the anionic state are highly polarisable, e.g. I^- and S^{2-}. Most non-molecular materials have mixed ionic and covalent bonding and, as discussed later, it is possible to assess the *ionicity* of a particular bond, i.e. the percentage of ionic character in the bond.

An additional factor, especially in some transition metal compounds, is the occurrence of *metallic bonding*. In metallic bonding, the outermost valence electrons are not localised on ions or in covalent bonds between atoms but are delocalised in energy bands which cover the entire structure. Most elements of the periodic table are metallic; this includes all *s*-, *d*-, *f*- and many *p*-block elements. However, of great interest in solid state chemistry is that many compounds, including oxides, are also metallic. These include tungsten bronzes such as Na_xWO_3 and high-T_c ceramic superconductors such as $YBa_2Cu_3O_7$. In these, variable valence transition

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

metals such as $W^{5+,\,6+}$ and $Cu^{2+,\,3+}$ are responsible for the metallic bonding, but ionic bonding is also present, especially involving ions such as Na^+, Y^{3+} and Ba^{2+}.

Van der Waals bonding is a feature of molecular materials in which strong covalent bonds hold the molecules together and weak van der Waals bonds are responsible for intermolecular interactions and the fact that molecules such as N_2 are able to liquify and solidify at low temperatures. Without van der Waals bonds, gaseous elements such as He would not be able to liquify since there are no other bond types present. Van der Waals bonds are generally weak, which is why molecular materials have low melting and boiling points.

Hydrogen bonds, such as between H_2O molecules, are rather stronger since these involve ionic or polar interactions between partially charged $H^{\delta+}$ and $O^{\delta-}$ atoms on adjacent molecules. H_2O molecules are in a permanently dipolar state because of electronegativity differences between H and O, whereas van der Waals bonds between non-polar molecules are attributed to weak, spontaneous polarity of molecules such as N_2 caused by instantaneous fluctuations in the charge clouds of the valence electrons.

3.2 Ionic Bonding

Purely ionic bonding rarely occurs. Even structures that are regarded as classically ionic, e.g. NaCl and CaO, usually have some partial covalent character. The degree of covalent bonding increases with increasing valence and ions with a *net* charge greater than $+1$ or -1 are unlikely to exist. Thus, while NaCl may reasonably be represented as Na^+Cl^-, TiC (which also has the NaCl structure) certainly does not contain Ti^{4+} and C^{4-} ions and the main bonding type in TiC must be non-ionic. This brings us to a dilemma. Do we continue to use the ionic model in the knowledge that for many structures, e.g. Al_2O_3 and $CdCl_2$, a large degree of covalent bonding must be present? If not, we must find an alternative model for the bonding. In this section, ionic bonding is given prominence because of its apparent wide applicability and its usefulness as a *starting point* for describing structures which in reality often have considerable covalent bonding.

3.2.1 Ions and ionic radii

It is difficult to imagine discussing crystal structures without having definitive information on the sizes of ions in crystals. However, crystal chemistry has undergone a minor revolution because the long-established tables of ionic radii of Pauling, Goldschmidt and others are now thought to be seriously in error; at the same time, our concepts of ions and ionic structures have undergone revision. In the most recent comprehensive compilations of ionic radii, of Shannon and Prewitt (1969, 1970), cations are larger and anions smaller than previously thought. For example, Pauling radii of Na^+ and F^- are 0.98 and 1.36 Å, respectively, whereas Shannon and Prewitt give values of 1.14–1.30 Å, depending on the coordination number, for Na^+ and 1.19 Å for F^-.

These changes have arisen because high-quality X-ray crystallography gives fairly accurate maps of the distribution of electron density throughout crystals. Thus, one can effectively 'see' ions and tell something about their size, shape and nature. Figure 3.1 shows an electron density 'contour map' of LiF for a section passing through the structure parallel to one unit cell face. The map passes through the centres of Li^+ and F^- ions located on (100) planes. Figure 3.2 shows the variation of electron density with distance along the line that connects adjacent Li^+ and F^- ions. From Fig. 3.1 and Fig. 3.2 and similar diagrams for other structures, the following conclusions about ions in crystals may be drawn:

(a) Ions are essentially spherical.
(b) Ions may be regarded as composed of two parts: a central core in which most of the electron density is concentrated and an outer sphere of influence which contains very little electron density.

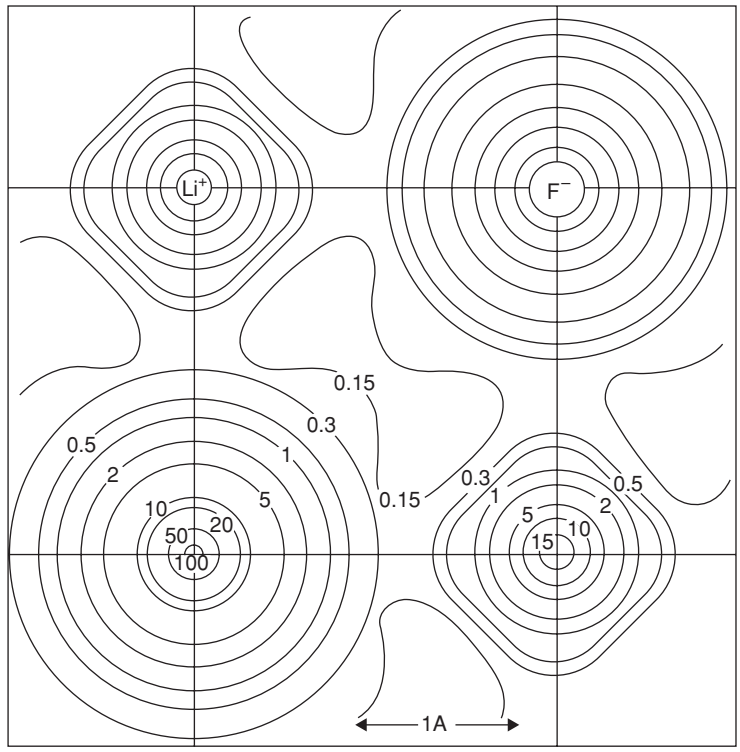


Figure 3.1 *Electron density contour map of LiF (rock salt structure): a section through part of the unit cell face. The electron density (electrons Å^{-3}) is constant along each of the contour lines. Reproduced with permission from Krug, Witte and Wolfel,* **Zeit. Phys. Chem., Frankfurt**, 4, *36, © 1955 Oldenbourg Verlag.*

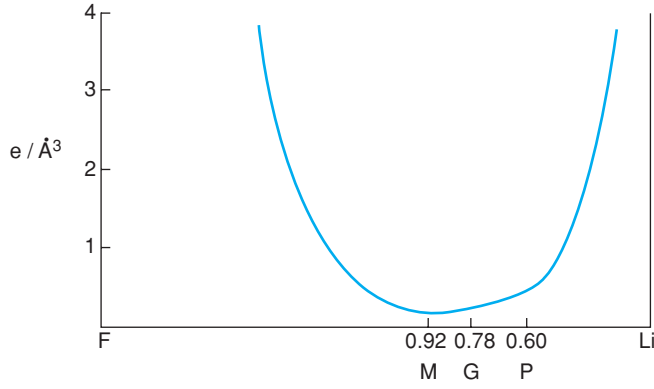


Figure 3.2 *Variation of electron density along the line connecting adjacent Li and F nuclei in LiF. P = Pauling radius of Li$^+$; G = Goldschmidt radius; M = minimum in electron density. Reproduced with permission from Krebs,* **Fundamentals of Inorganic Crystal Chemistry**, *© 1968 McGraw-Hill.*

(c) Assignment of radii to ions is difficult; even for ions which are supposedly in contact, it is not obvious (Fig. 3.2) where one ion ends and another begins.

Conclusion (b) is in contrast to the oft-stated assumption that 'ions are charged, incompressible, non-polarisable spheres'. Certainly, ions are charged, but they cannot be regarded as hard spheres with a clearly defined radius. Their electron density does not decrease abruptly to zero at a certain distance from the nucleus but instead decreases roughly exponentially with increasing radius. Instead of being incompressible, ions are fairly elastic, by virtue of flexibility in the outer sphere of influence of an ion while the inner core remains unchanged. This flexibility is necessary in order to explain variations of apparent ionic radii with coordination number and environment (see later). Within limits, ions can expand or contract as the situation demands.

From Fig. 3.1 and Fig. 3.2, most electron density is concentrated close to the nuclei of ions; in a crystal, therefore, most of the total volume is essentially free space and contains relatively little electron density. It is not completely ludicrous to draw an analogy between the structure of an atom (or ion) and that of our solar system. In both, most of the matter is concentrated in relatively small regions and the rest of the volume is made up of essentially empty space.

It is difficult to specify ionic radii exactly because, between adjacent anions and cations, the electron density passes through a broad, shallow minimum. For LiF, Fig. 3.2, the radii for Li^+ given by Pauling and Goldschmidt are marked together with the value corresponding to the minimum in the electron density along the line connecting Li^+ and F^-. Although the values of these radii vary from 0.60 to 0.92 Å, all lie within the broad electron density minimum of Fig. 3.2.

In spite of the difficulties in determining absolute radii, it is necessary to have a set of radii for reference. Fortunately, most sets of radii are *additive* and *self-consistent*; provided that one does not mix radii from different tabulations, it is possible to use any set of radii to evaluate interionic distances in crystals with reasonable confidence. Shannon and Prewitt give two sets of radii: one is based on $r_{O^{2-}} = 1.40$ Å and is similar to Pauling, Goldschmidt, etc.; the other is based on $r_{F^-} = 1.19$ Å (and $r_{O^{2-}} = 1.26$ Å) and is related to the values determined from X-ray electron density maps. Both sets are comprehensive for cations in different coordination environments but pertain only to oxides and fluorides. We use here the set based on $r_{F^-} = 1.19$ Å (and $r_{O^{2-}} = 1.26$ Å). Radii for some ions M^+ to M^{4+} are shown in Fig. 3.3 as a function of cation coordination number; it should be stressed that the more highly charged ions are unlikely to exist as such since their positive charge is reduced by polarisation of the anion and consequent partial covalent bonding between cation and anion.

As an alternative to using tabulations of ionic radii, typical bond distances for each cation may be used. These are summarised in Appendix F. For instance, the Ti–O bond in TiO_2 and a wide range of metal titanates almost always lies in the range 1.94–1.96 Å when the Ti occupies regular, undistorted TiO_6 octahedra. A useful way to compare the sizes of ions such as Ti^{4+}, Zn^{2+} and Al^{3+} is therefore to compare their typical bond lengths to oxygen.

The following trends in ionic radii, with position in the Periodic Table, formal charge and coordination number, occur:

(a) For the s- and p-block elements, radii increase with atomic number for any vertical group, e.g. octahedrally coordinated alkali metal ions.
(b) For any isoelectronic series of cations, radii decrease with increasing charge, e.g. Na^+, Mg^{2+}, Al^{3+} and Si^{4+}.
(c) For any element which can have more than one oxidation state, the cation radius decreases with increasing oxidation state, e.g. $V^{2+} > V^3 > V^{4+} > V^{5+}$.
(d) For an element which can have various coordination numbers, the cationic radius increases with increasing coordination number.

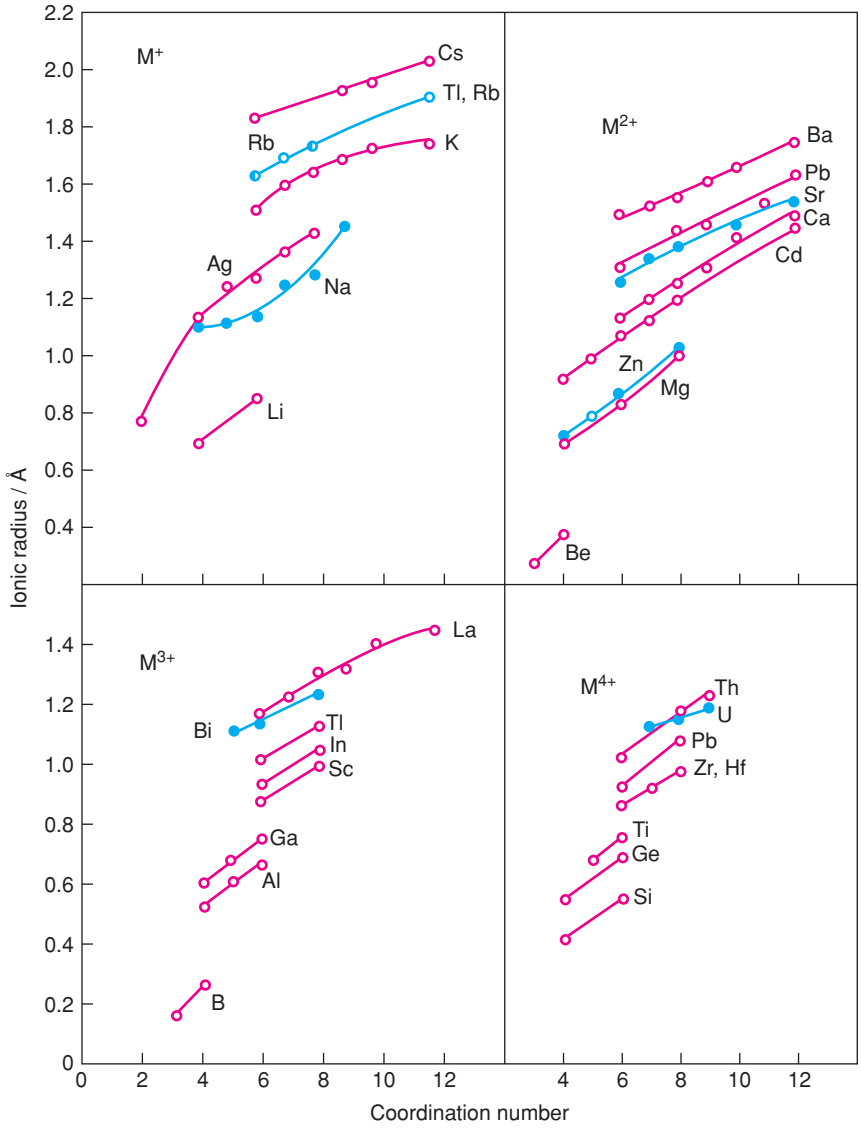


Figure 3.3 ***Ionic radii as a function of coordination number for cations M^+ to M^{4+}. Data based on $r_{F^-} = 1.19$ Å and $r_{O^{2-}} = 1.26$ Å. Reproduced with permission from Shannon and Prewitt,*** **Acta Cryst.,** B25, ***925, © 1970 International Union of Crystallography.***

(e) In the 'lanthanide contraction', ions with the same charge but increasing atomic number show a reduction in size (due to ineffective shielding of the nuclear charge by the d and, especially, f electrons), e.g. octahedral radii, La^{3+} (1.20 Å) ... Eu^{3+} (1.09 Å) ... Lu^{3+} (0.99 Å). Similar effects occur across some series of transition metal ions.
(f) The radius of a particular transition metal ion is smaller than that of the corresponding main group ion for the reasons given in (e), e.g. octahedral radii, Rb^{+} (1.63 Å) and Ag^{+} (1.29 Å) or Ca^{2+} (1.14 Å) and Zn^{2+} (0.89 Å).
(g) Certain pairs of elements positioned diagonally to one another in the Periodic Table have similar cationic size (and chemistry), e.g. Li^{+} (0.88 Å) and Mg^{2+} (0.86 Å). This is due to a combination of effects (a) and (b).

3.2.2 Ionic structures – general principles

Consider the following as a guide to ionic structures:

(a) Ions are charged, elastic and polarisable spheres.
(b) Ionic structures are held together by electrostatic forces and are arranged so that cations are surrounded by anions, and vice versa.
(c) In order to maximise the net electrostatic attraction between ions (i.e. the lattice energy), coordination numbers are as high as possible, provided that the central ion 'maintains contact' with its neighbouring ions of opposite charge.
(d) Next nearest neighbour interactions are of the anion–anion and cation–cation type and are repulsive. Similarly charged ions arrange themselves to be as far apart as possible, which leads to structures of high symmetry with a maximised volume.
(e) Local electroneutrality prevails; the valence of an ion is equal to the sum of the electrostatic bond strengths between it and adjacent ions of opposite charge.

Point (a) was considered in Section 3.2; ions are obviously charged, are elastic because their size varies with coordination number and are polarisable when departures from purely ionic bonding occur. Thus, the electron density map for LiF, Fig. 3.1, shows a small distortion from spherical shape of the Li^{+} ion, attributable to a small amount of covalent bonding between Li^{+} and F^{-}.

Points (b)–(d) imply that the forces which hold ionic crystals together are the same as would be obtained by regarding the crystal as a 3D array of point charges and considering the net coulombic energy of the array. From Coulomb's law, the force F between two ions of charge $Z_{+}e$ and $Z_{-}e$, separated by distance r, is given by

$$F = \frac{(Z_{+}e)(Z_{-}e)}{r^{2}} \tag{3.1}$$

A similar equation applies to each pair of ions in the crystal and evaluation of the resulting force between all the ions leads to the lattice energy of the crystal (see later).

Point (c) includes the proviso that nearest neighbour ions should be 'in contact'. Given the nature of electron density distributions in ionic crystals, it is difficult to quantify what is meant by 'in contact'. It is nevertheless important since, although the apparent size of ions varies with coordination number, most ions, smaller ones especially, appear to have a maximum coordination number; for Be^{2+} this is four and for Li^{+} it is six. Ions are flexible, therefore, but within fairly narrow limits.

The idea of *maximising* the volume of ionic crystals, point (d), is rather unexpected since one is accustomed to regarding ionic structures and derivative *cp* structures, especially, as having *minimum* volume. There is no conflict, however. The prime bonding force in ionic crystals is the nearest neighbour cation–anion *attractive* force and this force is maximised at a small cation–anion separation (when ions become too close, additional repulsive forces come into play, thereby reducing the net attractive force). Superposed on this is the effect of next nearest neighbour *repulsive* forces between like ions. With the constraints that (a) cation–anion distances are minimised and (b) coordination numbers are maximised, like ions arrange themselves to be as far apart as possible in order to reduce their mutual repulsion. This leads to regular and highly symmetrical arrays of ions which tend to have maximised volumes.

An excellent example of a structure whose volume is maximised is rutile, TiO_2 (see Section 1.17.6). The oxide layers are buckled compared with those in an ideal *hcp* array, which causes the next nearest neighbour coordination number of O (by O) to be reduced from 12 (as in *hcp*) to 11 (as in *tp*). The coordination of Ti by O, and vice versa, is unaffected by this distortion but the overall volume of the structure increases by 2–3%. Hence the oxide layers buckle so that the volume can be maximised.

Point (e) is Pauling's *electrostatic valence rule*, the second of a set of rules formulated by Pauling for ionic crystals. Basically, the rule means that the charge on a particular ion, e.g. an anion, must be balanced by an equal and opposite charge on the immediately surrounding cations. However, since these cations are also shared with other anions, it is necessary to estimate the amount of positive charge that is effectively associated with each cation–anion bond. For a cation M^{m+} surrounded by n anions, X^{x-}, the *electrostatic bond strength* (*ebs*) of each cation–anion bond is defined as

$$ebs = \frac{m}{n} \tag{3.2}$$

For each anion, the sum of the *ebs* of the surrounding cations must balance the charge on the anion, i.e.

$$\sum \frac{m}{n} = x \tag{3.3}$$

For example:

(a) Spinel, $MgAl_2O_4$, contains octahedral Al^{3+} and tetrahedral Mg^{2+} ions; each O is surrounded tetrahedrally by three Al^{3+} ions and one Mg^{2+} ion. We can check that this must be so, as follows:

$$\text{For } Mg^{2+} : \qquad ebs = \frac{2}{4} = \frac{1}{2}$$

$$\text{For } Al^{3+} : \qquad ebs = \frac{3}{6} = \frac{1}{2}$$

Therefore,

$$\sum ebs\ (3Al^{3+} + 1Mg^{2+}) = 2$$

(b) We can show that three SiO_4 tetrahedra *cannot* share a common corner in silicate structures:

$$\text{For } Si^{4+} : \qquad ebs = \frac{4}{4} = 1$$

Table 3.1 ***Electrostatic bond strengths of some cations***

Cation	Coordination number(s)	*ebs*
Li^{+}	4, 6	1/4, 1/6
Na^{+}	6, 8	1/6, 1/8
Be^{2+}	3, 4	2/3, 1/2
Mg^{2+}	4, 6	1/2, 1/3
Ca^{2+}	8	1/4
Zn^{2+}	4	1/2
Al^{3+}	4, 6	3/4, 1/2
Cr^{3+}	6	1/2
Si^{4+}	4	1
Ge^{4+}	4, 6	1, 2/3
Ti^{4+}	6	2/3
Th^{4+}	8	1/2

Therefore, for an O that bridges two SiO_4 tetrahedra, $\Sigma ebs = 2$, which is, of course, acceptable. However, three tetrahedra sharing a common O would give $\Sigma ebs = 3$ for that O, which is completely unacceptable.

This rule of Pauling provides an important guide to the polyhedral linkages that are and are not possible in crystal structures. Table 3.1 lists some common cations with their formal charge, coordination number, *CN*, and *ebs*. Table 3.2 lists some allowed and unallowed combinations of polyhedra about a common oxide ion. Many other combinations are possible and the reader may like to deduce some, bearing in mind that there are also topological restrictions on the number of possible polyhedral combinations; thus the maximum number of octahedra that can share a common corner is six (as in rock salt), etc.

Pauling's third rule concerns the topology of polyhedra and was considered in Chapter 1. Pauling's first rule states: 'A coordinated polyhedron of anions is formed about each cation; the cation–anion distance is determined by the radius sum and the *CN* of the cation by the radius ratio'. The idea that cation–anion distances are determined by the radius sum is implicit to every tabulation of ionic radii since a major objective of such tabulations is to be able to predict, correctly, interatomic distances. Let us now consider *CN*s and the *radius ratio rules*.

Table 3.2 ***Allowed and unallowed combinations of corner-sharing oxide polyhedra***

Allowed	Example	Unallowed
$2SiO_4$tet.	Silica	$>2SiO_4$tet.
$1MgO_4$tet. + $3AlO_6$oct.	Spinel	$3AlO_4$tet.
$1SiO_4$tet. + $3MgO_6$oct.	Olivine	$1SiO_4$tet. + $2AlO_4$tet.
$8LiO_4$tet.	Li_2O	$4TiO_6$oct.
$2TiO_6$oct. + $4CaO_{12}$dod.	Perovskite	
$3TiO_6$oct.	Rutile	

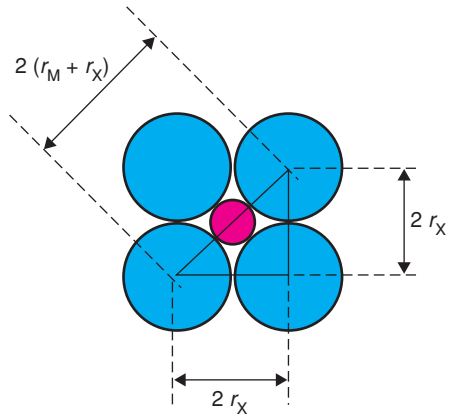


Figure 3.4 ***Radius ratio calculation for octahedral coordination.***

3.2.3 The radius ratio rules

In ideally ionic structures, the *CN*s of the ions are determined by electrostatic considerations. Cations surround themselves with as many anions as possible, and vice versa. This maximises the attractions between neighbouring ions of opposite charge and hence maximises the lattice energy of the crystal (see later). This requirement led to the formulation of the *radius ratio rule* for *ionic structures* in which the structure adopted for a particular compound depends on the *relative* sizes of the ions. This rule is useful qualitatively but, as we shall see, cannot usually be applied rigorously to explain or predict.

There are two guidelines to be followed in using radius ratio rules:

1. A cation must be in contact with its anionic neighbours.
2. Neighbouring anions may or may not be in contact.

Guideline 1 places a *lower* limit on the size of cation which may occupy a particular site, since a situation in which a cation may 'rattle' inside its anion polyhedron is assumed to be unstable. Using these guidelines, one may calculate the range of cation sizes that can occupy the various interstitial sites in an anion array. If you have worked systematically through problems 1.32 to 1.35, then the following section should be a revision.

Let us first calculate the minimum radius for an octahedral cation site, in which both anion–anion and anion–cation contacts occur. The cation is in contact with all six anion neighbours; Fig. 3.4 shows a section through an octahedron, in which the central cation has four coplanar anion neighbours. Applying the Pythagoras theorem:

$$(2r_x)^2 + (2r_x)^2 = [2(r_M + r_x)]^2 \tag{3.4}$$

i.e.

$$2r_x\sqrt{2} = 2(r_M + r_x)$$

and

$$r_M/r_x = \sqrt{2} - 1 = 0.414 \tag{3.5}$$

For radius ratios <0.414, the cation is too small for an octahedral site and, instead, should occupy a site of smaller *CN*. For radius ratios >0.414, the cation would push the anions apart, and this happens increasingly up to a radius ratio of 0.732. At this value, the cation can have eight anionic neighbours, all of which are in contact with the cation.

To calculate the minimum radius ratio for eight coordination, it is necessary to recognise that cations and anions are in contact along the body diagonal of the cube as in CsCl, Fig. 1.36, i.e.

$$2(r_M + r_x) = \text{(cube body diagonal)} \tag{3.6}$$

In addition, anions are in contact along the cell edge, a, i.e.

$$a = 2r_x \tag{3.7}$$

Therefore,

$$2(r_M + r_x) = 2r_x\sqrt{3}$$

and

$$r_M/r_x = \sqrt{3} - 1 = 0.732 \tag{3.8}$$

To calculate the minimum radius ratio for tetrahedral coordination, we may regard a tetrahedron as a cube with alternate corners missing, Fig. 1.25. Anions and cations are in contact along the cube *body* diagonals, equation (3.6), but anions are in contact along the cube *face* diagonals, i.e.

$$2r_x = \text{(face diagonal)} \tag{3.9}$$

Combining equations (3.6) and (3.9) and rearranging gives

$$(2r_x)^2 + (\sqrt{2}r_x)^2 = [2(r_M + r_x)]^2$$

i.e.

$$2(r_M + r_x) = \sqrt{6}x_x$$

and

$$r_M/r_x = \frac{\sqrt{6} - 2}{2} = 0.225 \tag{3.10}$$

The minimum radius ratios for various *CN*s are given in Table 3.3. Note that $CN = 5$ is absent from the table; in *cp* structures, it is not possible to have a *CN* of 5 in which all M–X bonds are of the same length.

The radius ratio rules are successful in predicting trends in *CN* and structure type but can be used only as a qualitative guide. Radius ratios depend very much on which table of ionic radii is consulted and there appears to be no clear advantage in using either one of the more traditional sets or the modern set of values based on X-ray diffraction results. For example, for RbI, $r^+/r^- = 0.69$ or 0.80, according to the tables based on $r_{O^{2-}} = 1.40$ and 1.26 Å, respectively. Thus one value would predict six coordination (rock salt), as observed,

Table 3.3 ***Minimum radius ratios***

Coordination	Minimum r_M:r_x
Linear, 2	–
Trigonal, 3	0.155
Tetrahedral, 4	0.225
Octahedral, 6	0.414
Cubic, 8	0.732
Dodecahedral, 12	1.000

but the other predicts eight coordination (CsCl). On the other hand, LiI has $r^+/r^- = 0.28$ and 0.46, according to the same tables; one value predicts tetrahedral coordination and the other octahedral (as observed). For the larger cations, especially caesium, $r^+/r^- > 1$, and it may be then more realistic to consider the inverse ratio, r^-/r^+, in, for instance, CsF.

A more convincing example of the relevance of radius ratios is provided by oxides and fluorides of general formula MX_2. Possible structure types, with their cation *CN*s, are silica (4), rutile (6) and fluorite (8). A selection of oxides in each group is given in Table 3.4, together with the radius ratios calculated from Fig. 3.3 (based on $r_{O^{2-}} = 1.26$ Å). Changes in *CN* are expected at radius ratios of 0.225, 0.414 and 0.732. Bearing in mind that the calculated ratios depend on the table of radii that is used, the agreement between theory and practice is reasonable. For example, GeO_2 is polymorphic and has both silica and rutile structures; the ratio calculated for tetrahedral coordination of Ge is borderline between the values predicted for $CN = 4$ and 6.

3.2.4 Borderline radius ratios and distorted structures

The structural transition from $CN = 4$ to 6 which occurs with increasing cation size is often clear-cut. Thus GeO_2 has a borderline radius ratio and also exhibits polymorphism. Both polymorphs have highly symmetric structures with $CN = 4$ and 6, respectively. Polymorphs with $CN = 5$ do not occur with GeO_2.

Table 3.4 ***Structures and radius ratios of oxides MO_2***

Oxide	Calculated radius ratio[a]		Observed structure type	
CO_2	~0.1	($CN = 2$)	Molecular	($CN = 2$)
SiO_2	0.32	($CN = 4$)	Silica	($CN = 4$)
GeO_2	0.43	($CN = 4$)	Silica	($CN = 4$)
	0.54	($CN = 6$)	Rutile	($CN = 6$)
TiO_2	0.59	($CN = 6$)	Rutile	($CN = 6$)
SnO_2	0.66	($CN = 6$)	Rutile	($CN = 6$)
PbO_2	0.73	($CN = 6$)	Rutile	($CN = 6$)
HfO_2	0.68	($CN = 6$)	Fluorite	($CN = 8$)
	0.77	($CN = 8$)		
CeO_2	0.75	($CN = 6$)	Fluorite	($CN = 8$)
	0.88	($CN = 8$)		
ThO_2	0.95	($CN = 8$)	Fluorite	($CN = 8$)

[a]Since cation radii vary with *CN*, Fig. 3.3, radius ratios may be calculated for different *CN*s. The *CN*s used here are given in parentheses. Calculations are based on $r_{O^{2-}} = 1.26$ Å.

In other borderline cases, however, distorted polyhedra and/or *CN*s of 5 are observed. Thus V^{5+} (radius ratio = 0.39 for $CN = 4$ or 0.54 for $CN = 6$) has an environment in one polymorph of V_2O_5 which is a gross distortion of octahedral; five V–O bonds are of reasonable length, in the range 1.5–2.0 Å, but the sixth is much longer, 2.8 Å, and the coordination is better regarded as distorted square pyramidal. It appears that V^{5+} is rather small to occupy an octahedral site comfortably and, instead, a structure occurs which is transitional between tetrahedral and octahedral.

Similar distortions occur between $CN = 6$ and 8. Thus, ZrO_2 has a borderline radius ratio (0.68 for $CN = 6$; 0.78 for $CN = 8$) and, although it may have the fluorite structure at very high temperatures (>2000 °C), with $CN = 8$ for Zr, in its room temperature form as the mineral baddleyite, $CN = 7$ for Zr.

Less severe distortions occur when a cation is only slightly too small for its anion environment. The regular anion coordination is maintained but the cation may rattle or undergo small displacements within its polyhedron. In $PbTiO_3$ (radius ratio for Ti = 0.59 for $CN = 6$), Ti is displaced by ~0.2 Å off the centre of its octahedral site towards a corner oxygen. The direction of displacement is reversible in an applied electric field, giving rise to the important property of ferroelectricity (see Chapter 8). It is amazing that large sections of the electronics and electroceramics industries depend upon slight structural distortions associated with borderline radius ratios; examples include piezoelectric sensors, high-permittivity dielectrics/capacitor materials and ferroelectric-based memory switching devices.

The concept of 'maximum contact distance' was proposed by Dunitz and Orgel. If the metal–anion distance increases above this distance, then the cation is free to rattle. If the metal–anion distance decreases, the metal ion is subjected to compression. However, the maximum contact distance does not correspond to the sum of ionic radii, as they are usually defined, and this has been a difficult concept to quantify.

3.2.5 Lattice energy of ionic crystals

Ionic crystals may be regarded as regular 3D arrays of point charges. The forces that hold them together are entirely electrostatic and may be calculated by summing all the electrostatic repulsions and attractions in the crystal. The *lattice energy*, U, is the net potential energy of the arrangement of charges that forms the structure. It is equivalent to the energy required to sublime the crystal and convert it into a collection of gaseous ions, e.g.

$$NaCl(s) \rightarrow Na^+(g) + Cl^-(g), \ \Delta H = U$$

The value of U depends on the crystal structure, the charge on the ions and the internuclear separation between anions and cations. Two principal kinds of force determine ionic crystal structures:

(i) Electrostatic forces of attraction and repulsion. Two ions M^{Z+} and X^{Z-} separated by a distance r experience an attractive force, F, given by Coulomb's law:

$$F = \frac{Z_+ Z_- e^2}{r^2} \tag{3.11}$$

Their coulombic potential energy, V, is given by

$$V = \int_{\infty}^{r} F \mathrm{d}r = -\frac{Z_+ Z_- e^2}{r} \tag{3.12}$$

(ii) Short-range repulsive forces which are important when atoms or ions are so close that their electron clouds begin to overlap. Born suggested that this repulsive energy has the form

$$V = \frac{B}{r^n} \tag{3.13}$$

where B is a constant and n is in the range 5–12. Because n is large, V falls rapidly to zero with increasing r.

The lattice energy is calculated by combining the net electrostatic attraction and the Born repulsion energies and finding the internuclear separation, r_e, which gives the maximum U value. The procedure is as follows.

Consider the NaCl structure, Fig. 1.29(a). Between each pair of ions there is an electrostatic interaction given by equation (3.11). We wish to sum all such interactions in the crystal and calculate the net attractive energy. Let us first consider one particular ion, e.g. Na^+ in the body centre of the unit cell, and calculate the interaction between it and its neighbours. Its nearest neighbours are six Cl^- ions in the face centre positions at distance r ($2r$ is the value of the unit cell edge). The attractive potential energy is given by

$$V = -6\frac{e^2 Z_+ Z_-}{r} \tag{3.14}$$

The next nearest neighbours are 12 Na^+ ions at edge centre positions, distance $\sqrt{2}r$; this gives a repulsive potential energy term

$$V = +12\frac{e^2 Z_+ Z_-}{\sqrt{2}r} \tag{3.15}$$

The third nearest neighbours are eight Cl^- ions at the cube corners, distance $\sqrt{3}r$; these are attracted to the central Na^+ ion according to

$$V = -8\frac{e^2 Z_+ Z_-}{\sqrt{3}r} \tag{3.16}$$

The net attractive energy between our Na^+ ion and all other ions in the crystal is given by an infinite series:

$$V = -\frac{e^2 Z_+ Z_-}{r}\left(6 - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \frac{6}{\sqrt{4}}\ldots\right) \tag{3.17}$$

This summation is repeated for each ion in the crystal, i.e. for $2N$ ions per mole of NaCl. Since each ion pair interaction is thereby counted twice, it is necessary to divide the final value by 2, giving

$$V = -\frac{e^2 Z_+ Z_-}{r}NA \tag{3.18}$$

where the *Madelung constant*, A, is the numerical value of the summation in parentheses from equation (3.17). The Madelung constant depends only on the geometric arrangement of point charges. It has the same value,

Table 3.5 **Madelung constants for some simple structures**

Structure type	A
Rock salt	1.748
CsCl	1.763
Wurtzite	1.641
Sphalerite	1.638
Fluorite	2.520
Rutile	2.408

1.748, for all compounds with the rock salt structure. Values of A for some other structures are given in Table 3.5.

If equation (3.18) were the only factor in the lattice energy, the structure would collapse in on itself since $V \propto -1/r$, equation (3.18) and Fig. 3.5. This catastrophe is avoided by the mutual repulsion between ions, of whatever charge, when they become too close and which is given by equation (3.13). The dependence of this repulsive force on r is shown schematically in Fig. 3.5. The total energy of the crystal, U, is given by summing equations (3.18) and (3.13) and differentiating with respect to r to find the maximum U value and equilibrium interatomic distance, r_e; i.e.

$$U = -\frac{e^2 Z_+ Z_- NA}{r} + \frac{BN}{r^n} \tag{3.19}$$

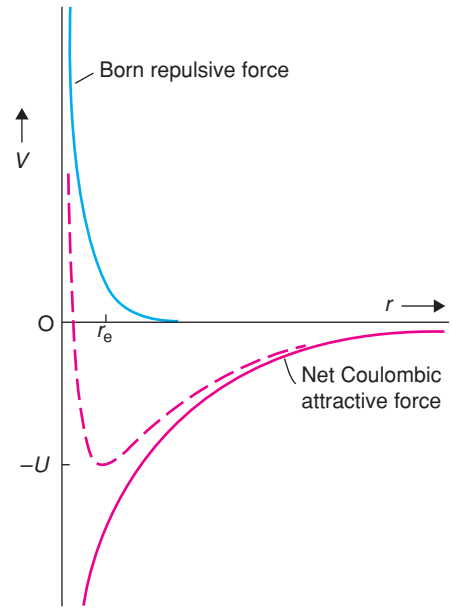


Figure 3.5 **Lattice energy (dashed line) of ionic crystals as a function of internuclear separation.**

Therefore,

$$\frac{\mathrm{d}U}{\mathrm{d}r} = \frac{e^2 Z_+ Z_- NA}{r^2} - \frac{nBN}{r^{n+1}} \tag{3.20}$$

When

$$\frac{\mathrm{d}U}{\mathrm{d}r} = 0 \tag{3.21}$$

then

$$B = \frac{e^2 Z_+ Z_- A r^{n-1}}{n} \tag{3.22}$$

and therefore

$$U = -\frac{e^2 Z_+ Z_- NA}{r_\mathrm{e}} \left(1 - \frac{1}{n}\right) \tag{3.23}$$

The dashed line in Fig. 3.5 shows the variation of U with r and gives the minimum U value when $r = r_\mathrm{e}$.

For most practical purposes, equation (3.23) is entirely satisfactory, but in more refined treatments certain modifications are made:

(a) The Born repulsive term is represented instead by an exponential function:

$$V = B \exp\left(\frac{-r}{\rho}\right) \tag{3.24}$$

where ρ is a constant, typically 0.35. When r is small ($r \ll r_\mathrm{e}$), equations (3.13) and (3.24) give very different values for V, but for realistic interatomic distances, i.e. $r \simeq r_\mathrm{e}$, the two values are similar. Use of equation (3.24) in the expression for U gives the *Born–Mayer equation*:

$$U = -\frac{e^2 Z_+ Z_- AN}{r_\mathrm{e}} \left(1 - \frac{\rho}{r_\mathrm{e}}\right) \tag{3.25}$$

(b) The zero point energy of the crystal is included in the calculation of U. It equals $2.25h\nu_{0,\mathrm{max}}$, where $\nu_{0,\mathrm{max}}$ is the frequency of the highest occupied vibrational mode in the crystal. Its inclusion leads to a small reduction in U.

(c) van der Waals attractive forces exist between ions due to induced dipole–induced dipole interactions. These take the form NC/r^6 and lead to an increase in U.

A more complete equation for U, after correcting for these factors, is

$$U = -\frac{Ae^2 Z_+ Z_- N}{r} + BNe^{-r/\rho} - CNr^{-6} + 2.25Nh\nu_{0,\mathrm{max}} \tag{3.26}$$

Table 3.6 ***Some lattice energies in kJ mol^{-1}; all have rock salt structure***

MgO	−3938	LiF	−1024	NaF	−911
CaO	−3566	LiCl	−861	KF	−815
SrO	−3369	LiBr	−803	RbF	−777
BaO	−3202	LiI	−744	CsF	−748

Source: Reproduced with permission from Ladd and Lee, *Progr. Solid State Chem.*, **1**, 37–82, 1963; **2**, 378–413, © 1965 Elsevier.

Typical values for the terms, in kJ mol^{-1}, for two chemicals with the rock salt structure are as follows (from Greenwood, 1968):

Substance	$NAe^2Z_+Z_-r^{-1}$	$NBe^{-r/\rho}$	NCr^{-6}	$2.25Nh\nu_{0,\max}$	U
NaCl	−859.4	98.6	−12.1	7.1	−765.8
MgO	−4631	698	−6.3	18.4	−3921

Hence the Born repulsive term contributes 10–15% to the value of U; the zero point vibrational and van der Waals terms contribute about 1% each and, being of opposite sign, tend to cancel each other. For most purposes, therefore, we use the simplified equation (3.23). Let us now consider each of the terms in equation (3.23) and evaluate their relative significance.

The magnitude of U depends on six parameters A, N, e, Z, n and r_e, four of which are constant for a particular structure. This leaves just two, the charge on the ions, Z_+Z_-, and the internuclear separation, r_e. Of these two, charge is the most important since the product (Z_+Z_-) is capable of much larger variation than r_e. For instance, U for a material with divalent ions should be four times as large as for an isostructural crystal with the same r_e but containing monovalent ions (compare the lattice energies of SrO and LiCl, Table 3.6, which have similar r_e, Table 1.8: $r_e = a/2$). For isostructural phases with the same Z values but increasing r_e, a decrease in U is expected (e.g. alkali metal fluorides, alkaline earth oxides with NaCl structure). Lattice energies for materials that show these two trends are given in Table 3.6.

Since U of a crystal is equivalent to its heat of dissociation, a correlation exists between U and the melting point (a better correlation may be sought between U and the sublimation energy, but such data are not so readily available). The effect of (Z_+Z_-) on the melting point is shown by the refractoriness of the alkaline earth oxides (m.p. of CaO = 2572 °C) compared with the alkali halides (m.p. of NaCl = 800 °C). The effect of r_e on melting points may be seen in isostructural series such as MgO (2800 °C), CaO (2572 °C) and BaO (1923 °C).

3.2.6 Kapustinskii's equation

Kapustinskii noted an empirical increase in the value of the Madelung constant, A, as the *CN* of the ions increased, e.g. in the series ZnS, NaCl, CsCl, Table 3.5. Since, for a particular anion and cation, r_e also increases with *CN*, Fig. 3.3, Kapustinskii proposed a general equation for U in which variations in A and r_e are auto-compensated. He suggested using the rock salt value for A and octahedral ionic radii (Goldschmidt) in calculating r_e; substituting $r_e = r_c + r_a$, $\rho = 0.345$, $A = 1.745$ and values for N and e into equation (3.25) gives the Kapustinskii equation:

$$U = \frac{1200.5 V Z_+ Z_-}{r_c + r_a}\left(1 - \frac{0.345}{r_c + r_a}\right) \text{ kJ mol}^{-1} \tag{3.27}$$

Table 3.7 ***Thermochemical radii (Å) of complex anions***

BF_4^-	2.28	CrO_4^{2-}	2.40	IO_4^-	2.49
SO_4^{2-}	2.30	MnO_4^-	2.40	MoO_4^{2-}	2.54
ClO_4^-	2.36	BeF_4^-	2.45	SbO_4^{3-}	2.60
PO_4^{3-}	2.38	AsO_4^{3-}	2.48	BiO_4^{3-}	2.68
OH^-	1.40	O_2^{2-}	1.80	CO_3^{2-}	1.85
NO_2^-	1.55	CN^-	1.82	NO_3^-	1.89

Source: Reproduced with permission from Kapustinskii, *Quart. Rev.*, 283–294, © 1956 Royal Society of Chemistry.

where V is the number of ions per formula unit (two in NaCl, three in PbF_2, etc.). This equation may be used to calculate U of any known or hypothetical ionic compound and, in spite of the assumptions involved, the answers obtained are surprisingly accurate.

The Kapustinskii equation was used to predict the stable existence of several previously unknown compounds. In cases where U was known from Born–Haber cycle calculations (next section), it was used to derive values for ionic radii. This has been particularly useful for complex anions, e.g. SO_4^{2-}, PO_4^{3-}, whose effective size in crystals is difficult to measure by other means. Radii determined in this way are known as *thermochemical radii*; some values are given in Table 3.7. It should be noted that radii for non-spherical ions such as CN^- represent gross simplifications and are really applicable only to other lattice energy calculations.

3.2.7 The Born–Haber cycle and thermochemical calculations

The lattice energy of a crystal is equivalent to its heat of formation from 1 mol of its ionic constituents in the gas phase:

$$Na^+(g) + Cl^-(g) \rightarrow NaCl(s), \quad \Delta H = U$$

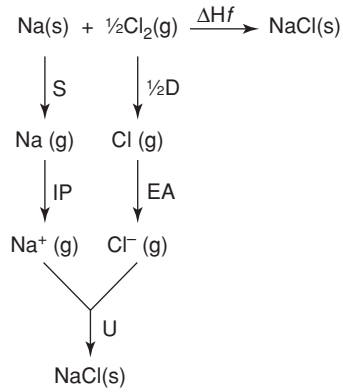
It cannot be measured experimentally. However, the heat of formation of a crystal, ΔH_f, can be measured relative to the reagents in their standard states:

$$Na(s) + \tfrac{1}{2}Cl_2(g) \rightarrow NaCl(s), \quad \Delta H = \Delta H_f$$

ΔH_f may be related to U by constructing a thermochemical cycle known as a Born–Haber cycle, in which ΔH_f is given by the summation of energy terms in a hypothetical reaction pathway. For NaCl, the individual steps in the pathway, commencing with the elements in their standard states, are as follows:

Sublimation of solid Na	$\Delta H = S$
Ionisation of gaseous Na atoms	$\Delta H = IP$
Dissociation of Cl_2 molecules	$\Delta H = \frac{1}{2}D$
Formation of the Cl^- ion	$\Delta H = EA$
Coalescence of gaseous ions to give crystalline NaCl	$\Delta H = U$

Addition of these five terms is equivalent to forming crystalline NaCl from solid Na and gaseous Cl_2, as follows:



From Hess' law:

$$\Delta H_f = S + {}^1\!/_2 D + IP + EA + U \tag{3.28}$$

The Born–Haber cycle and equation (3.28) have various uses:

(a) Six enthalpy terms are present in equation (3.28). If all six can be determined independently, then the cycle gives a check on internal consistency of the data. The values (in kJ mol^{-1}) for NaCl are as follows:

S	109
IP	493.7
$\frac{1}{2}D$	121
EA	−356
U	−764.4
ΔH_f	−410.9

Summation of the first five terms gives a calculated ΔH_f of −396.7 kJ mol^{-1}, which compares reasonably well with the measured ΔH_f value of −410.9 kJ mol^{-1}.

(b) If only five of the terms are known, the sixth may be evaluated using equation (3.28). An early application was in the calculation of electron affinities, for which data were not then available.

(c) The stability of an unknown compound may be estimated. It is necessary to assume a structure in order to calculate U; although there are obviously errors involved, e.g. in choosing r_e, these are usually unimportant compared with the effect of some of the other energy terms in equation (3.28). Having estimated U, ΔH_f may be calculated. If ΔH_f is large and positive, it is clear why the compound is unknown: it is unstable relative to its elements. If ΔH_f(calc.) is negative, however, it may be worthwhile to try to prepare the compound under certain conditions. Examples are given in the next section.

(d) Differences between U values obtained by the Born–Haber cycle (i.e. thermochemical lattice energies) and calculated from an ionic model of the crystal structure may be used as evidence for non-ionic bonding effects. Data for the silver halides, Table 3.8, and for thallium and copper halides (not given) show that differences between the two lattice energies are least for the fluorides and greatest for the iodides. This is attributed to covalent bonding in the iodides leading to an increase in thermochemical U values. A correlation also exists between the insolubility of the Ag salts, especially AgI, in water and

Table 3.8 ***Lattice energies (kJ mol^{-1}) of Group I halides***

Compound	U_{calc}	$U_{Born–Haber}$	ΔU
AgF	−920	−953	33
AgCl	−832	−903	71
AgBr	−815	−895	80
AgI	−777	−882	105

Source: Reproduced with permission from Waddington, *Adv. Inorg. Chem. Radiochem.*, **1**, 157–221, © 1959 Elsevier.

the presence of partial covalent bonding. Data for the corresponding alkali halides show that differences between thermochemical and calculated lattice energies are small; therefore, the ionic bonding model may be applied satisfactorily to them. While covalent bonding is present in AgCl and AgBr, it is not strong enough to change the crystal structure from that of rock salt to one of lower *CN*. AgI is different, however; it is polymorphic and exists in at least three structure types, all of which have low *CN*s, usually four. Changes in structure and *CN* due to increased covalent bonding are described later.

(e) Certain transition elements have *crystal field stabilisation energies*, CFSE, due to their *d* electron configuration, and this gives an increased lattice energy in their compounds. For example, the difference between experimental and calculated *U* in CoF_2 is 83 kJ mol^{-1}, in fair agreement with the CFSE value calculated for the high-spin state of Co in CoF_2 of 104 kJ mol^{-1}. Ions which do not exhibit CFSE effects have the configurations d^0 (e.g. Ca^{2+}), d^5 high spin (e.g. Mn^{2+}) and d^{10} (e.g. Zn^{2+}). Further details are given in Section 3.2.15.1.

(f) The Born–Haber cycle has many other uses, e.g. in solution chemistry to determine energies of complexation and hydration of ions. These usually require knowledge of the lattice energy of the appropriate solids, but since they do not provide any new information about solids, they are not discussed further.

3.2.8 Stabilities of real and hypothetical ionic compounds

3.2.8.1 Inert gas compounds

One may ask whether it is worthwhile trying to synthesise, for example, ArCl. Apart from ΔH_f, the only unknown in equation (3.28) is *U*. Suppose that hypothetical ArCl had the rock salt structure and the radius of the Ar^+ ion is between those of Na^+ and K^+. An estimated lattice energy for ArCl is, then, −745 kJ mol^{-1} (NaCl = −764.4: KCl = −701.4 kJ mol^{-1}). Substitution in equation (3.28) gives, in kJ mol^{-1}:

S	$^1/_2 D$	IP	EA	*U*	ΔH_f(calc.)
0	121	1524	−356	−745	+544

ArCl has a large positive ΔH_f(calc.) and would be thermodynamically unstable, by a large amount, relative to the elements.

Such a calculation also tells us *why* ArCl is unstable and cannot be synthesised. Comparing the calculations for ArCl and NaCl, the instability of ArCl is due to the very high ionisation potential of Ar (stability is strictly governed by free energies of formation, but ΔS is small and hence $\Delta G \simeq \Delta H$). The heats of formation calculated for several other hypothetical compounds are given in Table 3.9.

A large number of inert gas compounds are now known, following the preparation of $XePtF_6$ by Bartlett in 1962. Consideration of lattice energies and enthalpies of formation led Bartlett to try to synthesise $XePtF_6$ by direct reaction of Xe and PtF_6 gases. He had previously prepared O_2PtF_6 as an ionic salt, $(O_2)^+(PtF_6)^-$,

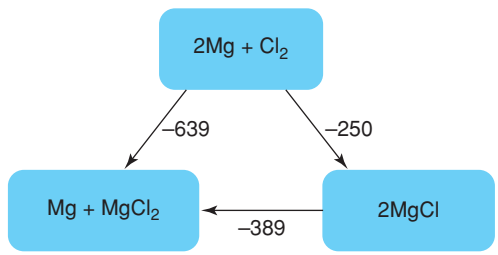
Table 3.9 ***Enthalpies of formation (kJ mol^{-1}) of some hypothetical (*) and real compounds***

HeF*	+1066	NeCl*	+1028	$CsCl_2^*$	+213	CuI_2	−21
ArF*	+418	NaCl	−411	CsF_2^*	−125	$CuBr_2$	−142
XeF*	+163	MgCl*	−125	AgI_2^*	+280	$CuCl_2$	−217
$MgCl_2$	−639	AlCl*	−188	$AgCl_2$	+96		
$NaCl_2^*$	+2144	$AlCl_3$	−694	AgF_2	−205		

by reacting (by accident) O_2 with PtF_6. From a knowledge of the similarity in the first ionisation potentials of molecular oxygen (1176 kJ mol^{-1}) and xenon (1169 kJ mol^{-1}), he reasoned, correctly, that the corresponding Xe compound should be stable.

3.2.8.2 Lower and higher valence compounds

Consider alkaline earth compounds. In these, the metal is always divalent. Since a great deal of extra energy is required to doubly ionise the metal atoms, it is reasonable to ask why monovalent compounds, such as 'MgCl', do not form. The data in Table 3.9 show that MgCl is indeed stable relative to the elements [ΔH(calc.) = −125 kJ mol^{-1}] but that $MgCl_2$ is much more stable ($\Delta H = -639$ kJ mol^{-1}). This is shown by the following sequence of enthalpies:



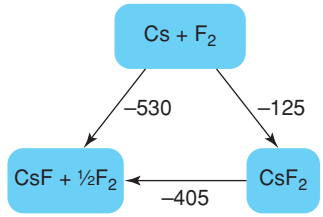
In any attempt to synthesise MgCl, attention should be directed towards keeping the reaction temperature low and/or isolating the MgCl product, in order to prevent it from reacting further or disproportionating. Similar trends are observed for other hypothetical compounds such as ZnCl, Zn_2O, AlCl and $AlCl_2$.

From a consideration of the factors that affect the stability of compounds, the following conclusions may be drawn about compounds with metals in unusual oxidation states:

(a) The formation and stability of compounds with lower than normal valence states are favourable when (i) the second, and higher, ionisation potentials of the metal are particularly high and (ii) the lattice energy of the corresponding compounds with the metal in its normal oxidation state is reduced.
(b) In order to prepare compounds with the metal in a higher than normal oxidation state and during which it may be necessary to break into a closed electron shell, it is desirable to have (i) low values for the second (or higher) ionisation potential of the metal atoms and (ii) large lattice energies of the resulting higher valence compounds.

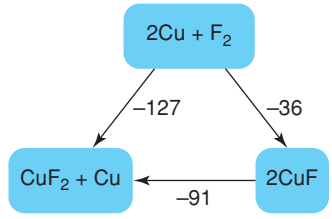
As examples of these trends, calculations for the alkaline earth monohalides show that while all are unstable relative to the dihalides, the enthalpy of disproportionation is least in each case for the iodide [$U_{(MI_2)} < U_{(MBr_2)}$, etc.: effect a, ii]. On the other hand, higher valence halogen compounds of the Group I

elements are most likely to occur with caesium and the copper subgroup elements (effect b, i), in combination with fluorine (effect b, ii). Thus, from Table 3.9, all caesium dihalides, apart from CsF_2, have positive ΔH_f values and would be unstable. CsF_2 is stable in principle, with its negative ΔH_f value, but has not been prepared because its decomposition to CsF has a large negative ΔH:



For the silver dihalides, ΔH_f becomes less positive and finally negative across the series AgI_2 to AgF_2. This again shows the effect of r_e on U and hence on ΔH_f (effect b, ii). Unlike CsF_2, AgF_2 is stable since AgF and AgF_2 have similar enthalpies of formation and the disproportionation enthalpy of AgF_2 (to $AgF + {}^1/_2F_2$) is approximately zero.

The copper halides are particularly interesting. The divalent state of Cu, in which the d^{10} shell of Cu is broken into, is the most common state and again the dihalides show decreasing stability across the series CuF_2 to CuI_2, Table 3.9: CuI_2 appears not to exist and its calculated ΔH_f is barely negative. In the monovalent state, the situation is reversed and all the halides apart from CuF are known. CuF is calculated to be stable relative to the elements but not relative to CuF_2:



These examples show that several factors affect the formulae and stability of compounds: ionisation potentials, lattice energies (via internuclear distances and ionic charges) and the relative stability of elements in different oxidation states. Often a delicate balance between opposing factors controls the stability or instability of a compound and, as with the copper halides, detailed calculations are needed to assess the factors involved.

3.2.9 Effect of partial covalent bonding on crystal structures

Covalent bonding, partial or complete, occurs when the outer electronic charge density on an anion is polarised towards and by a neighbouring cation. The net effect is that an electron pair that would be associated entirely with the anion in a purely ionic structure is displaced to occur between the anion and cation. In cases of partial covalent bonding, some electron density is common to both atoms but the rest is associated with the more electronegative atom. In purely covalent bonding, there is no distinction between cations and anions. The electron distribution in a particular bond is symmetrical between the two participating atoms and the bonds are non-polar.

Table 3.10 **Non-polar covalent radii**

Element	r_c/Å	Element	r_c/Å	Element	r_c/Å	Element	r_c/Å
H	0.32	Mg	1.38	As	1.19	Te	1.35
Li	1.34	Al	1.26	Se	1.16	I	1.33
Be	0.91	Si	1.17	Br	1.14	Cs	2.35
B	0.82	P	1.10	Rb	2.16	Ba	1.98
C	0.77	S	1.04	Sr	1.91	Tl	1.48
N	0.74	Cl	0.99	Ag	1.50	Pb	1.47
O	0.70	K	1.96	Cd	1.46	Bi	1.46
F	0.68	Ca	1.74	Sn	1.40		
Na	1.54	Ge	1.22	Sb	1.38		

Source: after R.T. Sanderson, 1976, Chemical bonds and bond energy, Academic Press; R.T. Sanderson 'The nature of Ionic solids.' *J. Chem. Ed.* 44 516 (1967).

Non-polar covalent radii can be measured accurately and represent a point of reference with which to compare other radii. Thus the atomic, non-polar covalent radius of carbon, given by half the C–C single bond length, is constant at 0.77 Å in materials as diverse as diamond and gaseous hydrocarbons. Non-polar radii, r_c, are listed in Table 3.10.

Some clear-cut examples of the influence of partial covalent bonding on crystal structures of materials which might, at first glance, be expected to be ionic are as follows:

(a) *SrO, BaO, HgO*. SrO and BaO both have the rock salt structure with octahedral M^{2+} ions. Based on size considerations alone, and if it were ionic, HgO should have the same structure. However, mercury is only *two* coordinate in HgO; linear O–Hg–O segments occur and may be rationalised on the basis of *sp* hybridisation of Hg. The ground state of atomic Hg is

$$(\text{Xe core})\, 4f^{14}\, 5d^{10}\, 6s^2$$

The first excited state, corresponding to Hg (II), is

$$(\text{Xe core})\, 4f^{14}\, 5d^{10}\, 6s^1\, 6p^1$$

Hybridisation of the 6*s* and one 6*p* orbitals gives rise to two, linear *sp* hybrid orbitals, each of which forms a normal, electron pair covalent bond by overlap with an orbital on oxygen. Hence Hg has a *CN* of 2 in HgO.

(b) *AlF_3, $AlCl_3$, $AlBr_3$, AlI_3*. These show a smooth transition from ionic to covalent bonding as the electronegativity difference between the two elements decreases. Thus, AlF_3 is a high-melting, essentially ionic solid with a distorted octahedral coordination of Al^{3+} ions; its structure is related to that of ReO_3. $AlCl_3$ has a layered, polymeric structure in the solid state similar to that of $CrCl_3$, which is related to the $CdCl_2$ and CdI_2 structures. The bonds may be regarded as part ionic/part covalent. $AlBr_3$ and AlI_3 have molecular structures with dimeric units of formula Al_2X_6. Their structure and shape are shown in Fig. 1.26 and the bonding between Al, and Br, or I is essentially covalent.

(c) Halides of other elements, e.g. Be, Mg, Ga, In, also show variations in bond type and structure, depending on the halide. The trends are always the same: the fluorides have the largest difference in electronegativity between the two elements and their structures are the most ionic. Increasing covalent character occurs in

the series fluoride–chloride–bromide–iodide. The occurrence of partial covalent bonding in an otherwise ionic structure may sometimes be detected by abnormally large values of the lattice energy, Table 3.8. In others, it is clear from the nature of the structure and the coordination numbers of the atoms that the bonding cannot be purely ionic.

It is difficult to quantify the degree of partial covalence in a particular structure, although chemists intuitively feel that most so-called ionic structures must have a considerable degree of covalent character. Two atomic parameters that undoubtedly have a strong influence on bond character are ionisation potential (i.e. the ease with which atoms ionise to form cations) and electron affinity (i.e. the strength with which atoms attract electrons to form anions). Two other atomic parameters related to these that also strongly influence bond type are effective nuclear charge and electronegativity. Two approaches which have had considerable success in quantifying ionicity or partial covalent bonding and that use the concepts of effective nuclear charge and electro-negativity are the coordinated polymeric model of Sanderson and the ionicity plots of Mooser and Pearson.

3.2.10 Effective nuclear charge

The *effective nuclear charge* of an atom is the positive charge that would be felt by a foreign electron on arriving at the periphery of the atom. Atoms are, of course, electrically neutral overall but, nevertheless, the valence electrons are not very effective in shielding the outside world from the positive charge on the nucleus. Consequently, an incoming electron (e.g. an electron that belongs to a neighbouring atom and is coming to investigate the possibility of bond formation!) feels a positive, attractive charge. Were this not the case and the surface of an atom were completely shielded from the nuclear charge, then atoms would have zero *electron affinity* and no bonds, ionic or covalent, would ever form.

The effective nuclear charge is greatest in elements which have a single vacancy in their valence shell, i.e. in the halogens. In the inert gases, the outermost electron shell is full and no foreign electrons can enter it. Consequently, incoming electrons would have to occupy vacant orbitals which are essentially 'outside' the atom and the effective nuclear charge experienced in such orbitals would be greatly reduced. Calculations of *screening constants* were made by Slater, who found that outermost electrons are much less efficient at screening nuclear charge than are electrons in inner shells. The screening constants of outermost electrons are calculated to be approximately one-third; this means that for each unit increase in atomic number and positive nuclear charge across a series, e.g. from sodium to chlorine, the additional positive charge is screened by only one-third. Therefore, the effective nuclear charge increases in steps of two-thirds and the valence electrons experience an increasingly strong attraction to the nucleus on going from sodium to chlorine. Similar effects occur throughout the Periodic Table: the effective nuclear charge is small for the alkali metals and increases to a maximum in the halogens.

Many atomic properties correlate with effective nuclear charge:

(a) Ionisation potentials increase from left to right across the Periodic Table.
(b) Electron affinities become increasingly negative in the same direction.
(c) Atomic radii decrease from left to right, Table 3.10.
(d) Electronegativities increase from left to right.

3.2.11 Electronegativity and partially charged atoms

The electronegativity of an atom is a measure of the net attractive force experienced by an outermost electron towards the nucleus. Electronegativity, originated by Pauling, is a parameter that correlates with the polarity

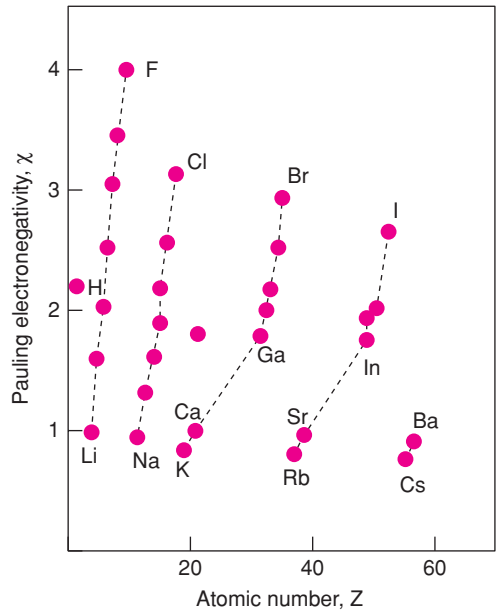


Figure 3.6 ***Variation of Pauling electronegativities with position in the Periodic Table. Reproduced with permission from* Shriver and Atkins, Inorganic Chemistry*, 4th edition, © 2006 Oxford University Press.***

of bonds between unlike atoms. Atoms of high electronegativity attract electrons (in a covalent bond) more than do atoms of low electronegativity and, hence, they acquire a partial negative charge. The magnitude of the partial charge depends on the initial difference in electronegativity between the two atoms. Pauling observed a correlation between the strengths of polar bonds and the degree of polarity in the bonds. The variation of Pauling electronegativities with position in the Periodic Table is shown in Fig. 3.6. A useful concept to aid our understanding of bond formation is the *principle of electronegativity equalisation*. It may be stated as: 'When two or more atoms initially different in electronegativity combine chemically, they adjust to have the same intermediate electronegativity within the compound.' Thus, in a bond between unlike atoms, the bonding electrons are preferentially and partially transferred from the less electronegative to the more electronegative atom, resulting in partially charged atoms.

A list of the partial charges that are calculated to be associated with oxygen in a variety of oxides is given in Table 3.11, where values cover almost the entire range between 0 and -1. Although oxides are traditionally regarded as containing the oxide ion, O^{2-}, calculations show that the actual charge carried by an oxygen never exceeds -1 and is usually much less than -1. This is, of course, consistent with electron affinity data for oxygen. Although these are *gas-phase* data, it is clear that while the reaction, $O + e \rightarrow O^-$ is exothermic, $O^- + e \rightarrow O^{2-}$ is endothermic and, therefore, unfavourable.

From the principle of electronegativity equalisation, electrons in a hypothetical covalent bond are partially transferred to the more electronegative atom. This removal of electrons from the electropositive atom leads to an increase in its effective nuclear charge, a decrease in its size and hence an increase in its effective electronegativity. Likewise, as the electronegative atom acquires electrons, so its ability to attract still more electrons diminishes and its electronegativity decreases. In this way, the electronegativities of the two atoms adjust themselves until they are equal. This principle of electronegativity equalisation may be applied equally to diatomic gas molecules, in which only one bond is involved, or to 3D solid structures, in which each atom

Table 3.11 ***Partial charge on oxygen in some solid oxides***

Compound	$-\delta_O$	Compound	$-\delta_O$	Compound	$-\delta_O$	Compound	$-\delta_O$
Cu_2O	0.41	BeO	0.36	BaO	0.68	La_2O_3	0.56
Ag_2O	0.41	PbO	0.36	Ga_2O_3	0.19	CO_2	0.11
Li_2O	0.80	SnO	0.37	Tl_2O_3	0.21	GeO_2	0.13
Na_2O	0.81	FeO	0.40	In_2O_3	0.23	SnO_2	0.17
K_2O	0.89	CoO	0.40	B_2O_3	0.24	PbO_2	0.18
Rb_2O	0.92	NiO	0.40	Al_2O_3	0.31	SiO_2	0.23
Cs_2O	0.94	MnO	0.41	Fe_2O_3	0.33	MnO_2	0.29
HgO	0.27	MgO	0.50	Cr_2O_3	0.37	TiO_2	0.39
ZnO	0.29	CaO	0.56	Sc_2O_3	0.47	ZrO_2	0.44
CdO	0.32	SrO	0.60	Y_2O_3	0.52	HfO_2	0.45
CuO	0.32						

is surrounded by and bonded to several others. It illustrates how covalent bonds that are initially non-polar may become polar due to electronegativity equalisation.

An alternative approach is to start with purely ionic bonds and consider how they may acquire some covalent character. In an ionic structure, M^+X^-, the cations are surrounded by anions (usually 4, 6 or 8). However, the cations have empty valence shells and are potential electron pair acceptors; likewise, anions with their filled valence shells are potential electron pair donors. The cations and anions therefore interact in the same way as do Lewis acids and bases: the anions, with their lone pairs of electrons, coordinate to the surrounding cations. The strength of this interaction, and hence the degree of covalent bonding which results, is again related to the electronegativities of the two atoms. Thus, electronegative cations such as Al^{3+} are much stronger electron pair acceptors (and Lewis acids) than are electropositive cations such as K^+. The *coordinated polymeric model* of structures, proposed by Sanderson, is based on this idea of acid–base interactions between ions and forms a bridge between the ionic and covalent extremes of bond type.

3.2.12 Coordinated polymeric structures – Sanderson's model

In Sanderson's model, all bonds in non-molecular structures are regarded as polar covalent. The atoms carry partial charges whose values may be calculated using a scale of electronegativities developed by Sanderson. From the partial charges, the relative contributions of ionic and covalent bonding to the total bond energy may be estimated. Since all non-molecular crystals contain partially charged atoms, the ionic model represents an extreme form of bonding that is not attained in practice. Thus in KCl, the charges on the atoms are calculated to be ± 0.76 instead of ± 1.0 for a purely ionic structure.

Sanderson developed a method to evaluate the partial charges and atomic radii in a large number of solid compounds. Data are given in Table 3.12 for a number of mono- and divalent chlorides; the charge on Cl varies from -0.21 in $CdCl_2$ to -0.81 in CsCl; at the same time the calculated radius of Cl varies from 1.24 to 1.95 Å. These compare with a non-polar covalent radius of 0.99 Å and an ionic radius of 2.18 Å, Table 3.10. Although these radii and partial charges may not be quantitatively correct, because of the assumptions involved in their calculation, they nevertheless provide useful insights. Most of the compounds in Table 3.12 are normally regarded as ionic, and if the partial charge data are in any way correct, they clearly show that it is unrealistic and misleading to assign a radius to the chloride ion which is constant for all solid chlorides.

Table 3.12 ***Partial charge and radius of the chlorine atom in some solid chlorides***

Compound	$-\delta_{Cl}$	r_{Cl}/Å	Compound	$-\delta_{Cl}$	r_{Cl}/Å
$CdCl_2$	0.21	1.24	$BaCl_2$	0.49	1.57
$BeCl_2$	0.28	1.26	LiCl	0.65	1.76
CuCl	0.29	1.34	NaCl	0.67	1.79
AgCl	0.30	1.35	KCl	0.76	1.90
$MgCl_2$	0.34	1.39	RbCl	0.78	1.92
$CaCl_2$	0.40	1.47	CsCl	0.81	1.95
$SrCl_2$	0.43	1.50			

3.2.13 Mooser–Pearson plots and ionicities

While the radius ratio rules are rather unsatisfactory for predicting the structure adopted by, for instance, particular AB compounds, an alternative approach by Mooser and Pearson has had considerable success. This focuses on the directionality or covalent character of bonds. The two factors regarded as influencing covalent bond character in crystals are (i) the average principal quantum number, $\bar{n}$, of the atoms involved and (ii) their difference in electronegativity, $\Delta\chi$. In constructing Mooser–Pearson plots, these two parameters are plotted against each other, as shown for AB compounds in Fig. 3.7. The most striking feature is the clean separation of compounds into four groups corresponding to the structures: zinc blende (ZnS, B), wurtzite (ZnS, W), rock salt and CsCl. One's intuitive feeling is that zinc blende (or sphalerite) is the most covalent and either rock salt or CsCl is the most ionic. Mooser–Pearson plots present this intuitive feeling in diagrammatic form.

The ionic character in the bonds, termed *ionicity*, increases from the bottom left to the top right of the diagram, as shown by the arrow. Thus, ionicity is not governed by electronegativity alone but also depends

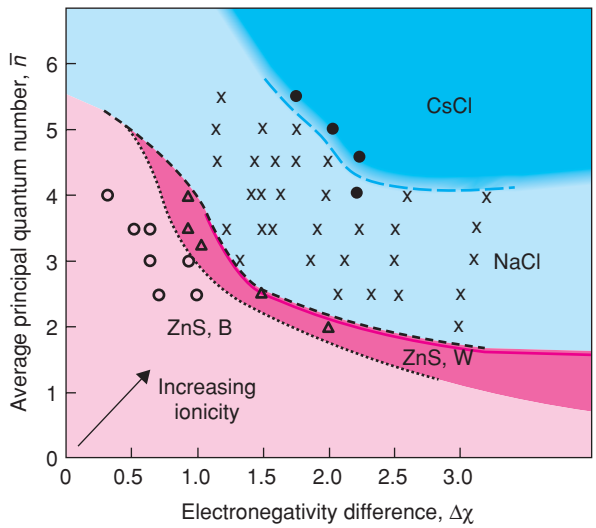


Figure 3.7 ***Mooser–Pearson plot for AB compounds containing A group cations. The arrow indicates direction of increasing bond ionicity. Adapted with permission from* Acta Cryst.*, 12, 1015, © 1959 The International Union of Crystallography.***

on the principal valence shell of the atoms and hence on atomic size. There is a general trend for highly directional covalent bonds to be associated with light elements, i.e. at the bottom of Fig. 3.7, and with small values of $\Delta\chi$, i.e. at the left-hand side of Fig. 3.7.

The fairly sharp crossover between the different '*structure fields*' in Fig. 3.7 suggests that for each structure type there are critical ionicities which place a limit on the compounds that can have that particular structure. Theoretical support for Mooser–Pearson plots has come from the work of Phillips and van Vechten, who measured optical absorption spectra of some AB compounds and, from these, calculated electronegativities and ionicities. The spectral data gave values of band gaps, E_g (see later). For isoelectronic series of compounds, e.g. ZnSe, GaAs and Ge, the band gaps have contributions from (a) a homopolar band gap, E_h, as in pure Ge, and (b) a charge transfer, C, between A and B, termed the 'ionic energy'. These are related by

$$E_g^2 = E_h^2 + C^2 \tag{3.29}$$

E_g and E_h are measured from the spectra and hence C can be calculated. C is related to the energy required for electron charge transfer in a polar bond and hence is a measure of electronegativity, as defined by Pauling. A scale of ionicity, f_i, was devised:

$$f_i = \frac{C^2}{E_g^2} \tag{3.30}$$

Values of f_i range from zero ($C = 0$ in a homopolar covalent bond) to one ($C = E_g$ in an ionic bond) and give a measure of the fractional ionic character of a bond. Phillips analysed the spectroscopic data for 68 AB compounds with either octahedral or tetrahedral structures and found that the compounds fall into two groups separated by a critical ionicity, f_i, of 0.785.

The link between Mooser–Pearson plots and Phillips–Van Vechten ionicities is

$$\Delta x \text{ (Mooser–Pearson)} \simeq C \text{ (Phillips)}$$

and

$$\bar{n} \text{ (Mooser–Pearson)} \simeq E_h \text{ (Phillips)}$$

The explanation of the latter is that as $\bar{n}$ increases, the outer orbitals become larger and more diffuse and the energy differences between outer orbitals (s, p, d and/or f) decrease; the band gap, E_h, decreases until metallic behaviour occurs at $E_h = 0$. Use of $\bar{n}$ gives the average behaviour of anion and cation. The Phillips–Van Vechten analysis has so far been restricted to AB compounds, but this has provided a theoretical justification for the more widespread use and application of the readily constructed Mooser–Pearson plots.

3.2.14 Bond valence and bond length

The structures of most *molecular* materials – organic and inorganic – may be described satisfactorily using valence bond theory in which single, double, triple and occasionally partial bonds occur between atoms. It may be difficult to apply valence bond theory to *non-molecular* inorganic materials, even though the bonding in them may be predominantly covalent, in cases where insufficient bonding electrons are available for each bond to be treated as an electron pair single bond.

An empirical but very useful approach to describing partial bonds in such cases has been developed by Pauling, Brown, Shannon, Donnay and others, and involves the evaluation of *bond orders* or *bond valences*

in a structure. Bond valence is defined in a similar way to electrostatic bond strength in Pauling's electrostatic valence rule for ionic structures, Section 3.3, and represents an extension of Pauling's rule to structures that are not necessarily ionic. Bond valences are defined empirically, using information on atom oxidation states and experimental bond lengths; no reference is made, at least not initially, to the nature of the bonding, whether it be covalent, ionic or some blend of the two.

Pauling's electrostatic valence rule requires the sum of the electrostatic bond strength, *ebs*, between an anion and its neighbouring cations to equal the formal charge on the anion, equations (3.2) and (3.3). This rule may be modified to include structures which are not necessarily ionic by replacing (a) the *ebs* by the *bond valence* (*bv*) and (b) the formal charge on the anion by the valence of that atom (valence being defined as the number of electrons that take part in bonding). This leads to the *valence sum rule*, which relates the valence, V_i, of atom i to the bond valence, bv_{ij}, between atom i and neighbouring atom j:

$$V_i = \sum_j bv_{ij} \tag{3.31}$$

The valence of an atom must equal the sum of the *bv*s for all the bonds that it forms. For cases in which bv_{ij} is an integer, this rule becomes the familiar rule that is used for evaluating the number of bonds around an atom in molecular structures, i.e. the valence of an atom is equal to the number of bonds that it forms (counting double bonds as two bonds, etc.). In non-molecular structures, however, integral *bv*s are the exception rather than the rule.

The *ebs* in Pauling's rule is given by the ratio cation charge:cation coordination number. For structures in which the cation coordination is irregular or the cation–anion bonds are not all of the same length, only an average *ebs* is obtained. One advantage of the *bv* approach is that each bond is treated separately and hence irregularities or distortions in coordination environments are taken into account.

For a given pair of elements, an inverse correlation between *bv* and bond length exists, as shown in Fig. 3.8 for bonds between oxygen and atoms of the second row in their group valence states, i.e. Na^{I}, Mg^{II}, Al^{III}, Si^{IV}, P^{V} and S^{VI}. While each atom and oxidation (or valence) state has its own *bv*–bond length curve, 'universal

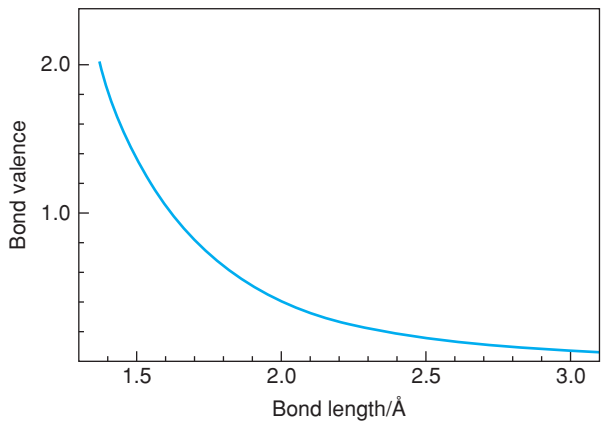


Figure 3.8 ***Bond valence–bond length universal correlation curve for bonds between oxygen and second row atoms: Na, Mg, Al, Si, P and S. Reproduced with permission from Brown,* Chem. Soc. Rev.*, 7(3), 359, © 1978 Royal Society of Chemistry.***

curves' such as that in Fig. 3.8 may be used for isoelectronic series of ions. Various analytical expressions have been used to fit curves such as Fig. 3.8, including

$$bv_{ij} = \left(\frac{R_0}{R}\right)^N \tag{3.32}$$

where R is the bond length and R_0 and N are constant (R_0 is the value of the bond length for unit bv); for the elements represented by Fig. 3.8, $R_0 = 1.622$ and $N = 4.290$.

From Fig. 3.8, bond length increases with decreasing bv; since, for a given atom, bv must decrease with increasing CN, a correlation also exists between CN and bond length. This correlation was presented, in slightly different form, in Fig. 3.3, in which cation radii increase with increase in CN. Since the data in Fig. 3.3 assume a constant radius of the fluoride or oxide ion, the ordinate in Fig. 3.3 could be changed from the ionic radius to the metal–oxygen bond length.

Curves such as that in Fig. 3.8 are important because they lead to an increased rationalisation and understanding of crystal structures. They have several applications associated with structure determination, for example:

(a) As a check on the correctness of a proposed structure, since the valence sum rule should be obeyed by all atoms of the structure to within a few percentage points.
(b) To locate hydrogen atoms, which are often 'invisible' in X-ray structure determinations because of the very low scattering power of H. The bv sums around each atom are determined. If any atom shows a large discrepancy between the atom valence and the bvs, then H is likely to be bonded to such atoms.
(c) To distinguish between Al^{3+} and Si^{4+} positions in aluminosilicate structures. Al^{3+} and Si^{4+} cannot be distinguished by X-ray diffraction because of their very similar scattering power, but in sites of similar CN they give different bvs; e.g. in regular MO_4 tetrahedra, Si–O bonds have a bv of 1 but Al–O bonds have a bv of 0.75. Site occupancies may therefore be determined using the valence sum rule and/or by comparing the M–O bond lengths with values expected for Si–O and Al–O.

3.2.15 Non-bonding electron effects

In this section, the influence on structure of two types of non-bonding electrons is considered: the d electrons in transition metal compounds and the s^2 pair of electrons in compounds of the heavy p-block elements in low oxidation states. These two types of electron do not take part in bonding as such, but nevertheless exert a considerable influence on the coordination number and environment of the parent metal atom.

3.2.15.1 d-Electron effects

In transition metal compounds, the majority of the d electrons on the metal atom do not usually take part in bond formation but do influence the coordination environment of the metal atom and are responsible for properties such as colour and magnetism. For present purposes, basic *crystal field theory* (CFT) is adequate to describe qualitatively the effects that occur. It is assumed that the reader is acquainted with CFT and only a summary is given here.

1. Crystal field splitting of energy levels

In an octahedral environment, the five d orbitals on a transition metal atom are not degenerate but split into two groups, the t_{2g} group of lower energy and the e_g group of higher energy, Fig. 3.9. If possible, electrons occupy orbitals singly, according to Hund's rule of maximum multiplicity. For d^4 to d^7 atoms or ions, two

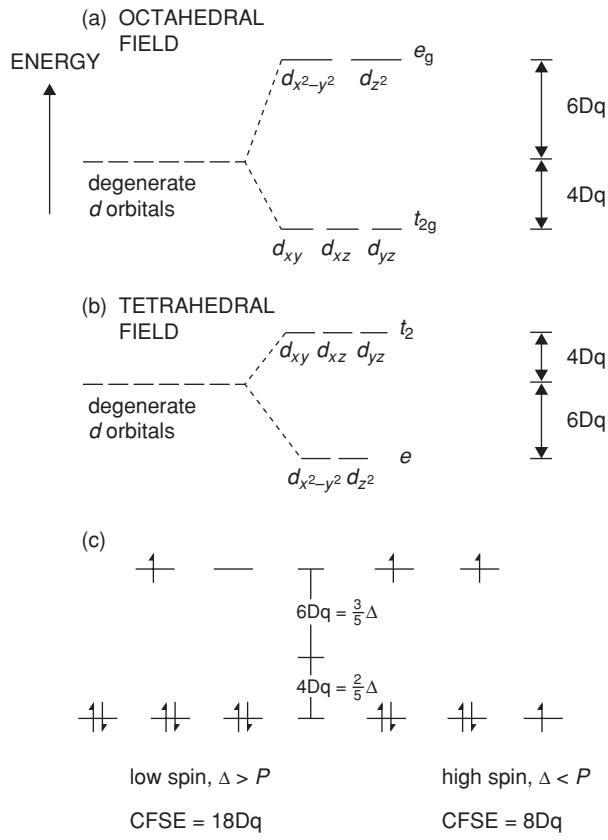


Figure 3.9 ***Splitting of* d *energy levels in (a) an octahedral and (b) a tetrahedral field. (c) Low and high spin states for a* d^7 *transition metal ion in an octahedral environment.***

possible configurations occur, giving *low-spin* (LS) and *high-spin* (HS) states; these are shown for a d^7 ion in Fig. 3.9(c). The increased energy, Δ, required to place an electron in an e_g orbital, and hence maximise the multiplicity, has to be balanced against the repulsive energy or *pairing energy*, P, which arises when two electrons occupy the same t_{2g} orbital. The magnitude of Δ depends on the ligand or anion to which the metal is bonded: for *weak field* anions, Δ is small and the HS configuration occurs, and vice versa for *strong field* ligands. Δ also depends on the metal and, in particular, to which row it belongs: generally $\Delta(5d) > \Delta(4d) > \Delta(3d)$. Consequently HS behaviour is rarely observed in the 4*d* and 5*d* series. Δ values may be determined experimentally from electronic spectra. The possible spin configurations for the different numbers of *d* electrons are given in Table 3.13.

The radii of transition metal ions depend on their *d* electron configuration, as shown in Fig. 3.10(a) for octahedrally coordinated divalent ions. With increasing atomic number, several trends occur. First, there is a gradual decrease in radius as the *d* shell is filled, as shown by the dashed line that passes through Ca^{2+}, Mn^{2+} (HS) and Zn^{2+}. For these three ions, the distribution of *d* electron density is spherically symmetrical because the *d* orbitals are either empty (Ca), singly occupied (Mn) or doubly occupied (Zn). This decrease in radius is associated with poor shielding of the nuclear charge by the *d* electrons; hence a greater effective nuclear charge is experienced by the outer, bonding electrons resulting in a steady contraction with increasing

Table 3.13 **d-*Electron configurations in octahedrally coordinated metal atoms***

No. of electrons	Low spin, $\Delta > P$		High spin, $\Delta < P$		Gain in orbital energy for low spin	Example
	t_{2g}	e_g	t_{2g}	e_g		
1	↑		↑			V^{4+}
2	↑↑		↑↑			Ti^{2+}, V^{3+}
3	↑↑↑		↑↑↑			V^{2+}, Cr^{3+}
4	↑↓↑↑		↑↑↑	↑	Δ	Cr^{2+}, Mn^{3+}
5	↑↓↑↓↑		↑↑↑	↑↑	2Δ	Mn^{2+}, Fe^{3+}
6	↑↓↑↓↑↓		↑↓↑↑	↑↑	2Δ	Fe^{2+}, Co^{3+}
7	↑↓↑↓↑↓	↑	↑↓↑↓↑	↑↑	Δ	Co^{2+}
8	↑↓↑↓↑↓	↑↑	↑↓↑↓↑↓	↑↑		Ni^{2+}
9	↑↓↑↓↑↓	↑↓↑	↑↓↑↓↑↓	↑↓↑		Cu^{2+}
10	↑↓↑↓↑↓	↑↓↑↓	↑↓↑↓↑↓	↑↓↑↓		Zn^{2+}

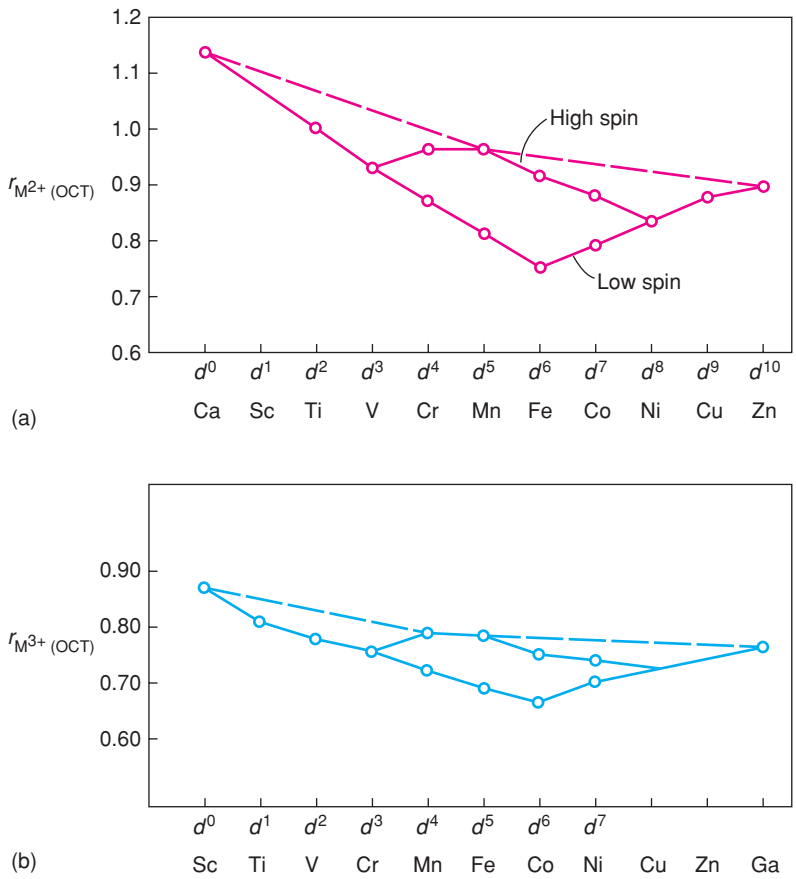


Figure 3.10 ***Radii in octahedral coordination of (a) divalent and (b) trivalent transition metal ions relative to r_F = 1.19 Å. Data from Shannon and Prewitt,* Acta Cryst., B25*, 925, © 1970 The International Union of Crystallography.***

atomic number. Similar effects occur across any horizontal row of the Periodic Table, but are particularly well documented for the transition metal series.

For the ions d^1 to d^4 and d^6 to d^9, the d electron distribution is not spherical. The shielding of the nuclear charge by these electrons is reduced even further and the radii are smaller than expected. For example, Ti^{2+} has the configuration $(t_{2g})^2$, with two singly occupied t_{2g} orbitals. In octahedral Ti^{2+}, these non-bonding electrons occupy regions of space that are directed away from the (Ti^{2+}–anion) axes. Comparing Ti^{2+} with Ca^{2+}, for instance, Ti^{2+} has an extra nuclear charge of +2 but the two extra electrons in the t_{2g} orbitals do not shield the bonding electrons from this extra charge. Hence, Ti–O bonds in TiO are shorter than Ca–O bonds in CaO due to the stronger attraction between Ti^{2+} and the bonding electrons. This trend continues in V^{2+}, Cr^{2+} (LS), Mn^{2+} (LS) and Fe^{2+} (LS), all of which contain only t_{2g} electrons, Table 3.13. Beyond Fe^{2+} (LS), the electrons begin to occupy e_g orbitals and these shield the nuclear charge more effectively. Consequently, the radii increase again in the series Fe^{2+} (LS), Co^{2+} (LS), Ni^{2+}, Cu^{2+} and Zn^{2+}.

For the high-spin ions, a different trend is observed. On passing from V^{2+} to Cr^{2+} (HS) and Mn^{2+} (HS), electrons enter the e_g orbitals, thereby shielding the nuclear charge and giving an increased radius. However, on passing from Mn^{2+} (HS) to Fe^{2+} (HS), Co^{2+} (HS) and Ni^{2+}, the additional electrons occupy t_{2g} orbitals and the radii decrease once again.

Trivalent transition metal ions show a similar trend but of reduced magnitude, Fig. 3.10(b). However, the various effects that occur are effectively transferred to the right by one atom; hence the smallest trivalent ion is Co^{3+} (LS) instead of Fe^{2+} (LS) which is the smallest divalent ion.

So far, only octahedral coordination of transition metal ions has been considered. Tetrahedral coordination is also common but a different energy level diagram applies to the d electrons. A tetrahedral field splits the d orbitals into two groups, but in the opposite manner to an octahedral field: three orbitals have higher energy d_{xy}, d_{xz} and d_{yz} whereas the other two, $d_{x^2-y^2}$ and d_{z^2} have lower energy, Fig. 3.9(b).

Crystal field splitting of d orbitals in transition metal ions may result in a *crystal field stabilisation energy* (CFSE) and an increased lattice energy of ionic compounds. For example, CoF_2 has the rutile structure with octahedral Co^{2+} (d^7HS), Fig. 3.9(c). The energy difference, Δ, between t_{2g} and e_g orbitals is set equal to 10 Dq; the t_{2g} orbitals are stabilised by an amount 4 Dq whereas the e_g orbitals are destabilised by an amount 6 Dq. Relative to five degenerate orbitals without crystal field splitting, Fig. 3.9(a), the CFSE of Co^{2+} in HS and LS states may be calculated. For the LS state, the CFSE is $6 \times 4\ \text{Dq} - 1 \times 6\ \text{Dq} = 18\ \text{Dq}$. For the HS state, the CFSE is $5 \times 4\ \text{Dq} - 2 \times 6\ \text{Dq} = 8\ \text{Dq}$.

The occurrence of CFSE leads to increased lattice energy. The CFSE calculated for Co^{2+} (HS) in rutile is 104 kJ mol^{-1}. This compares fairly well with the value of 83 kJ mol^{-1} given by the difference between the lattice energy determined from a Born–Haber cycle (2959 kJ mol^{-1}) and that calculated using the Born–Mayer equation (2876 kJ mol^{-1}).

The lattice energies of first-row transition element difluorides are shown in Fig. 3.11. Similar trends are observed for other halides. Ions that do not exhibit CFSE are d^0 (Ca), d^5 (HS) (Mn) and d^{10} (Zn), and their lattice energies fall on the lower, dashed curve. Most ions show some degree of CFSE and their lattice energies fall on the upper, solid curve. For the fluorides, Fig. 3.11, the agreement between the calculated CFSE and the difference in lattice energy ΔU [i.e. U (Born–Haber)–U (Born–Mayer)] is reasonable and indicates that the bonding may be treated as ionic. For the other halides, $\Delta U \gg$ CFSE, which indicates that other effects, perhaps covalent bonding, must be present.

2. Jahn–Teller distortions

In many transition metal compounds, the metal coordination is distorted octahedral and two axial bonds are either shorter or longer than the other four bonds. The Jahn–Teller effect is responsible for these distortions in d^9, d^7 (LS) and d^4 (HS) ions. Consider the d^9 ion Cu^{2+} whose configuration is $(t_{2g})^6(e_g)^3$. One e_g orbital

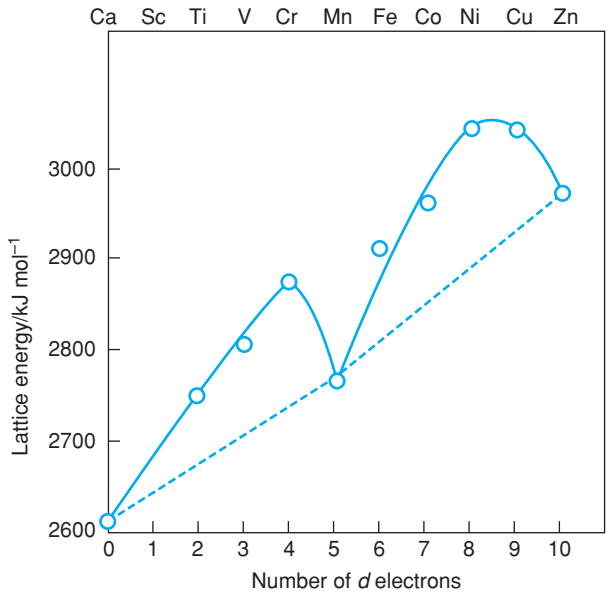


Figure 3.11 ***Lattice energies of transition metal difluorides determined from Born–Haber cycle calculations. Reproduced with permission from Waddington,* Adv. Inorg. Chem. Radiochem.*, 1, 157–221, © 1959 Elsevier.***

contains two electrons and the other contains one. The singly occupied orbital can be either d_{z^2} or $d_{x^2-y^2}$ and in a free ion situation both would have the same energy. However, since the metal ion is not a free ion but is octahedrally coordinated and the two e_g levels are occupied unequally, they are no longer degenerate. The e_g orbitals are high energy orbitals (relative to t_{2g}) since they point directly towards the surrounding ligands; electrons in the doubly occupied orbital experience stronger repulsion and have higher energy than electrons in the singly occupied orbital. This has the effect of lengthening the metal–ligand bonds in the directions of the doubly occupied orbital, e.g. if the d_{z^2} orbital is doubly occupied, the two metal–ligand bonds along the z axis are longer than the other four metal–ligand bonds. The energy level diagram for this latter situation is shown in Fig. 3.12(a). Lengthening of the metal–ligand bond along the z axis leads to a *lowering* of energy of the d_{z^2} orbital. The distorted structure is stabilised by an amount $^1/_2\delta_2$ relative to the regular octahedral arrangement and, hence, the distorted structure becomes the observed, ground state.

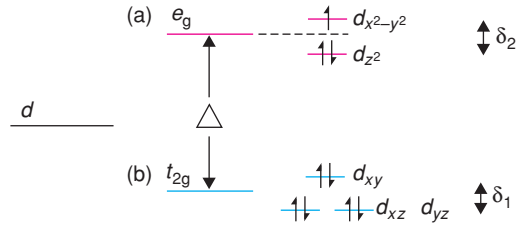


Figure 3.12 ***Energy level diagram for the* d *levels in a* d^9 *ion experiencing a Jahn–Teller distortion. The two bonds parallel to* z *are longer than the other four; hence the* d_{z^2} *orbital is of lower energy than the* $d_{x^2-y^2}$ *orbital.***

High-spin d^4 and low-spin d^7 ions also have odd numbers of e_g electrons and show Jahn–Teller distortions. It is not clear which type of distortion is preferred (i.e. two short and four long bonds, or vice versa) and the actual shapes in a particular structure must be determined experimentally. The degeneracy of the t_{2g} levels may also be removed by the Jahn–Teller effect, but the magnitude of the splitting, δ_1 in Fig. 3.12(b), is small and the effect is relatively unimportant.

The common coordination environment of Cu^{2+} is distorted octahedral with four short and two long bonds. The distortion varies from compound to compound. In CuF_2 (distorted rutile structure), the distortion is fairly small (four F atoms at 1.93 Å, two at 2.27 Å), but is larger in $CuCl_2$ (four Cl atoms at 2.30 Å, two at 2.95 Å) and extreme in tenorite, CuO, which is effectively square planar (four O atoms at 1.96 Å, two at 2.78 Å).

The importance of Jahn–Teller distortions in Cu^{2+} and Cr^{2+} (d^4) compounds is seen by comparing the structures of the oxides and fluorides of the first-row divalent, transition metal ions. For the oxide series, MO (M^{2+} is Ti, V, Cr, Mn, Fe, Co, Ni and Cu), all have the rock salt structure with regular octahedral coordination apart from (a) CuO which contains grossly distorted (CuO_6) octahedra and possibly (b) CrO, whose structure is not known. For the fluoride series, MF_2, all have the regular rutile structure apart from CrF_2 and CuF_2, which have distorted rutile structures. Other examples of distorted octahedral coordination due to the Jahn–Teller effect are found in compounds of Mn^{3+} (HS) and Ni^{3+} (LS).

3. Square planar coordination

The d^8 ions – Ni^{2+}, Pd^{2+}, Pt^{2+} – commonly have square planar or rectangular planar coordination. To understand this, consider the *d* energy level diagram for such ions in (a) octahedral and (b) distorted octahedral fields:

(a) The normal configuration of a d^8 ion in an octahedral field is $(t_{2g})^6(e_g)^2$. The two e_g electrons singly occupy the d_{z^2} and $d_{x^2-y^2}$ orbitals, which are degenerate, and the resulting compounds, with unpaired electrons, are paramagnetic.

(b) Consider, now, the effect of distorting the octahedron and lengthening the two metal–ligand bonds along the *z* axis. The e_g orbitals lose their degeneracy, and the d_{z^2} orbital becomes stabilised by an amount $^1/_2\delta_2$, Fig. 3.12. For small elongations along the *z* axis, the pairing energy required to doubly occupy the d_{z^2} orbital is larger than the energy difference between d_{z^2} and $d_{x^2-y^2}$, i.e. $P > \delta_2$. There is no gain in stability if the d_{z^2} orbital is doubly occupied, therefore, and no reason why small distortions from octahedral coordination should be stable. With increasing elongation along the *z* axis, however, a stage is reached where $P < \delta_2$, and the doubly occupied d_{z^2} orbital becomes stabilised and is the preferred ground state for a d^8 ion. The distortion from octahedral coordination may be sufficiently large that the coordination becomes square planar as in tenorite, CuO (above); in many cases, e.g. PdO, there are no axial ligands along *z* and, the transformation from octahedral to square planar coordination is complete. Because they have no unpaired electrons, square planar compounds are diamagnetic.

Square planar coordination is more common with 4*d* and 5*d* transition elements than with 3*d* elements because the 4*d* and, especially, 5*d* orbitals are more diffuse and extend to greater radial distances from the nucleus. Consequently, the magnitude of the crystal (or ligand) field splitting (Δ, δ) caused by a particular ligand, e.g. O^{2-}, increases in the series $3d < 4d < 5d$. Thus, NiO has the rock salt structure with regular octahedral coordination of Ni^{2+} whereas PdO and PtO both have square planar coordination for the metal atoms (4*d* and 5*d*). The only known compound of Pd with octahedral Pd^{2+} is PdF_2 (rutile structure) and no octahedral Pt^{2+} or Au^{3+} compounds are known.

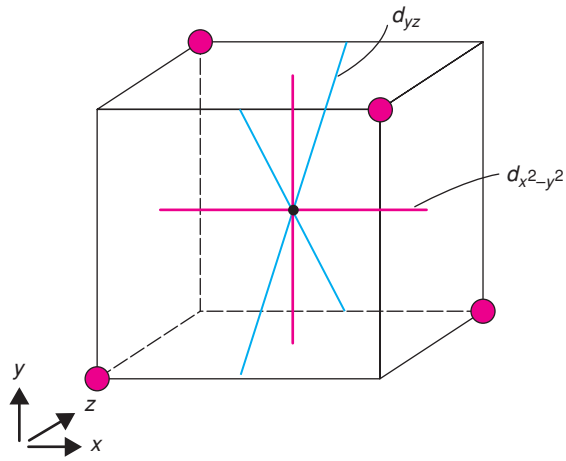


Figure 3.13 ***Orientation of* d *orbitals in a tetrahedral field.***

4. Tetrahedral coordination

As stated earlier, a tetrahedral field splits the d energy levels, but in the opposite sense to an octahedral field, Fig. 3.9(b). Further, the magnitude of the splitting, Δ, is generally less in a tetrahedral field since none of the d orbitals point directly towards the four ligands. Rather, the d_{xy}, d_{xz} and d_{yz} orbitals are somewhat closer to the ligands than are the other two orbitals, Fig. 3.13 (only two orbital directions, d_{yz} and $d_{x^2-y^2}$ are shown). Jahn–Teller distortions again occur, especially when the upper t_2 orbitals contain one, two, four or five electrons [e.g. d^3(HS), d^4(HS), d^5 and d^9]. Details are not given since various types of distortion are possible (e.g. tetragonal or trigonal distortions) and have not been as well studied as octahedral distortions. A common type of distortion is flattening or elongation of the tetrahedron in the direction of one of the twofold axes of the tetrahedron, Fig. 1.25(c). An example is the flattened $CuCl_4$ tetrahedron in Cs_2CuCl_4.

5. Tetrahedral versus octahedral coordination

Most transition metal ions prefer octahedral or distorted octahedral coordination because of their large CFSE in octahedral sites. This can be estimated as follows. In octahedral coordination, each t_{2g} electron experiences a stabilisation of $(4/10)\Delta^{oct}$ and each e_g electron a destabilisation of $(6/10)\Delta^{oct}$, Fig. 3.9(a). Thus Cr^{3+}, $d^3(t_{2g}^3)$, has a CFSE of $1.2\Delta^{oct}$ whereas Cu^{2+}, $d^9(t_{2g})^6(e_g)^3$, has a CFSE of $0.6\Delta^{oct}$. In tetrahedral coordination, each e electron has a stabilisation of $(6/10)\Delta^{tet}$ and each t_2 electron has a destabilisation of $(4/10)\Delta^{tet}$, Fig. 3.9(b). Hence Cr^{3+} would have a CFSE of $0.8\Delta^{tet}$ and Cu^{2+} would have a CFSE of $0.4\Delta^{tet}$. With the guideline that

$$\Delta^{tet} \approx 0.4\Delta^{oct}$$

the values of Δ^{oct} and Δ^{tet} for ions may be used to predict site preferences. More accurate values may be obtained spectroscopically and are given in Table 3.14 for some transition metal oxides. High-spin d^5 ions, and also d^0 and d^{10} ions, have no particular preference for octahedral or tetrahedral sites insofar as crystal field effects are concerned. Ions such as Cr^{3+}, Ni^{2+} and Mn^{3+} show the strongest preference for octahedral coordination: thus tetrahedral coordination is rare for Ni^{2+}.

Table 3.14 ***Crystal field stabilisation energies (kJ mol^{-1}) estimated for transition metal oxides***

Ion		Octahedral stabilisation	Tetrahedral stabilisation	Excess octahedral stabilisation
Ti^{3+}	d^1	87.4	58.5	28.9
V^{3+}	d^2	160.1	106.6	53.5
Cr^{3+}	d^3	224.5	66.9	157.6
Mn^{3+}	d^4	135.4	40.1	95.3
Fe^{3+}	d^5	0	0	0
Mn^{2+}	d^5	0	0	0
Fe^{2+}	d^6	49.7	33.0	16.7
Co^{2+}	d^7	92.8	61.9	30.9
Ni^{2+}	d^8	122.1	35.9	86.2
Cu^{2+}	d^9	90.3	26.8	63.5

Source: Data reproduced with permission from Dunitz and Orgel, *Adv. Inorg Radiochem.*, **2**, 1–60, © 1960 Elsevier.

Good examples of the coordination preferences of ions are found in the sites which they occupy in the spinel structure. Spinels have the formula AB_2O_4 (Chapter 1) and may be:

(a) *normal*: A tetrahedral, B octahedral;
(b) *inverse*: A octahedral, B tetrahedral and octahedral;
(c) some intermediate between normal and inverse.

The parameter γ is the fraction of A ions on octahedral sites. For normal spinel, $\gamma = 0$; for inverse spinel, $\gamma = 1$; for a random arrangement of A and B ions, $\gamma = 0.67$. Lattice energy calculations show that, in the absence of CFSE effects, spinels of the type 2, 3 (i.e. A = M^{2+}, B = M^{3+}, e.g. $MgAl_2O_4$) tend to be normal whereas spinels of type 4, 2 (i.e. A = M^{4+}, B = M^{2+}, e.g. $TiMg_2O_4$) tend to be inverse. However, these preferences may be changed by the intervention of CFSE effects, as shown by the γ values of some 2, 3 spinels in Table 3.15. Examples are as follows:

(a) All chromate spinels are normal with octahedral Cr^{3+}, consistent with the large CFSE for Cr^{3+}; ions such as Ni^{2+} are forced into tetrahedral sites in $NiCr_2O_4$.
(b) Co_3O_4 ($\equiv CoO{\cdot}Co_2O_3$) is normal because low-spin Co^{3+} gains more CFSE in octahedral sites than Co^{2+} loses in tetrahedral sites. Mn_3O_4 is also normal. Magnetite, Fe_3O_4, however, is inverse: whereas Fe^{3+} has no CFSE in either site, Fe^{2+} has a preference for octahedral sites.

Table 3.15 ***Values of the γ parameter of some spinels***

M^{3+}	M^{2+}	Mg^{2+}	Mn^{2+}	Fe^{2+}	Co^{2+}	Ni^{2+}	Cu^{2+}	Zn^{2+}
Al^{3+}		0	0.3	0	0	0.75	0.4	0
Cr^{3+}		0	0	0	0	0	0	0
Fe^{3+}		0.9	0.2	1	1	1	1	0
Mn^{3+}		0	0	0.67	0	1	0	0
Co^{3+}		–	–	–	0	–	–	0

Source: Data reproduced from Greenwood, *Ionic Crystals, Lattice Defects and Nonstoichiometry*, Butterworths © 1968; Dunitz and Orgel, *Adv. Inorg. Radiochem.*, **2**, 1, © 1960.

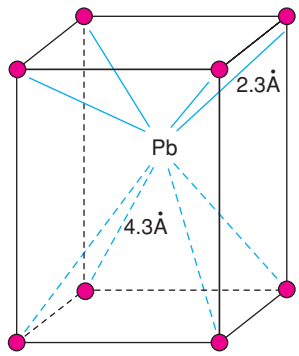


Figure 3.14 ***The structure of red PbO, showing the presence of the inert pair effect by the variation in Pb–O bond distances.***

Spinels are usually cubic but some show tetragonal distortions in which one cell edge is of different length to the other two. The Jahn–Teller effect causes such distortions in Cu^{2+}-containing spinels, $CuFe_2O_4$ (the tetragonal unit cell parameter ratio, $c/a = 1.06$) and $CuCr_2O_4$ ($c/a = 0.9$). $CuFe_2O_4$ is an inverse spinel in which the Jahn–Teller effect distorts the CuO_6 octahedra so that the two Cu–O bonds along z are longer than the four Cu–O bonds in the xy plane. On the other hand, $CuCr_2O_4$ is normal and the CuO_4 tetrahedra are flattened in the z direction, again due to the Jahn–Teller effect, causing a shortened c axis.

3.2.15.2 Inert pair effect

The heavy, post-transition elements, especially Tl, Sn, Pb, Sb and Bi, commonly exhibit a valence that is two less than the group valence (e.g. the divalent state in Group IV elements, Sn and Pb, compared with the group valence of 4). This so-called *inert pair effect* manifests itself structurally by a distortion of the metal ion coordination environment. Pb^{2+} has the configuration (Xe core) $4f^{14}5d^{10}6s^2$, and the $6s^2$ pair is 'stereochemically active' since these electrons are not in a spherically symmetrical orbital but stick out to one side of the Pb^{2+} ion (perhaps in some kind of s–p hybridised orbital).

Various kinds of distorted polyhedra occur as a consequence. Sometimes, the lone pair comes between the metal ion and some of its immediate anionic neighbours and causes a variation of bond length about the metal ion, e.g. red PbO has a structure that is a tetragonal distortion of the CsCl structure, Fig. 3.14. Four oxygens are situated at 2.3 Å from the Pb^{2+} ion, which is a reasonable Pb–O bond length, but the other four are at 4.3 Å. Although the lone pair is not directly visible, its presence is apparent from the distorted nature of the cubic coordination of Pb^{2+}.

A related effect occurs in SnS, which has a distorted rock salt structure. The SnS_6 octahedra are distorted along a [111] direction such that three sulfur atoms on one side are at ~2.64 Å but the other three are repelled by the lone pair to a distance of ~3.31 Å.

Another common distortion occurs when the lone pair simply takes the place of an anion and its associated pair of bonding electrons. Five-coordinated structures result in, for example, TlI, in which one corner (anion) of the octahedron 'TlI_6' is missing.

3.3 Covalent Bonding

The two main approaches used to describe and explain covalent bonding are molecular orbital, MO, theory and valence bond, VB theory. Each has special features but neither is able to explain fully all aspects of covalent

bonding. Coordination numbers and shapes are explained simply and predictively using a combination of hybridisation ideas and valence shell electron pair repulsion, VSEPR, but these are unable to account for various magnetic properties or electronic spectra. By contrast, MO theory is able to account for magnetic behaviour and electronic transitions but is unable to rationalise and predict molecular shapes. The main features of MO and VB theories are described but first, some key points on the electronic structure of atoms are summarised.

3.3.1 Particle-wave duality, atomic orbitals, wavefunctions and nodes

Sometimes, it is convenient to regard electrons as small, hard particles with definite mass. More often, they are regarded as having both particle and wave characteristics in which their momentum, mv, (m = mass, v = velocity) and wavelength, λ, are linked by the de Broglie relation:

$$\lambda = h/mv \tag{3.33}$$

and Planck's constant, h ($= 6.6261 \times 10^{-33}$ J s). Thus, electrons can diffract, just like other wave motions. They are never exclusively particles or waves, however, as summarised in one aspect of *Heisenberg's uncertainty principle: it is not possible to determine the position and momentum of electrons simultaneously.*
For instance, in order to 'see' an electron, it is necessary to look at it with 'light' of appropriate wavelength, but the electron is so small that short-wavelength, high-energy radiation would be required to see it (from $E = h\nu = hc/\lambda$) and energy transfer between the radiation and the electron would occur, changing the momentum of the electron. Simply put, looking at an object as small as an electron changes its momentum. Heisenberg's uncertainty principle is given by the probability expression

$$\text{(uncertainty in position)} \times \text{(uncertainty in momentum)} \geq h/4\pi \tag{3.34}$$

This means that the consequence of increased efforts to pinpoint an electron is an increase in its momentum, or vice versa, attempts to slow an electron effectively mean that it could be anywhere.

Given the limitations to which the position and momentum of an electron can be determined, it is usual to talk about probabilities of finding an electron in a particular place, which are quantified as *electron density* distributions. Electrons in atoms are located in *atomic orbitals*, but for each orbital all that one can state is the orbital shape and the probability of finding an electron at any point in the orbital.

Orbital shapes and electron density distributions are determined mathematically as solutions to *Schrödinger's equation*, which describes the wave properties of electrons and accounts for quantisation of energy levels (necessary, for instance, to explain atomic spectra in which only certain energies are allowed). Quantisation of wave motion and energy is, in fact, a familiar everyday phenomenon. With stringed musical instruments, only certain frequencies, or energies, of the resulting sound waves are allowed. By tightening the strings, this fixes the *boundary conditions* under which higher order harmonics can be heard. In the same way, electrons in atoms occupy orbitals with specific energies. The Periodic Table of the elements and the description of their electronic configurations and energies are a direct consequence of the wave nature of electrons and allowed solutions to Schrödinger's equation.

We shall not delve into the mathematical complexities of Schrödinger's equation but shall simply use some of the key results and consequences. The *sign of the wavefunction*, ψ, represented as either + or –, dictates whether the wave funtions of adjacent electrons are able to interact constructively as in bonding orbitals or destructively as in antibonding orbitals (there are clear analogies here with diffraction of, for example, light or X-rays). Note that the signs, + and –, do not refer to electronic charge but can be likened to the relative phase of a wave motion, which changes sign every 180°.

The electron density at any point is given by the square of the wavefunction, ψ^2 (just as with any wave motion, its intensity is given by the square of the wave amplitude). In the simplest case of a hydrogen atom

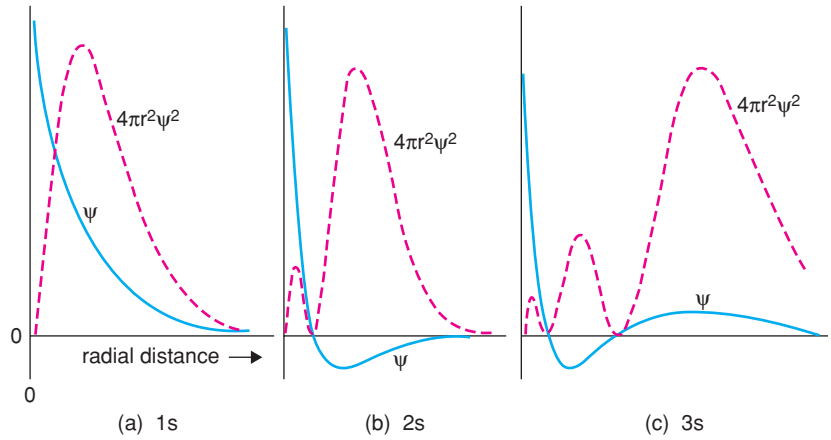


Figure 3.15 ***Plot of ψ and $4\pi r^2\psi^2$ for 1s, 2s and 3s orbitals of a hydrogen atom. Adapted with permission from* Chemical Bonding*, Mark J. Winter, © 1994 Oxford University Press.***

with a single electron and for which the Schrödinger equation can be solved exactly, the variations of ψ and ψ^2 as a function of distance from the nucleus are shown schematically in Fig. 3.15 for an electron in either the ground state 1*s* orbital or in excited state 2*s*, 3*s* orbitals. ψ^2 is multiplied by $4\pi r^2$ (which is the surface area of a sphere of radius *r*) to represent the total electron density at distance *r* from the nucleus. In the excited state orbitals, the wavefunction changes sign at a certain distance from the nucleus, at which point, known as a *node*, the electron density decreases to zero. Comparing the different *s* orbitals, the electron density for the 1*s* orbital passes through a maximum at a certain distance from the nucleus. For the 2*s* orbital, there are two maxima separated by a node and the outermost maximum contains most of the electron density. The plot of electron density as a function of distance from the nucleus is also known as the *radial distribution function*, RDF.

Although the electron density in an RDF is either positive or zero, the sign of the wavefunction changes at a node and this provides the key to the resulting energies when atomic orbitals, AOs, on adjacent atoms interact to give molecular orbitals, MOs. For *s* orbitals which are spherically symmetrical, the sign of the wavefunction is the same for a given distance from the nucleus, but this is not true for *p*, *d* and *f* orbitals, which are non-spherical. In these orbitals, the electron density is concentrated in certain regions, called *lobes*, and different lobes for a given orbital may have different signs as shown schematically in Fig. 3.16.

3.3.2 Orbital overlap, symmetry and molecular orbitals

Covalent bonding results when AOs on adjacent atoms overlap significantly to form MOs that extend over two or more atoms. If the signs of the wavefunctions and the AOs are the same in the overlap region, *bonding orbitals* result. If they are of opposite sign, the resulting MOs are *antibonding*. If the interactions are a mixture of same sign and opposite sign, or if the degree of overlap is very small, the resulting orbitals are *non-bonding*. Some general rules are that:

1. The number of resulting MOs is the same as the number of contributing AOs.
2. The energies of bonding MOs are lower than those of the contributing AOs whereas the energies of the antibonding orbitals are correspondingly higher. Non-bonding orbitals have essentially the same energy as the individual AOs.

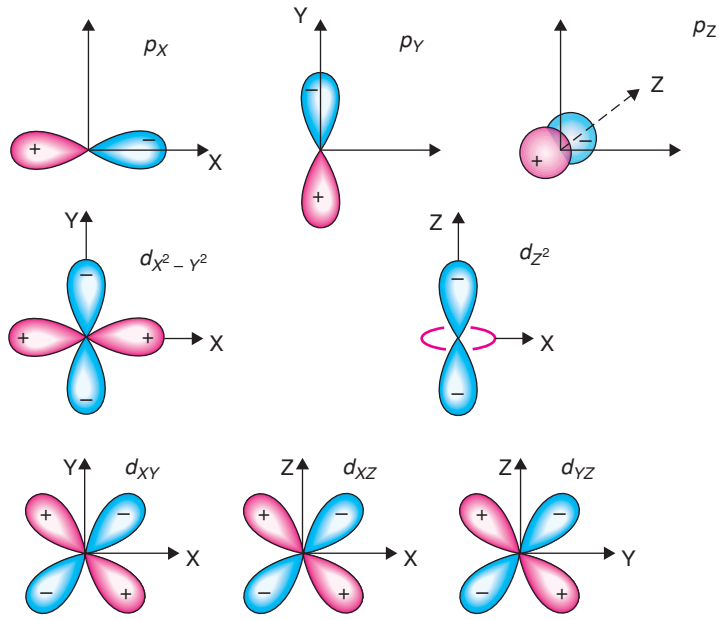


Figure 3.16 ***The three* p *and five* d *orbitals. Note the axes for labelling each orbital.***

3. MOs in which the associated electron density is concentrated in the space between adjacent atoms are referred to as σ orbitals (and σ^* for associated antibonding orbitals) whereas MOs for which the principal overlap regions are not along the line connecting atomic nuclei are referred to as π (and π^*) orbitals.

Before considering covalent bonding in solids, it is very useful and instructive to consider bonding in some small molecules. The simplest example is shown in Fig. 3.17 for overlap of the 1*s* AOs on two hydrogen atoms to form the H_2 molecule. Two electrons, one from each H atom, occupy the σ 1*s* orbital whereas there are no electrons in the σ^* 1*s* orbital. There is a node between lobes of the σ^* 1*s* orbital which means that even if electrons were available to occupy this orbital, there would be no electron density in the overlap region.

Another way of thinking about the energetics of bonding and antibonding orbitals is that electrons in these orbitals may or may not shield the positively charged atomic nuclei from each other. The shielding is effective for σ orbitals, partially effective for π orbitals (later) and ineffective for antibonding orbitals. This is reflected in the resulting energies of electrons in these orbitals (orbitals do not have any physical substance themselves, especially when they are empty, but simply provide a virtual structure with which to describe the probability of finding electrons and their energies).

In molecules such as H_2 in which a bonding orbital contains a pair of electrons but the corresponding antibonding orbital is empty, we may say that a single covalent bond is present between the hydrogen atoms. If, however, the bonding orbital contains only one electron (as in the H_2^+ molecular ion) or if the associated antibonding orbital contains a single electron, which partially offsets the pair of electrons in the bonding orbital (as in the hypothetical H_2^- ion), then the resulting *bond order* would equate to half a single bond. If both the antibonding and bonding orbitals contained two electrons, there would be no effective net bonding, the molecule would have zero stability and would probably not exist (as in the case of non-existent He_2) unless the atoms were held together by van de Waals bonds. Bond orders >1 are possible in which both σ

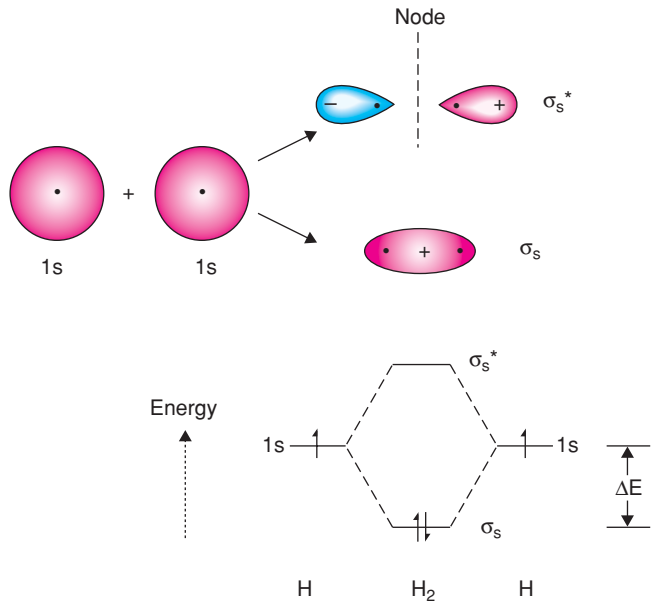


Figure 3.17 ***Bonding σ_s and antibonding σ_s^* MOs on H_2, and their relative energies, formed from two 1s AOs on the individual H atoms.***

bonding and π bonding orbitals contain electrons but without any electrons in the associated antibonding orbitals.

Bonding and antibonding MOs for the *p* orbitals with their lobes pointing along the line connecting adjacent nuclei (referred to as the *z* direction) are referred to as σ_p and σ_p^*, Fig. 3.18(a) and (b). For the p_x and p_y AOs (which are oriented perpendicular to the internuclear line), the resulting MOs are π_p (c) and π_p^* (d).

For molecules in which both *s* and *p* outer shell AOs are involved in bonding, the schematic energy level diagram shown in Fig. 3.19 results. The σ_s and σ_s^* orbitals are of lowest energy and would be occupied first by electrons. Next comes the σ_p and π_p bonding orbitals followed by their respective antibonding orbitals. The π orbitals are *doubly degenerate* since the two sets of orbitals formed by interaction of the p_x and p_y AOs have the same energies.

To evaluate the bond order, the combined total of available valence shell electrons in the *s* and *p* orbitals on the two atoms that form the MOs are fed into the MOs, two at a time, starting from the σ_s orbital. For example, in the case of fluorine, F_2, which has the electronic configuration (He core) $2s^2\ 2p^5$, there are 14 valence shell electrons available and these occupy the MOs in Fig. 3.19(a) in the sequence shown. The highest occupied MOs are the π_p^* MOs leaving the σ_p^* MO empty. The net effect is that the F_2 molecule is held together by a single covalent bond with a pair of electrons in the σ_p orbital.

The energy level diagram for the oxygen, O_2, molecule is shown in Fig. 3.19(b) but there are two electrons less to occupy the MOs [oxygen is (He core) $2s^2\ 2p^4$]. Following *Hund's rule of maximum multiplicity*, the *highest occupied molecular orbitals*, HOMOs, are the two π_p^* MOs which have only one electron each. Consequences of this are that first, the O_2 molecule has a bond order of 2, and is said to be double bonded, and second, O_2 is paramagnetic due to the unpaired electrons in the π_p^* orbitals. The resulting bond orders have a direct effect on bond strengths and bond lengths. Thus, the bond energy for doubly bonded O_2 is 493 kJ mol^{-1} with bond length 1.207 Å, whereas singly bonded F_2 has a bond energy 155 kJ mol^{-1} and a bond length 1.412 Å.

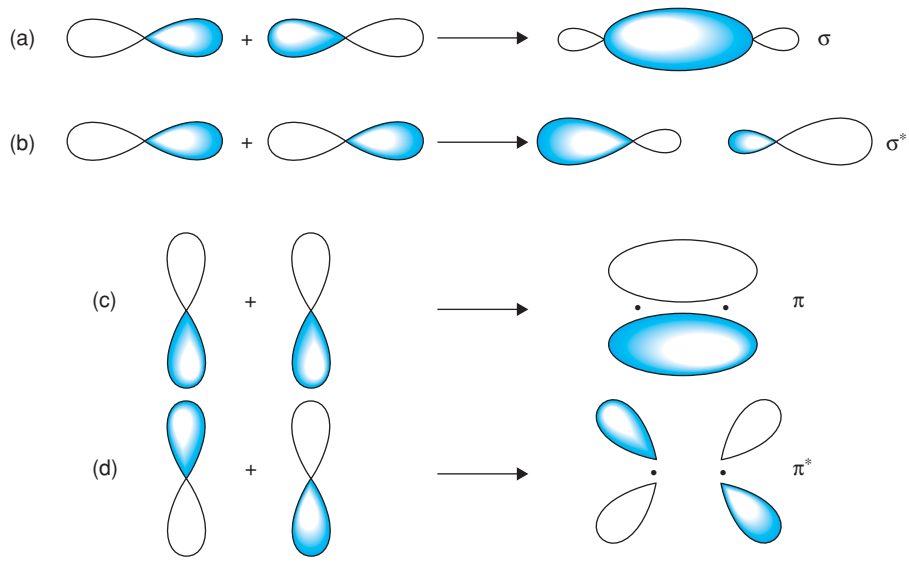


Figure 3.18 ***Overlap of 2p orbitals to give (a) σ_p bonding orbital, (b) σ_p^* antibonding orbital, (c) π bonding orbital and (d) π^* antibonding orbital.***

For the energy level diagrams shown in Fig. 3.19, there is no interaction between *s* and *p* orbitals. In many cases, interaction between *s* and *p* orbitals does occur, however, as shown by detailed calculations of orbital energies which give the sequence of MOs on the energy scale and by various electronic properties. The idea of orbital mixing also forms the basis of *hybridisation* (next section).

The interaction of σ_s and σ_p MOs and their resulting energies is shown schematically in Fig. 3.20(a). It is no longer appropriate to refer to the new orbitals as bonding and antibonding and they are therefore labelled simply as σ_1 and σ_2. Similarly, σ_p^* and σ_s^* MOs can mix to give two new hybrid orbitals σ_3 and σ_4, Fig. 3.20(b). π MOs are not involved in such mixing interactions.

The resulting energy level diagram, which is appropriate for molecules such as N_2, is shown in Fig. 3.21. The sequence of energy levels is now different to that shown in Fig. 3.19 in that the σ_3 orbital has higher energy than the π_p orbitals. This does not affect the bond order of N_2, which is 3 (two electron pairs in the π_p orbitals and one pair in the σ_3 orbital); π^* and σ_4 orbitals are empty. The bond energy of N_2 is 942 kJ mol^{-1} and the bond length 1.098 Å.

We next consider diatomic molecules in which the two atoms are not the same. Consequently, the energies of their individual AOs are different, orbital mixing of the resulting MOs occurs to various extents and the resulting energy level diagrams are often asymmetric. In addition, if the two atoms have significantly different electronegativities, bonding is a mixture of covalent and ionic.

A good example of a polar covalent bond is in the HF molecule, for which the energy levels based on MO theory are shown in Fig. 3.22. This illustrates a number of features, which are important in bonding of solids such as oxides, as follows:

- The energies of the participating AOs depend very much on the position in the Periodic Table and, in particular, how much the electrons in these AOs are shielded from the nuclear charge. The hydrogen 1*s* electron experiences unit positive charge from the hydrogen nucleus and is of higher energy than the 2*s*, 2*p*

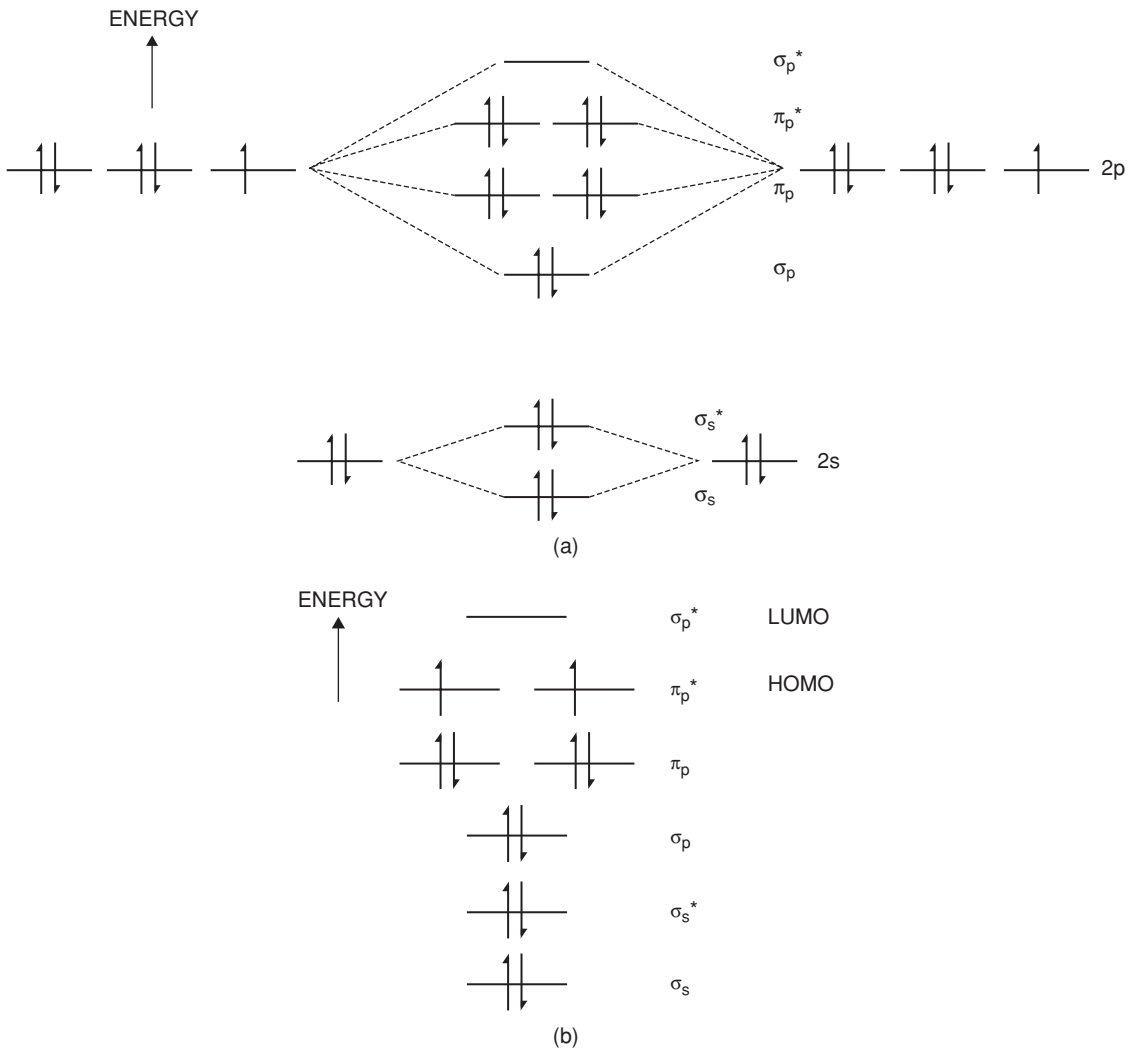


Figure 3.19 ***Energy level diagram for MOs on (a) the F_2 molecule formed from overlap of valence shell 2s and 2p orbitals and (b) O_2. HOMO, highest occupied molecular orbital; LUMO, lowest unoccupied molecular orbital.***

electrons of fluorine, which are not well shielded by the inner 1*s* electrons from the nuclear charge of 9+. The 2*p* orbitals are partially shielded from this nuclear charge by the 2*s* electrons, whose orbitals penetrate more deeply towards the nucleus and hence the 2*p* electrons have higher energy than the 2*s* electrons.

- F is a much more electronegative element than H and so the bonding electrons are strongly attracted towards F, giving a net negative charge $\delta-$ on F and an associate positive charge $\delta+$ on H.
- The 2*s* electrons on F do not take any effective part in MO formation since they are rather tightly held to the F nucleus and hence remain in an individual AO of the same energy in the MO diagram, Fig. 3.22. These may therefore be regarded as occupying a non-bonding orbital.

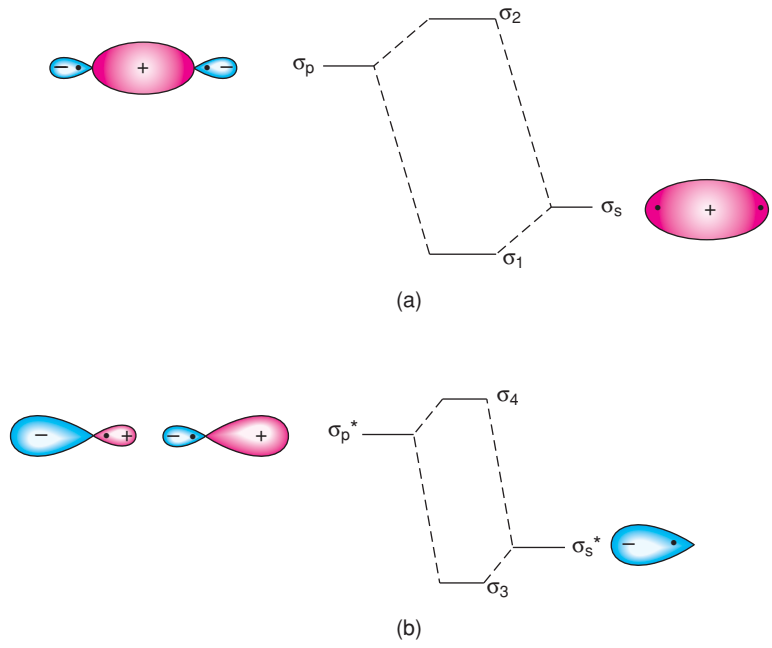


Figure 3.20 *The mixing of (a) σ_p and σ_s MOs and (b) σ_p^* and σ_s^* MOs to give four new MOs: σ_1, σ_2, σ_3 and σ_4.*

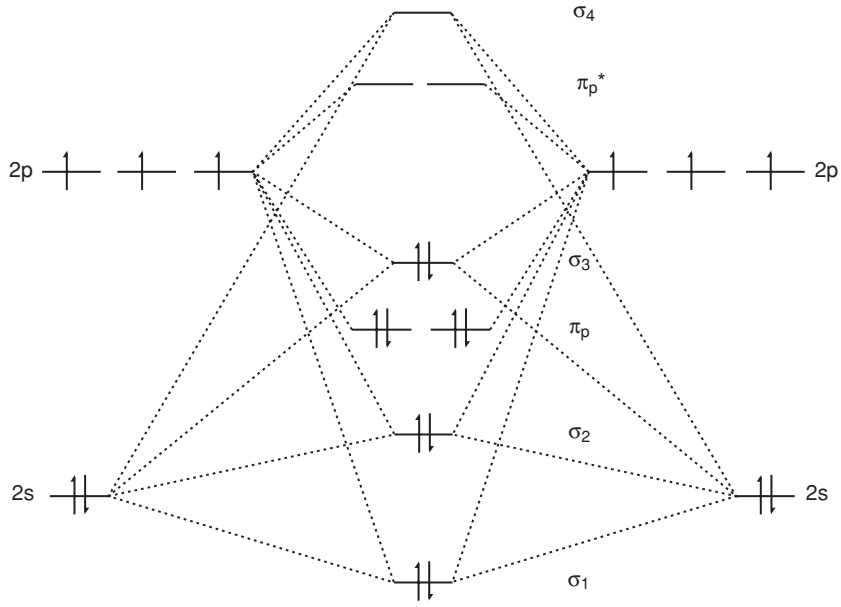


Figure 3.21 *Energy level diagram for N_2 showing the four hybrid σ MOs, σ_1, σ_2, σ_3 and σ_4 together with π, π^* orbitals.*

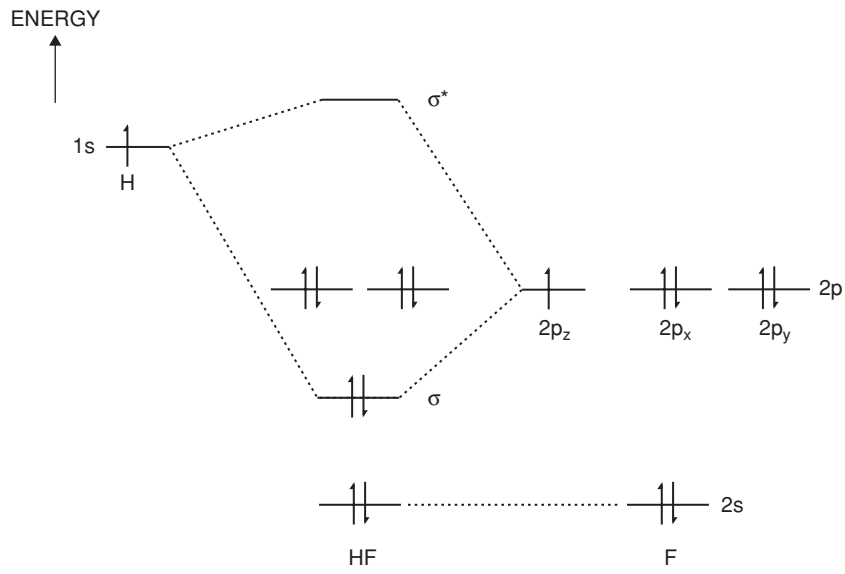


Figure 3.22 ***Energy level diagram for the HF molecule.***

- The σ bonding orbital formed by overlap of H 1*s* and F $2p_z$ orbitals is of lower energy than the contributing AOs but is closer in energy to that of the F $2p_z$ orbital. We may say that this σ orbital retains strong *p* characteristics of F.
- Conversely, the σ^* orbital has rather similar energy to the contributing H 1*s* AO and therefore retains strong s characteristics.
- The $2p_x$ and $2p_y$ AOs on F do not interact with the H 1*s* orbital and therefore remain as non-bonding orbitals in the energy level sequence.

MO theory has been very successful in accounting for the bonding, spectroscopic and magnetic properties of molecules and, as we shall see later, can be extended to crystalline solids and, in modified form, to metallic bonding. Next we consider an alternative approach, *valence bond theory*, which has many attractive features, not least its simplicity in relating bonding to coordination geometries and oxidation states.

3.3.3 Valence bond theory, electron pair repulsion, hybridisation and oxidation states

The valence bond approach to covalent bonding focuses on the outer shell valence electrons of atoms. Basically, they can do one of four things:

1. They form electron pair, single covalent bonds by the sharing of one electron each from the two atoms that form the bonds.
2. If two electrons from the same atom occupy an AO, they form a *lone pair* in a lobe-shaped non-spherical orbital. This lone pair generally occupies a region of space instead of a bonding pair; the lone pair is described as *stereochemically active*.
3. Two electrons from the same atom may act to form a single covalent bond between two atoms; such a bond is called a *dative bond*.

Table 3.16 ***Coordination numbers and geometries***

No. of electron pairs[a]	Ideal shape	Examples
2	Linear	CO_2
3	Triangular	BCl_3
4	Tetrahedral	CH_4, SiO_4^{4-}
5	Trigonal bipyramidal	PCl_5
6	Octahedral	SF_6, AlF_6^{3-}

[a]Single covalent bonding pairs and lone pairs.

4. Pairs of electrons, one from each atom, may take part in multiple bond formation to add to the single bond between the same pair of atoms. Double bonds, triple bonds and occasionally even quadruple bonds are formed in this way.

Coordination numbers and geometries around a central atom are given by the combined total of covalent bonding pairs and lone pairs of electrons, working on the principle that electron–electron repulsion causes these pairs of electrons to arrange themselves to be as far apart as possible in space. The ideal shapes, with some examples for coordination numbers ranging from 2 to 6, are shown in Table 3.16. However, lone pairs are held more tightly to a nucleus than bonding pairs, which effectively means that they occupy more space and cause distortions from ideal symmetries. The strength of electron pair repulsions is in the following sequence:

lone pair–lone pair > lone pair–bonding pair > bonding pair–bonding pair

Two familiar examples of the distorted shapes of small molecules are NH_3 and H_2O, Fig. 3.23. NH_3 has one lone pair on N and three N–H covalent bonding pairs but the lone pair exerts a greater influence on the shape and the H–N–H bond angles are reduced from the ideal values of 109.5° to 107°. In H_2O, again the H–O–H bond angle is reduced from 109.5° to 104.5°.

The above section gives a brief summary of valence shell electron pair repulsion, VSEPR, theory, which is widely used to account for the shapes of molecules. Multiple bonds do not have a major influence on shapes since their geometric disposition is essentially the same as that of a single bond; nevertheless, since they contain more than two electrons (four for a double bond, six for a triple bond) they do exert a greater repulsion on adjacent electron pairs.

From our knowledge of the shapes and dispositions of AOs on particular atoms, it quickly becomes apparent that modification of some of the AOs is required to account for VSEPR geometries. For instance, in the methane, CH_4, molecule, the three 2*p* AOs and the 2*s* AO on C are not expected to be arranged spatially

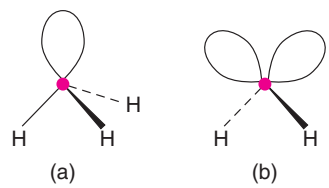


Figure 3.23 ***Shapes of the (a) NH_3 and (b) H_2O molecules.***

Table 3.17 Orbital hybridisation and resulting shapes

Hybridisation	Shape
$s + p \rightarrow$ two sp hybrids	Linear
$s + 2p \rightarrow$ three sp^2 hybrids	Triangular
$s + 3p \rightarrow$ four sp^3 hybrids	Tetrahedral
$s + 3p + d \rightarrow$ five sp^3d hybrids	Trigonal bipyramidal
$s + 3p + 2d \rightarrow$ six sp^3d^2 hybrids	Octahedral
$s + 3p + 3d \rightarrow$ seven sp^3d^3 hybrids	Pentagonal bipyramidal

in the form of a tetrahedron because the 2*p* AOs are at right-angles to each other and the 2*s* AO is spherically symmetrical. To account for the changes in orbital structure that are required to permit formation of four equal C–H bonds in methane, and many other similar cases, the concept of *hybridisation* was introduced. In hybridisation, *s* and *p* (and *d*) AOs are combined to give an equal number of hybrid AOs, which are then able to overlap with AOs on adjacent atoms leading to covalent bond formation.

The distribution of lobe-shaped hybrid orbitals that are postulated to form on hybridisation of different numbers of *s*, *p* and *d* AOs are summarised in Table 3.17. The shapes are the same as those given in Table 3.16 and the two tables are self-consistent since the number of occupied hybrid orbitals governs the coordination geometry.

To complete the process of assigning electrons to either bonding pairs or lone pairs, knowledge of the Periodic Table, element ground states and possible excited states and the usual oxidation state adopted by the elements in compound formation is required. To describe the electronic ground states of elements, the simplified scheme used here is as follows: only those electrons that form part of the outer valence shell are specified. The inner, core electrons are simply lumped together in the configuration expected for the appropriate noble gas element. For example, the first row *p*-block elements may be represented as (He core) $2s^2\ 2p^n$ where $n = 1$ (B), 2 (C), 3 (N), 4 (O), 5 (F). The first row transition series may be represented as (Ar core) $4s^2\ 3d^n$ where $n = 1$ (Sc), 2 (Ti), etc., and so on.

An old-fashioned but simple, pictorial and useful way to represent valence shell electrons is by their occupancy of a series of boxes, with one box for each AO, as summarised for some examples in Fig. 3.24. Of course, electrons do not occupy boxes, but this scheme does allow us to distinguish between lone pairs and unpaired electrons and between ground states and excited states, and also to identify which empty orbitals could be occupied by acceptance of an electron pair donated from an adjacent atom to form a dative bond.

In example 1, carbon has two unpaired electrons in 2*p* orbitals and would therefore be divalent unless electrons in the 2*s* orbital could be used in bond formation. To obtain the normal tetravalent state of carbon, one of the 2*s* electrons is regarded as being promoted to the empty 2*p* orbital, giving four orbitals each with a single electron. Following hybridisation of these four orbitals to give four equal sp^3 hybrid orbitals arranged tetrahedrally, carbon is able to form four electron-pair covalent bonds by accepting an electron into each of the sp^3 hybrid orbitals from surrounding atoms, as in CH_4.

Example 2 is phosphorus, which in its ground state would be trivalent since it has three half-occupied 3*p* orbitals; many trivalent phosphorus compounds are known, which are held together by three electron-pair covalent bonds. The shapes of these molecules are usually pyramidal rather than planar, however, since the lone pair of electrons originating from the 3*s* orbital occupies a stereochemically-active sp^3 hybrid orbital as a consequence of hybridisation; the resulting bonding pair–lone pair repulsions give rise to the pyramidal shape.

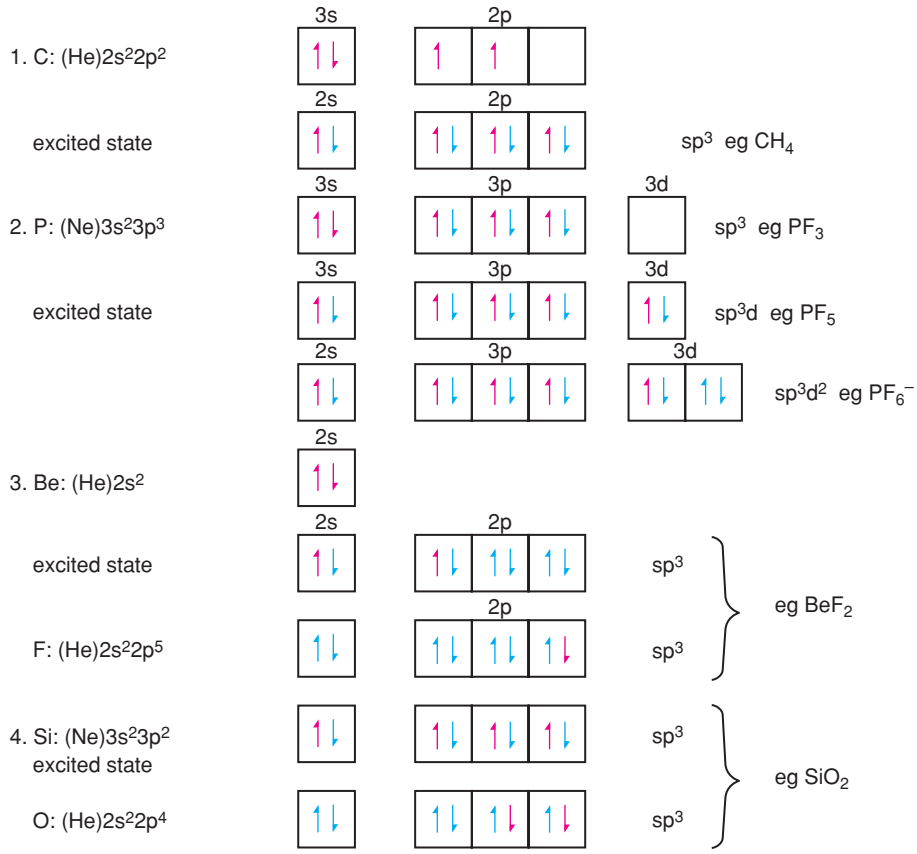


Figure 3.24 ***Valence bond descriptions of some inorganic compounds and structures.***

An additional feature of compound formation by phosphorus is that *orbital expansion* can occur in which one of the *s* electrons is promoted to an empty 3*d* orbital, giving rise to five half-filled orbitals. This allows phosphorus to exhibit either the group valence of five by forming five electron-pair covalent bonds or a valence of three that is two less than the group valence.

An additional complication with phosphorus is that a second 3*d* orbital is potentially able to accept an electron pair from a negatively charged ligand such as F^-, forming an octahedral complex such as PF_6^-, in which the sp^3d^2 hybridisation scheme gives six equal hybrid orbitals and an octahedral shape.

Example 3 is beryllium, for which one of the 2*s* electrons is promoted to an empty 2*p* orbital to give a valence state of two but the two empty 2*p* orbitals are able to accept electron pairs from negatively charged donors giving sp^3 tetrahedral coordination. The bonding in crystalline BeF_2, which is isostructural with SiO_2 and has a 3D framework structure formed from corner-sharing BeF_4 tetrahedra, may be rationalised using this valence bond scheme. Thus, each F atom forms a single covalent bond by sharing one electron with a Be electron and a dative bond by donating an electron pair into an empty sp^3 hybrid orbital on a second Be atom. Hence F is two-coordinate but the Be–F–Be linkage is non-linear because F has two lone pairs in addition to two bonding pairs.

Example 4 is SiO_2 with a similar structure to BeF_2 but, in this case Si in the excited state has four half-occupied sp^3 orbitals, each of which shares an electron with a surrounding oxygen. Each oxygen has two

half-occupied orbitals for covalent bond formation with Si but again the Si–O–Si bond is non-linear because of the two lone pairs on each O atom.

Not all inorganic crystal structures can be described in terms of electron-pair covalent bonds, but SiO_2 and BeF_2 are two good examples since, with the low coordination numbers of Si, Be and O, F, there are sufficient electrons for each linkage to be treated as an electron-pair covalent bond.

It is interesting that the hybridisation concept which forms the basis of VB theory has many similarities to the ideas of orbital mixing that is a key part of MO theory. VB theory is attractive since it allows us to rationalise and predict coordination geometries, including the presence of stereochemically active lone pairs. It includes dative bond formation in which both electrons in a covalent bond are provided by a more electronegative atom. It also relates the number of lone pairs and bonding electron pairs directly to element oxidation states. VB theory does not, however, have anything to say about the spectroscopic properties associated with electronic transitions between energy levels or magnetic properties associated with unpaired electrons and for which MO theory is much more useful. In addition, the ideas of MO theory as applied to small molecules can be readily extended to the bonding in those crystal structures which exhibit metallic bonding, as described in the next section.

3.4 Metallic Bonding and Band Theory

Metallic structures and bonding are characterised by delocalised valence electrons and these are responsible for the high electrical conductivity of metals. This contrasts with ionic and covalent bonding in which the valence electrons are localised on particular atoms, ions or molecular orbitals and are not free to migrate through the structure. The bonding theory used to account for delocalised electrons is the band theory of solids. We shall see qualitatively how band theory can be applied to certain groups of solids, especially metals and semiconductors.

In a metal such as Al, the inner core electrons – 1*s*, 2*s* and 2*p* – are localised in discrete orbitals on the individual Al atoms. However, the 3*s* and 3*p* electrons that form the valence shell occupy levels that are delocalised over the entire metal crystal. These levels are like giant MOs, each of which can contain up to two electrons. In a metal there must be an enormous number of such levels which are separated from each other by very small energy differences. Thus in a crystal of Al that contains *N* atoms, each contributes one 3*s* orbital and the result is a band of *N* closely spaced energy levels. This band is called the 3*s* *valence band*. The 3*p* levels similarly form a delocalised 3*p* band of energy levels.

The band structure of other materials may be regarded similarly. The differences between metals, semiconductors and insulators are shown schematically in Fig. 3.25 and depend on:

(a) the band structure of each;
(b) whether the valence bands are full or only partly full;
(c) the magnitude of any energy gap between full and empty bands.

The band theory of solids is well supported by results obtained using X-ray and photoelectron spectroscopy and by two independent theoretical approaches.

First, the ‘chemical approach’ is to take MO theory, as it is usually applied to small, finite-sized molecules, and extend it to infinite, 3D structures. In the application of MO theory of diatomic molecules, an AO from atom 1 overlaps with an AO on atom 2, resulting in the formation of two MOs that are delocalised over both atoms. One MO is bonding and has lower energy than that of the AOs. The other is antibonding and is of higher energy.

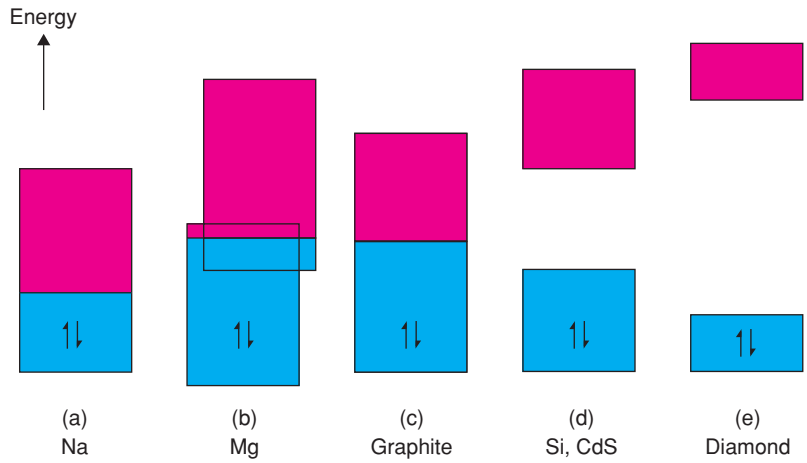


Figure 3.25 ***Band structure of (a) and (b) metals, (c) semimetals, (d) semiconductors and (e) insulators. Purple shading: empty energy levels; blue: occupied levels.***

Extension of this approach to larger molecules leads to an increase in the number of MOs. For each AO that is put into the system, one MO is created. As the number of MOs increases, the average energy gap between adjacent MOs must decrease until, essentially, a continuum of energy levels occurs. The gap between bonding and antibonding orbitals may also decrease and disappear entirely for outer shell orbitals.

Metals may therefore be regarded as infinitely large 'molecules' in which an enormous number of energy levels or 'molecular' orbitals is present, $\sim 6 \times 10^{23}$ for 1 mol of metal. It is no longer appropriate to refer to each level as a 'molecular' orbital since each is delocalised over all atoms in the crystal. Instead, they are usually referred to as energy levels or energy states.

The band structure of Na, calculated using the *tight binding* approximation, is shown in Fig. 3.26. The width of a particular energy band is related to the interatomic separation and hence to the degree of overlap between orbitals on adjacent atoms. Thus, at the observed interatomic separation, r_0, the 3*s* and 3*p* orbitals on adjacent atoms overlap to form broad 3*s* and 3*p* bands (shaded). The upper levels of the 3*s* band with energies in the range C to B have similar energies to the lower levels of the 3*p* band. Hence there is no discontinuity in energy between 3*s* and 3*p* bands. Such overlap of bands is important in explaining the metallic properties of elements such as the alkali and alkaline earth metals and Al, as shown in Fig. 3.25(b).

At the interatomic distance in Na metal, r_0, the 1*s*, 2*s* and 2*p* orbitals on adjacent Na atoms do not overlap. Instead, they remain as discrete AOs associated with individual atoms. They are represented in Fig. 3.26 as thin lines. If it were possible to reduce the internuclear separation from r_0 to r' by compression, then the 2*s* and 2*p* orbitals would also overlap to form bands of energy levels (shaded). The 1*s* levels would, however, still be present as discrete levels at distance r'. Similar effects are likely in other elements subjected to high pressure. For instance, it has been calculated that hydrogen would become metallic at a pressure $\geq 10^6$ atm due to overlap of 1*s* AOs to form a 1*s* band.

Sodium has the electronic configuration (Ne core) $3s^1$ and has one valence electron per atom. Since the 3*s* and 3*p* bands overlap, Fig. 3.26, the valence electrons are not confined to the 3*s* band but are distributed over the lower levels of both the 3*s* and 3*p* bands.

Second, the 'physical approach' to band theory is to consider the energy and wavelength of electrons in a solid. In the early *free electron theories* of Sommerfeld and Pauli, a metal is regarded as a potential well, inside which the more loosely held valence electrons are free to move. The energy levels that the electrons occupy are quantised (analogy with the quantum mechanical problem of a particle in a box) and the levels are filled from the bottom of the well with two electrons per level. The highest filled level at absolute zero is

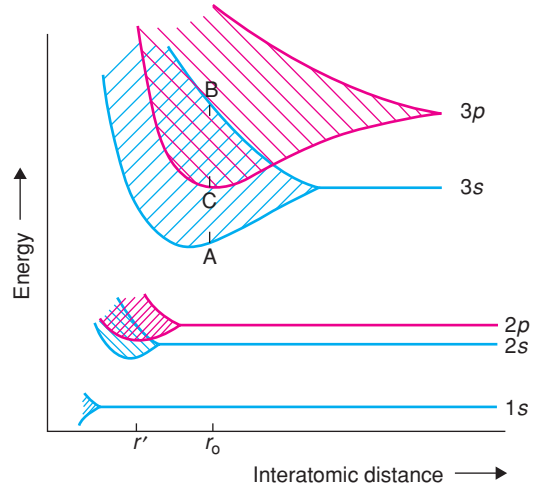


Figure 3.26 ***Effect of interatomic spacing on atomic energy levels and bands for Na, calculated using tight binding theory. Shaded areas represent bands of energy levels formed by significant overlap of AOs on adjacent atoms.***

known as the *Fermi level*. The corresponding energy is the *Fermi energy*, E_F, Fig. 3.27(a). The *work function*, ϕ, is the energy required to remove the uppermost valence electrons from the potential well. It is analogous to the ionisation potential of an isolated atom.

The energy levels in the potential well may be calculated using the 'particle in a box' description of electrons as standing waves whose wavelengths may be related to the box dimensions. Specifically, the box length, l, should be an integral number of electron half-wavelengths, as shown for the three wavelengths of lowest energy in Fig. 3.27(b). With this constraint, the waves for a given energy are in-phase and, therefore, avoid destructive interference. The wavelengths are given by

$$\lambda = 2l/n : \; n = 1, 2, 3, \ldots \tag{3.35}$$

and the momentum, p, of the electrons by

$$\begin{aligned} p &= \mathrm{h}/\lambda \\ &= hn/2l \end{aligned} \tag{3.36}$$

The potential energy, PE, inside the box is zero and therefore the total electron energy is kinetic energy, KE. If electrons are regarded as particles whose energies are given by the kinetic theory of gases, then KE is given by:

$$\begin{aligned} E &= mv^2/2 \\ &= p^2/2m \end{aligned} \tag{3.37}$$

Using the definition of wavenumber, k, as

$$\begin{aligned} k &= 2\pi/\lambda \\ &= n\pi/l \end{aligned} \tag{3.38}$$

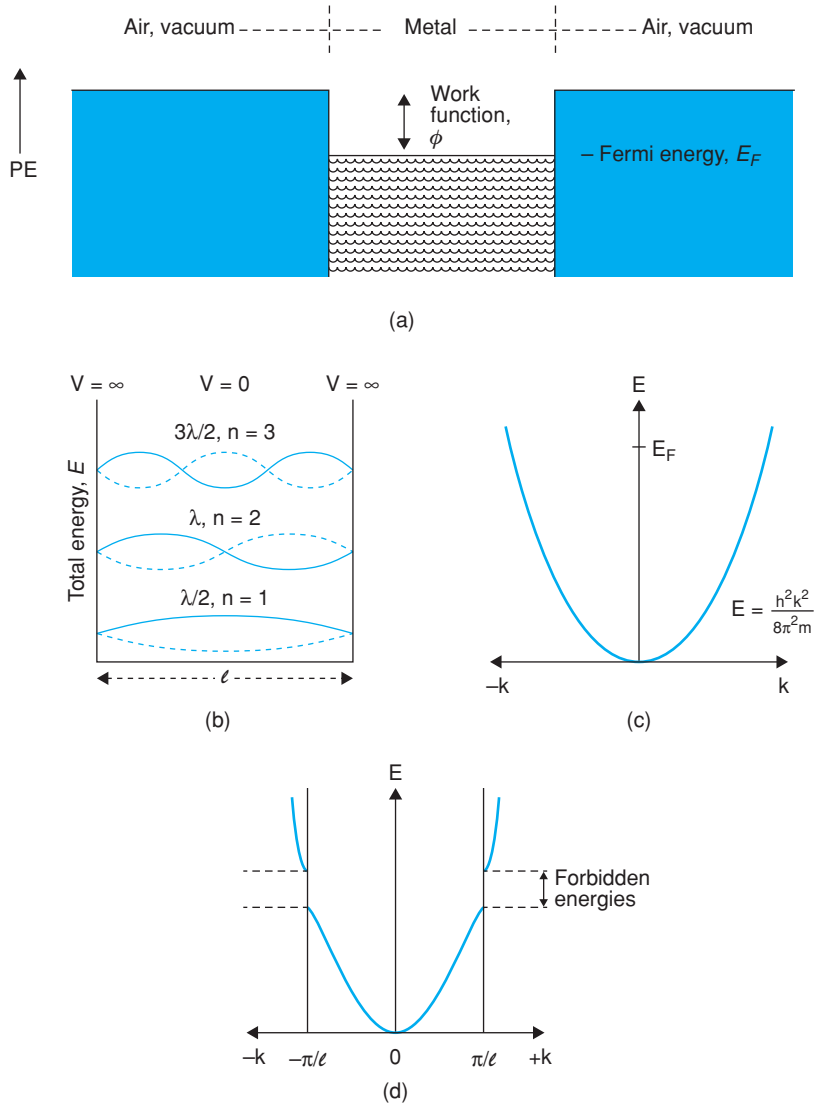


Figure 3.27 ***(a) Free electron theory of a metal; electrons in a potential well; (b) allowed wavelengths for electrons confined to a box of length l; (c) E–k parabolic relationship for a free electron in a solid lattice without perturbation from postively charged atomic nuclei; (d) the first Brillouin zone for a 1D crystal with interatomic spacing l.***

combining equations (3.36), (3.37) and (3.38) gives the relationship between E and either n, λ or k:

$$E = h^2/2\lambda^2 m = h^2k^2/8\pi^2 m = h^2n^2/8ml^2 \tag{3.39}$$

The E–k relationship, equation (3.39), takes the form of a parabola, as shown in Fig. 3.27(c). Figures 3.27(b) and (c) refer to the one-dimensional case of free electrons and in particular at the Fermi level, E_F, the energy

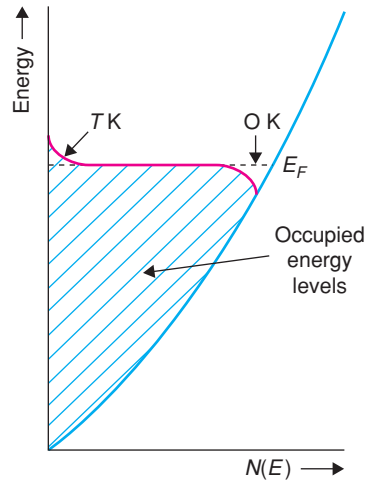


Figure 3.28 ***Density of states versus energy.***

is represented by two points on the positive and negative wings of the parabola (c). If there are N electrons in the solid, then $n = N/2$ and the Fermi energy, E_F, is given by

$$E_F = h^2N^2/32ml^2 \tag{3.40}$$

Extending the k-space to 3D, the two points at E_F become a Fermi surface which may be spherical for an isotropic conductor but can also adopt various non-spherical shapes. Equation (3.39) then becomes

$$E = h^2/8m\left(n_x^2/l_x^2 + n_y^2/l_y^2 + n_z^2/l_z^2\right) \tag{3.41}$$

An important diagram is the *density of states*, $N(E)$, diagram, which is a plot of the number of energy levels, $N(E)$, as a function of energy, E, Fig. 3.28. The number of available energy levels increases steadily with increasing energy in the free electron theory. Although the energy levels are quantised, there are so many and the energy difference between adjacent levels is so small that, effectively, a continuum occurs. At temperatures above absolute zero, some electrons in levels near to E_F have sufficient thermal energy to be promoted to empty levels above E_F. Hence at real temperatures, some states above E_F are occupied and others below E_F are vacant; what is particularly important is that a number of states near E_F are half occupied. The average occupancy of the energy levels at some temperature T above zero is shown as shading in Fig. 3.28.

The high electrical conductivity of metals is due to the drift of those electrons in half-occupied states close to E_F. Electrons in doubly occupied states lower down in the valence band cannot undergo net migration in a particular direction whereas those in singly occupied levels are free to move. The promotion of an electron from a full level below E_F to an empty one above E_F gives rise, effectively, to two mobile electrons. Measurements of the specific heat (or heat capacity) of metals as a function of temperature confirm that relatively few electrons (i.e. only those close to E_F) contribute to electrical conductivity at any one time since the magnitude of the specific heat can be accounted for almost entirely by that contribution due to the lattice vibrations. If all the electrons were mobile, a significantly higher specific heat would be expected.

The free electron theory is an oversimplification but is a very useful starting model. In the *nearly free electron theory*, the potential inside the crystal or potential well is regarded as periodic, Fig. 3.29, and not

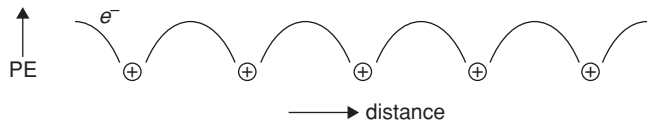


Figure 3.29 ***Potential energy of electrons as a function of distance through a solid.***

constant. The positively charged atomic nuclei are, of course, arranged in a regularly repeating manner in a crystal lattice. The potential energy of the electrons passes through a minimum at the positions of the nuclei, due to Coulombic attraction, and passes through a maximum midway between adjacent nuclei. Solution of the Schrödinger equation for a periodic potential function, as in Fig. 3.29, shows that an uninterrupted continuum of energy levels does not occur but, instead, only certain bands or ranges of energies are permitted for the electrons. The forbidden energies correspond to electron wavelengths that satisfy Bragg's law for diffraction in a particular direction in the crystal (see Section 5.3.4.2). Bragg's law relates interplanar separation, d, to the X-ray wavelength, λ, and the angles of incidence and reflection, θ, by

$$n\lambda = 2d \sin\theta \tag{3.42}$$

which, for normal incidence with $\theta = 90^\circ$, reduces to

$$n\lambda = 2d \tag{3.43}$$

This is the same equation as for standing waves that determine electron energy levels in the 1D 'particle in a box', of length l, Fig. 3.27(b). Effectively, this means that electrons whose wavelength satisfies equation (3.43) cannot propagate through a crystal and contribute to electrical conductivity but are constrained to be localised within interatomic distances. For electrons whose wavelengths do not quite satisfy equations (3.42) and (3.43), either because $\lambda \neq 2d/n$ or because $\sin\theta \neq 1$, theory shows that the associated energy levels are not available for long-range electron delocalisation and conduction. This leads to discontinuities in the E–k diagrams, Fig. 3.27(d), and forbidden energy levels or *band gaps* in $N(E)$–E diagrams, Fig. 3.30.

An alternative way to regard the energy levels is in terms of *k-space*. Electrons of lowest energy, corresponding to $n = 1, 2, 3 \ldots$, Fig. 3.27(b), are delocalised over the entire crystal structure (but do not contribute to electrical conductivity as they are in paired levels). As their energy increases with increase in n, and λ decreases, the situation is reached at which $\lambda = l/2$ and electron delocalisation over the entire crystal is no longer possible. At this point, the band gap opens. The range of energies, or k-space, covering $-\pi/l$ to $+\pi/l$, is known as the first *Brillouin zone*, as shown for the 1D case in Fig. 3.27(d). In 3D, interatomic separations, l, vary with direction and, as with Bragg's law, k-space adopts a more complex structure; the shape of the Brillouin zones are linked directly to the 3D crystal structure.

Similar conclusions about the existence of energy bands in solids are obtained from both the MO and periodic potential approaches. From either theory, one obtains a model with bands of levels for the valence electrons. In some materials, overlap of different bands occurs. In others, a forbidden gap exists between energy bands, as shown in Fig. 3.25.

Experimental evidence for the band structure of solids is obtained using a range of spectroscopic techniques in which electronic transitions between different levels are observed. For solids, X-ray emission and absorption give information about both core and valence electrons; information about valence electrons also comes from visible and UV spectroscopy. X-ray emission spectra of solids usually contain peaks or bands of various width. Transitions between inner levels appear as sharp peaks and indicate that these levels are discrete AOs.

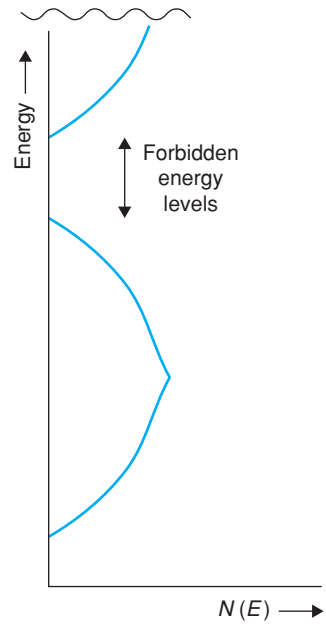


Figure 3.30 ***Density of states showing a band gap in semiconductors and insulators.***

Transitions involving valence shell electrons may give broad spectral peaks, especially for metals, indicating that valence electrons have a broad distribution of energies and are therefore located in bands.

3.4.1 Band structure of metals

In metals, the highest occupied band, the valence band, is only part full, as shown schematically for Na and Mg in Fig. 3.25. In metals such as Mg, energy bands overlap and both 3*s* and 3*p* bands contain electrons. Overlap of bands is responsible for the metallic properties of Mg. If the 3*s* and 3*p* bands did not overlap, then the 3*s* band would be full, the 3*p* band empty and Mg would not be metallic. This is the situation that holds in insulators and semiconductors. In other metals, such as Na, Fig. 3.25, the Fermi level is associated primarily with one partly filled band, in this case the 3*s* band.

3.4.2 Band structure of insulators

The *valence band* in insulators is full. It is separated by a large, forbidden gap from the next energy band, which is empty, Fig. 3.25. Diamond is an excellent insulator with a band gap of ~6 eV. Few electrons from the valence band have sufficient thermal energy to be promoted into the empty band above. Hence the conductivity is negligibly small. The origin of the band gap in diamond is similar to that for Si, discussed next.

3.4.3 Band structure of semiconductors: silicon

Semiconductors have a similar band structure to insulators but the band gap is not very large, usually in the range 0.5–3.0 eV. A few electrons have sufficient thermal energy to be promoted into the empty band.

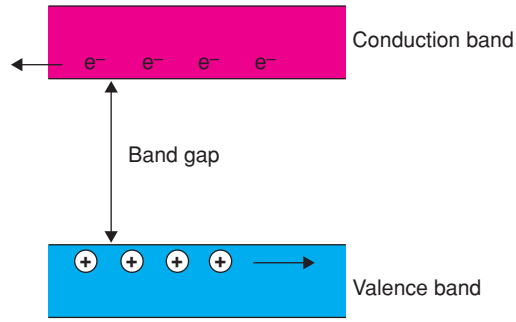


Figure 3.31 ***Positive and negative charge carriers.***

Two types of conduction mechanism may be distinguished in semiconductors, Fig. 3.31. Any electrons that are promoted into an upper, empty band, called the *conduction band*, are regarded as negative charge carriers and would move towards a positive electrode under an applied potential. The vacant electron levels that are left behind in the valence band may be regarded as *positive holes*. Positive holes move when an electron enters them, leaving its own position vacant as a fresh positive hole. Effectively, positive holes move in the opposite direction to electrons.

Intrinsic semiconductors are pure materials whose band structure is as shown in Fig. 3.31. The number of electrons, n, in the conduction band is governed entirely by (i) the magnitude of the band gap, E, and (ii) temperature, T, and is given by the expression:

$$n = n_0 \exp(-E/kT) \tag{3.44}$$

Pure Si is an intrinsic semiconductor. Its band gap, and those of other Group IV elements, are given in Table 3.18.

The band structure of Si is, however, quite different from that which might be expected by comparison with Al. In Al, the 3*s* and 3*p* levels overlap to give two broad bands, both of which are part full. If the trend continued, two similar bands would be expected in Si. The bands would be half full and Si would be metallic. Clearly this is not the case and instead Si contains two bands separated by a forbidden gap. Further, the lower band contains four electrons per Si and is full. If the forbidden gap merely corresponded to separation of *s* and *p* bands, the *s* band would contain only two electrons per Si. This cannot be the explanation, therefore.

A simplified explanation of the band structure of Si (and Ge, diamond, etc.) based on MO theory is as follows. It starts from the observation that each Si forms four equal bonds arranged tetrahedrally; these bonds

Table 3.18 ***Band gaps of Group IV elements***

Element	Band gap/eV	Type of material
Diamond, C	6.0	Insulator
Si	1.1	Semiconductor
Ge	0.7	Semiconductor
Grey Sn (>13 °C)	0.1	Semiconductor
White Sn (<13 °C)	0	Metal
Pb	0	Metal

or orbitals may be regarded as sp^3 hybridised. Each hybrid orbital overlaps a similar orbital on an adjacent Si to form a pair of MOs, one bonding, σ, and the other antibonding, σ^*. Each can contain two electrons, one from each Si. The only step that remains is to allow the individual σ MOs to overlap to form a σ band; this becomes the valence band. The σ^* orbitals overlap similarly and become the conduction band. The σ band is full since it contains four electrons per Si. The σ^* band is empty.

Further discussion of semiconductors and their applications is deferred to Chapter 8.

3.4.4 Band structure of inorganic solids

So far, we have concentrated on materials that are conductors of electricity. However, many inorganic solids can be treated profitably using band theory, whether or not they are electrical conductors. Band theory provides additional insight into the structures, bonding and properties of inorganic solids which complements that obtained with ionic/covalent models. Most inorganic materials are more complex structurally than metals and semiconducting elements. Consequently, often their band structures are known only approximately. Some examples follow.

3.4.4.1 III–V, II–VI and I–VII compounds

Above, we considered Group IV elements, especially Si; closely related are *III–V compounds*, such as GaP. Their valence shells are isoelectronic with Group IV elements; they are also semiconducting. Let us now take this one stage further and consider more extreme cases, with I–VII compounds such as NaCl and II–VI compounds such as MgO.

The bonding in these is predominantly ionic. They are white, insulating solids with negligible electronic conductivity. Addition of dopants tends to produce ionic rather than electronic conductivity. On the assumption that NaCl is 100% ionic, the ions have the following configurations:

$$\mathrm{Na}^+ : (\text{He core})2s^2 2p^6$$
$$\mathrm{Cl}^- : (\text{Ne core})3s^2 3p^6$$

Hence the 3*s*, 3*p* valence shell of Cl^- is full and that of Na^+ is empty. Adjacent Cl^- ions are approximately in contact in NaCl and the 3*p* orbitals may overlap somewhat to form a narrow 3*p* valence band which is full. This band is composed of anion orbitals only. The 3*s*, 3*p* orbitals on Na^+ ions may also overlap to form a band, the conduction band. This band is formed from cation orbitals only. It is empty under normal conditions since the band gap is large, ~7 eV. The band structure of NaCl is that of an insulator, Fig. 3.25, therefore, but with the additional detail that the valence band is composed of anion orbitals and the conduction band of cation orbitals. Any promotion of electrons from the valence band to the conduction band may also be regarded as back transfer of charge from Cl^- to Na^+.

This conclusion leads us to expect a correlation between the magnitude of the band gap and the difference in electronegativity between anion and cation. A large electronegativity difference favours ionic bonding. In such cases, back transfer of charge from anion to cation is difficult and hence ionic solids have large band gaps. The band gaps of a variety of inorganic solids are given in Table 3.19. A quantitative relation between band gap and ionicity proposed by Phillips and van Vechten, equation (3.29), exists. The band gap is made of two components: the 'homopolar band gap', which is the band gap that would be observed in the absence of any difference in electronegativity between the constituent elements, and the part associated with the degree of ionic character in the bonds.

Table 3.19 ***Band gaps (eV) of some inorganic solids***[a]

I–VII compound	Band gap/eV	II–VI compound	Band gap/eV	III–V compound	Band gap/eV
LiF	11	ZnO	3.4	AlP	3.0
LiCl	9.5	ZnS	3.8	AlAs	2.3
NaF	11.5	ZnSe	2.8	AlSb	1.5
NaCl	8.5	ZnTe	2.4	GaP	2.3
NaBr	7.5	CdO	2.3	GaAs	1.4
KF	11	CdS	2.45	GaSb	0.7
KCl	8.5	CdSe	1.8	InP	1.3
KBr	7.5	CdTe	1.45	InAs	0.3
KI	5.8	PbS	0.37	InSb	0.2
		PbSe	0.27	β-SiC	2.2
		PbTe	0.33	α-SiC	3.1

[a]Some of these values, especially for the alkali halides, are only approximate.

3.4.4.2 *Transition metal compounds*

In transition metal compounds, an additional important factor that influences the band structure and the resulting electronic properties is the presence of partly filled metal *d* orbitals. In some cases, these overlap to give a *d* band(s) and the material may have high conductivity. In other cases, *d* orbital overlap is limited and the orbitals are effectively localised on the individual atoms. An example of the latter is stoichiometric NiO. Its pale green colour is due to internal *d*–*d* transitions within the individual Ni^{2+} ions. It has a very low conductivity, $\sim 10^{-14}\ \Omega^{-1}\ cm^{-1}$ at 25 °C, and there is no evidence for significant overlap of the *d* orbitals to form a partly filled *d* band. Examples at the other extreme are TiO and VO. These have the same rock salt structure as NiO, but by contrast, *d* orbitals of the type d_{xy}, d_{xz}, d_{yz} on the M^{2+} ions overlap strongly,

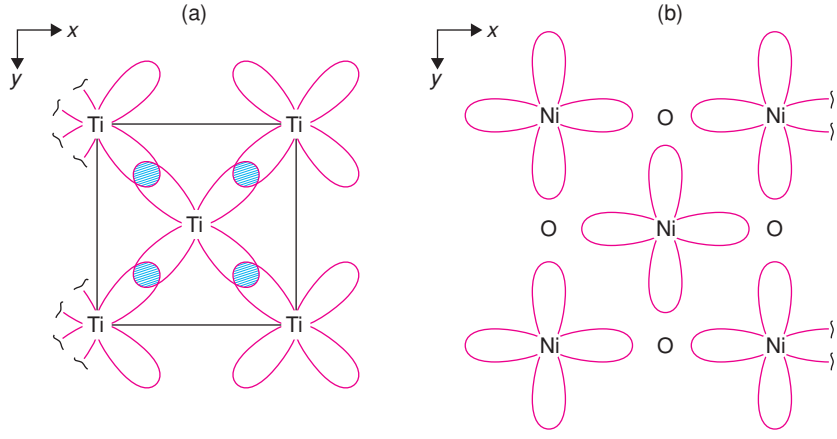


Figure 3.32 ***(a) Section through the TiO structure, parallel to a unit cell face, showing Ti^{2+} positions only. Overlap of d_{xy} orbitals on adjacent Ti^{2+} ions, and similar overlap of d_{xz} and d_{yz} orbitals, leads to a t_{2g} band. (b) Structure of NiO, showing $d_{x^2-y^2}$ orbitals pointing directly at oxide ions and, therefore, unable to overlap and form an e_g band.***

Fig. 3.32(a), to form a broad t_{2g} band. This band is only partly filled by electrons. Consequently, TiO and VO have metallic conductivity, $\sim 10^3\ \Omega^{-1}\ \text{cm}^{-1}$ at 25 °C.

An additional difference between TiO and NiO is that the t_{2g} band, capable of containing six electrons per metal atom, must be full in NiO. The two extra d electrons in Ni^{2+} are in e_g levels, d_{z^2}, $d_{x^2-y^2}$. These e_g orbitals point directly at the oxide ions, Fig. 3.32(b). Because of the intervening oxide ions, the e_g orbitals on adjacent Ni^{2+} ions cannot overlap simply to form a band. Hence e_g orbitals remain localised on the individual Ni^{2+} ions.

Some guidelines as to whether good overlap of d orbitals is likely have been given by Phillips and Williams; d band formation is likely to occur if:

(a) The formal charge on the cations is small.
(b) The cation occurs early in the transition series.
(c) The cation is in the second or third transition series.
(d) The anion is reasonably electropositive.

The reasoning behind these guidelines is fairly straightforward. Effects (a)–(c) keep the d orbitals spread out as far as possible and reduce the amount of positive charge that electrons in these orbitals 'feel' from their parent transition metal ion nucleus. Effect (d) is associated with reduced ionicity and band gap, as discussed earlier in this section. A variety of examples can be found to illustrate these guidelines:

- For (a), TiO is metallic whereas TiO_2 is an insulator. Cu_2O and MoO_2 are semiconductors whereas CuO and MoO_3 are insulators.
- For (b), TiO, VO are metallic whereas NiO and CuO are poor semiconductors.
- For (c), Cr_2O_3 is a poor conductor whereas lower oxides of Mo, W are good conductors.
- For (d), NiO is a poor conductor whereas NiS, NiSe, NiTe are good conductors.

The d electron structure of solid transition metal compounds is also sensitive to the crystal structure and the oxidation state of the transition metal. Some interesting examples are provided by complex oxides with the spinel structure:

(a) Both Fe_3O_4 and Mn_3O_4 are spinels but whereas Mn_3O_4 is an insulator, Fe_3O_4 has almost metallic conductivity. The structure of Fe_3O_4 may be written as

$$[Fe^{3+}]^{tet}[Fe^{2+},\ Fe^{3+}]^{oct}O_4 : \quad \text{inverse spinel}$$

whereas the structure of Mn_3O_4 is

$$[Mn^{2+}]^{tet}[Mn_2^{3+}]^{oct}O_4 : \quad \text{normal spinel}$$

Since Fe_3O_4 is inverse, it contains Fe^{2+} and Fe^{3+} ions distributed over the octahedral sites. These octahedral sites are close together since they belong to edge-sharing octahedra. Consequently, positive holes can migrate easily from Fe^{2+} to Fe^{3+} ions and hence Fe_3O_4 is a good conductor.

In Mn_3O_4, the spinel is normal and the closely spaced octahedral sites contain only Mn^{3+} ions. The tetrahedral sites, containing Mn^{2+} ions, share corners only with the octahedral sites. The Mn^{2+}–Mn^{3+} distance is greater, therefore, and electron exchange cannot take place easily.

(b) A related example, which is really an example of guideline (b) above, is provided by the lithium spinels, $LiMn_2O_4$ and LiV_2O_4. Their structural formulae are similar:

$$[Li^+]^{tet}[Mn^{3+},\ Mn^{4+}]^{oct}O_4$$
$$[Li^+]^{tet}[V^{3+},\ V^{4+}]^{oct}O_4$$

A mixture of +3 and +4 ions is present in the octahedral sites of both but *d* orbital overlap is greater for V than for Mn, which is reflected in the electrical properties: $LiMn_2O_4$ is a hopping semiconductor whereas LiV_2O_4 is metallic (see Chapter 8 for more details).

3.4.4.3 Fullerenes and graphite

Carbon polymorphs (or allotropes) and carbon compounds provide very good examples of how electrical properties can change from insulating, through semiconducting to metallic, depending on the degree of overlap of $p\pi$ orbitals on C and the extent of filling of the resulting MOs or bands. The idea of delocalisation of $p\pi$ electrons stems from the unusual structure and properties of benzene, C_6H_6. Early bonding models were based on alternating single and double C–C bonds with possible resonance between two bonding configurations, Fig. 3.33(a), followed by a model with delocalisation of the π electrons over all six C atoms in the benzene ring, (b). Overlap of *p* orbitals on adjacent C atoms is greatly assisted by the planarity of the benzene molecule, in which the p_z orbital on each C is oriented perpendicular to the plane of the molecule.

Graphite may be regarded as infinite layers of benzene molecules, Fig. 3.33(c), in which the π electrons, one for each C, are delocalised over the complete layer. Band structure calculations show that the overlapping p_z orbitals give rise to two bands of levels, π and π^*, with bonding and antibonding character, as a simple extension of the π, π^* MOs in the ethene (ethylene) molecule, $CH_2{=}CH_2$ (d, e). The π, π^* levels in graphite are not discrete levels, as in C_2H_4, but form bands of levels (f). The lower π band (valence band) is full and the upper π^* band (conduction band) is essentially empty. However, the two bands do overlap by about 0.04 eV in the 3D graphite structure, which allows electrons from the valence band to be readily promoted into the conduction band. Consequently, graphite exhibits electronic conductivity (the limited amount of band overlap makes it a *semimetal*, Fig. 3.25) and is a black, lustrous solid.

An interesting consequence of the band structure of graphite, linked to its layered crystal structure with a large amount of empty space in the van der Waals gap between adjacent layers, is the formation of graphite compounds by intercalation. In the formation of compounds such as C_8Na and C_8Br, Fig. 3.33(g), electrons are either added to ($C_8^-\ Na^+$) or withdrawn from ($C_8^+\ Br^-$) the graphite conduction/valence bands as part of a solid state redox reaction. Hence Na, as a reducing agent, donates electrons (h) and may be regarded as a *donor dopant*. Br, by contrast, is an *acceptor dopant* and creates empty electron levels at the top of the valence band (not shown).

The band structure of fullerene, C_{60}, is similar to that of graphite, but the valence and conduction bands are narrower and are separated by a band gap of 2.6 eV, Fig. 3.34. The delocalised π electron system of C_{60} is different to that of all other C-based molecular materials to date as the C_{60} molecules are not flat (as in benzene, graphite, etc.), but are curved. In graphite, each three-coordinate C is regarded as sp^2 hybridised, with bond angles of 120°, and the p_z orbitals are perpendicular to the plane of the C rings. In C_{60}, the curvature of the molecule means that there is some s–p_z orbital mixing in the π orbital system. Thus, the coordination number of each C is still three, but each C atom is above the plane of its three C neighbours rather than being coplanar with them; instead of three bond angles of 120°, as in a planar system, they are 120, 120 and 108° in C_{60}.

Electronic structure calculations show that the valence band or the HOMOs, symbol h_u, contain 10e in C_{60}. The conduction band, or the LUMOs, symbol t_{1u}, can contain up to 6e. C_{60} is also able to form a wide range of intercalation compounds in which C_{60} acts as an electron acceptor. Thus, in K_3C_{60} (see Section 1.15.6), the conduction band is half full, Fig. 3.34(b), and K_3C_{60} is metallic, in contrast to the insulating C_{60} parent.

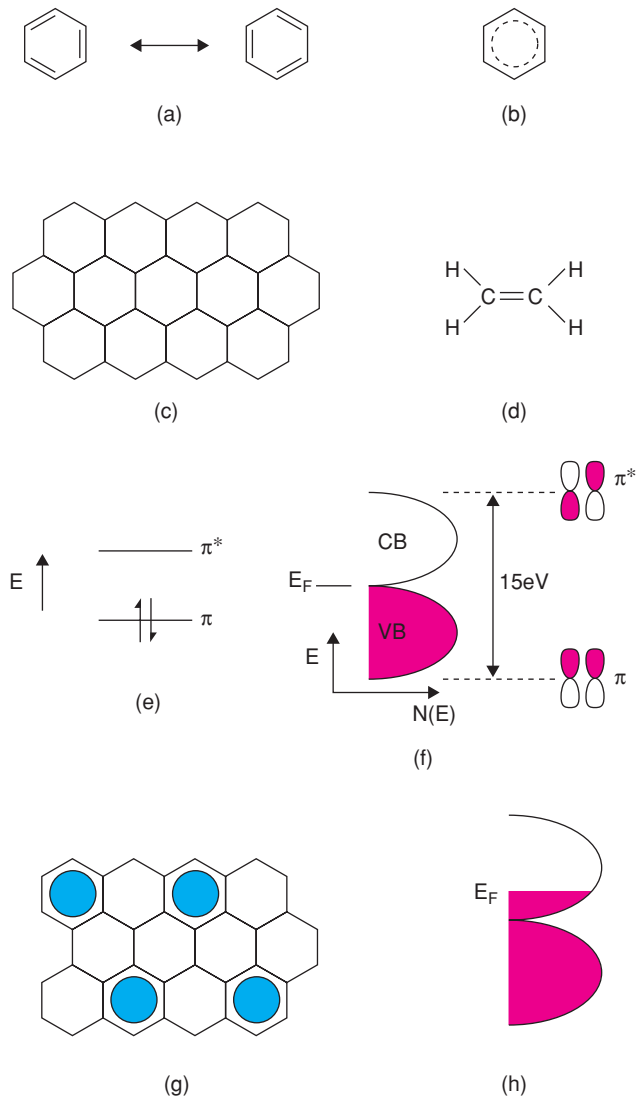


Figure 3.33 ***(a) Resonating bond model proposed initally to explain the structure and properties of benzene, (b) delocalised π electron model for benzene, (c) a layer of graphite, (d) the ethene molecule, (e) MOs in ethene, (f) band structure of graphite, (g) ordered arrangement of M (e.g. K, Br) between adjacent graphite layers in C_8M and (h) band structure in C_8Na.***

3.5 Bands or Bonds: a Final Comment

The three extreme types of bonding in solids are ionic, covalent and metallic. Within each category, many good examples can be found. Most inorganic solids do not belong exclusively to one particular category, and then the question arises as to the most appropriate way of describing their bonding.

The band model is clearly appropriate where there are freely mobile electrons, as in metals and some semiconductors. Experimental measurements of mobility show that these electrons are highly mobile and

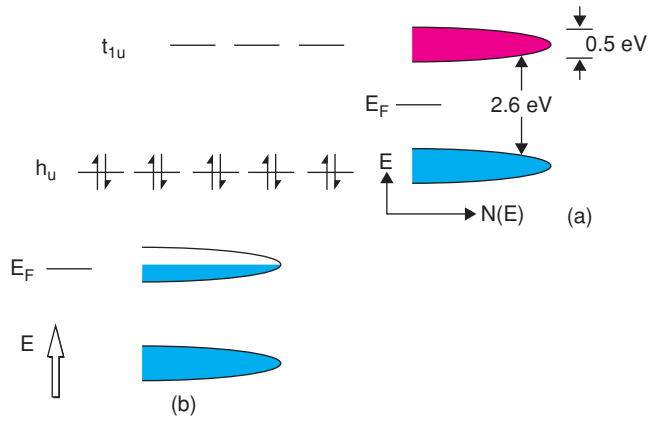


Figure 3.34 ***(a) Electronic structure of C_{60} and (b) band filling in K_3C_{60}.***

are not associated with individual atoms. In certain other semiconductors, such as doped nickel oxide (see Chapter 8), the band model appears not to be suitable since these materials are best regarded as hopping semiconductors in which the electrons do not have high mobility. Instead, it appears to be more appropriate to regard the *d* electrons as occupying discrete orbitals on the nickel ions.

It is important to remember, however, that the question of conduction in nickel oxide refers to only one or two sets of energy levels. NiO, like all materials, has many sets of energy levels. The lower lying levels are full and are discrete levels associated with the individual anions and cations. At higher energy, various excited levels are usually completely empty but may overlap to form energy bands. In asking whether or not a bond or band model is the most suitable, one has to be clear about the particular property or set of energy levels to which the question refers. Thus many ionically bonded solids may, under UV irradiation, show electronic conductivity that is best described in terms of band theory, whereas their crystal structures and lattice energies are best described in terms of ionic bonding.

This illustrates the general point that no single bonding theory is all-encompassing and able to explain everything concerning structures and properties. It is important, therefore, to be familiar with the different bonding theories, to recognise the key points of each and, where appropriate, to be able to mix ideas and terminology when discussing a particular material or property.

4

Synthesis, Processing and Fabrication Methods

4.1 General Observations

Many methods can be used to synthesise non-molecular inorganic solids. Some solids can be prepared by a variety of routes but others, especially those that are not thermodynamically stable, may be much more difficult to prepare and may require special methods. Non-molecular inorganic solids can also be prepared in various forms, as fibres, films, foams, ceramics, powders, nanoparticles and single crystals, as shown for one example, Al_2O_3, in Table 4.1. For these different forms, synthesis is not the problem; instead, it is optimisation of the processing and fabrication procedures that is crucially important. This, linked to the possibilities of doping, modification of properties and possibilities for new applications, makes the study of inorganic solids such a diverse, challenging and stimulating area of science.

In this chapter, the main methods used to synthesise inorganic solids are summarised, together with the methods used to fabricate solids in different forms. A few new materials are introduced, such as amorphous silicon and diamond films, which are important technologically and can be discussed easily within the framework of synthesis methods.

4.2 Solid State Reaction or Shake 'n Bake Methods

The oldest, simplest and still most widely used method to make inorganic solids is to mix together powdered reactants, perhaps press them into pellets or some other shape and then heat in a furnace for prolonged periods. The method is not sophisticated, hence the use of alternative names such as shake 'n bake or, beat 'n heat! It is nevertheless very effective and, for instance, almost all the high-T_c superconductors were first prepared by this method.

Solid state reactions are intrinsically slow because, although the reactants may be well mixed at the level of individual particles (e.g. on a scale of 1 μm or 10^{-3} mm), on the atomic level they are very inhomogeneous. In order to achieve atomic level mixing of reactants, either solid state counterdiffusion of ions between different particles or liquid- or gas-phase transport is necessary to bring together atoms of the different elements, and in the correct ratio, to form the desired product. Consider a typical solid state reaction, that of MgO and

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

Table 4.1 ***Different forms of inorganic solids, e.g. alumina, Al_2O_3***

Solid form	Properties and applications
Ceramic or film	Dielectric substrate
Fibre	Saffil insulation wool
Powder	Grinding and polishing medium
Cr-doped	Ruby gemstone and laser generation
Na-doped	β-Alumina, $NaAl_{11}O_{17}$ solid electrolyte

Al_2O_3 powders to form $MgAl_2O_4$ spinel. Let us consider the various processes involved in the reaction, Fig. 4.1.

4.2.1 Nucleation and growth, epitaxy and topotaxy

At every interface between contacting grains of MgO and Al_2O_3, all the ingredients necessary to form the $MgAl_2O_4$ product are present. The first stage, therefore, is *nucleation* of small crystals with the spinel stoichiometry and structure. For nuclei to be stable and not to disintegrate spontaneously as soon as they form, they must usually be several tens of ångströms across, i.e. they must be considerably larger than a single unit cell and must therefore contain many tens of atoms. The *critical size* of nuclei represents a balance between the negative free energy of formation of the spinel product and the positive surface energy of the nuclei. If the nuclei are too small, the surface area to volume ratio is too large and the nuclei are unstable. Nucleation tends to be difficult simply because of the large number of ions that must get together, in the correct arrangement, to form a stable nucleus.

Nucleation is greatly facilitated if it is *heterogeneous* and the product nuclei are able to form on the surface of an existing structure, especially if there is some similarity in the two structures. The MgO/Al_2O_3 reaction to form spinel provides two examples of possible heterogeneous nucleation, on the surfaces of either MgO or Al_2O_3 grains, Fig. 4.2.

At the $MgO/MgAl_2O_4$ interface, provided that the $MgAl_2O_4$ nucleus and the underlying MgO are in the same crystallographic orientation, the oxide ion arrangement can continue unchanged since it is *ccp* in both structures, with an ABC stacking sequence, Fig. 4.2(a). In MgO, the Mg^{2+} ions occupy fully the octahedral sites (rock salt structure), whereas in spinel, Mg occupies one-eighth of the tetrahedral sites with Al in half of

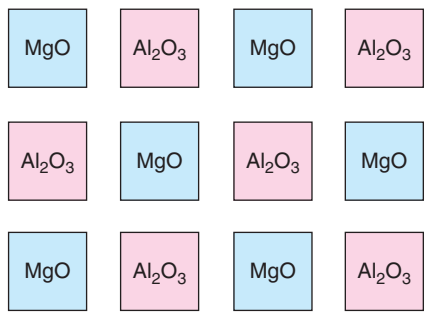


Figure 4.1 ***Idealised reaction mixture composed of grains of MgO and Al_2O_3. In practice, their shapes will be irregular, of different size and not arranged in such an orderly fashion.***

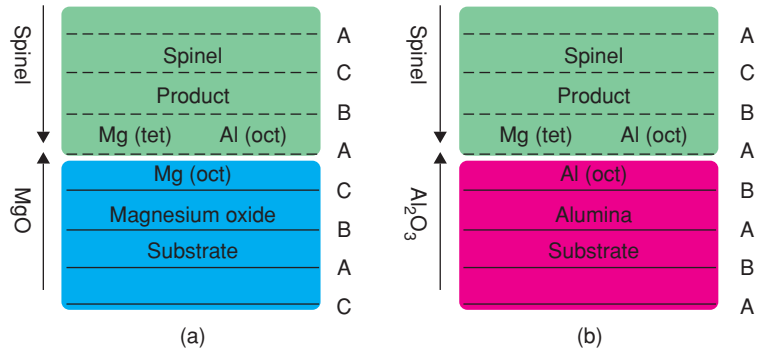


Figure 4.2 ***Nucleation of $MgAl_2O_4$ spinel on (a) MgO and (b) Al_2O_3. Letters A, B, C refer to close-packed layers of oxide ions.***

the octahedral sites. Formation of a spinel nucleus on top of an MgO structure (or *substrate*) is relatively easy because of the structural similarity of MgO and spinel and the absence of any necessity to create a stand-alone spinel nucleus with large surface area.

At a suitably oriented $Al_2O_3/MgAl_2O_4$ interface, the situation is similar, but now the oxygen stacking sequence changes from *hcp* (AB) in Al_2O_3 to *ccp* in spinel (Al_2O_3 has the corundum structure with *hcp* oxide ions and Al in two-thirds of the octahedral sites). Oriented nucleation of spinel on an alumina substrate is still easy (the change in stacking sequence resembles that found in stacking faults, Chapter 2), but the structural similarity between substrate and nucleus is limited to the 2D interface and is referred to as *epitaxy*. In the MgO/spinel case, the structural similarity extends to 3D since there is a common oxygen packing arrangement; this is referred to as *topotaxy*. When discussing chemical reactions in which major structural features of one of the reactants are preserved in the product, it is common to refer to these as *topochemical reactions*.

Epitactic and topotactic features are common in solid state reactions, crystal growth, oriented thin film deposition, intercalation reactions and phase transformations. They are fascinating since the two structures to either side of a common interface may have very different properties and can transform imperceptibly from one to the other. A beautiful analogue is provided by the 'fishes to birds' transformation in one of the Escher drawings, Fig. 4.3. In the crossover region, the structure has recognisable features of both fishes and birds but is not completely one or the other. Similar effects may occur at the interfaces between two solids; even if there are no atomic displacements at the interface, the coordination environments of atoms at the interface are different to either side of the interface. In practice, atomic displacements may also occur, further complicating the structural details within the interfacial region.

Oriented nucleation, of the kind described above, is relatively easy provided that (a) the appropriate atoms that form the product nuclei are nearby and (b) there is good *lattice matching* between substrate and nucleus; for instance, the distance separating O^{2-} ions should be similar in both structures. If lattice matching is good, a *coherent interface* may form; if lattice matching is not good, large strains rapidly develop, the coherency of the interface is lost and the nucleus becomes detached from the substrate. In our MgO/Al_2O_3 reaction mixture, Fig. 4.1, we may expect thin nuclei of spinel to form whenever grains of MgO and Al_2O_3 are close together or touching because all three structures have a close-packed oxide arrangement and lattice matching across the interfaces should be good.

Although the first few atomic layers of product nuclei may form easily, subsequent growth or thickening of the product is more difficult because, effectively, the two reactants, MgO and Al_2O_3, are no longer in contact

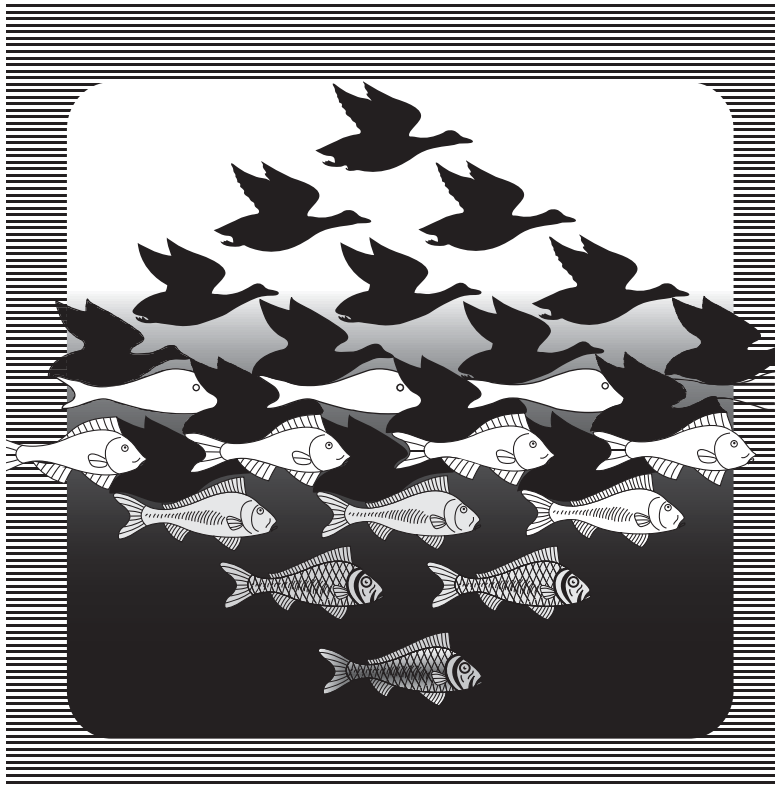


Figure 4.3 ***Fishes to birds: Escher drawing. Reprinted with permission from M.C. Escher's 'Sky and Water I' © 2013 The M.C. Escher Company-The Netherlands. All rights reserved. www.mcescher.com***

but are separated by a rather impenetrable spinel layer. A complex counter-diffusion process is then required in which Mg^{2+} ions diffuse away from, and Al^{3+} ions diffuse towards, the $MgO/MgAl_2O_4$ interface, and vice versa for the $MgAl_2O_4/Al_2O_3$ interface. Assuming for the moment that there is no way that the ions can avoid the long diffusion pathway through the spinel product layer, we have the scenario summarised in Fig. 4.4. As reaction proceeds, the spinel layer thickens, the diffusion pathlength increases and reaction slows.

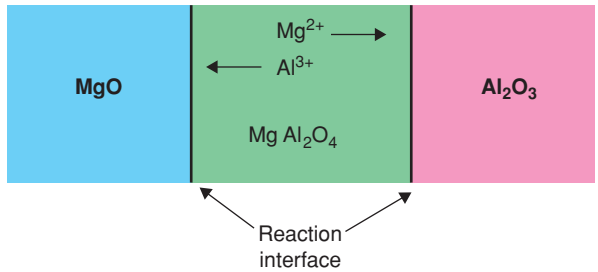


Figure 4.4 ***Spinel product layer separating MgO and Al_2O_3 reactant grains.***

In order to preserve local electroneutrality during reaction and throughout the product, it is necessary that, for every three Mg^{2+} ions that diffuse to the right-hand interface, two Al^{3+} ions must diffuse to the left-hand interface. The reactions that occur at the interfaces may be written, ideally, as follows:

interface $MgO/MgAl_2O_4$:

$$2Al^{3+} - 3Mg^{2+} + 4MgO \rightarrow MgAl_2O_4$$

interface $MgAl_2O_4/Al_2O_3$:

$$3Mg^{2+} - 2Al^{3+} + 4Al_2O_3 \rightarrow 3MgAl_2O_4$$

overall reaction:

$$4MgO + 4Al_2O_3 \rightarrow 4MgAl_2O_4$$

from which we see that the right-hand interface should move three times as quickly as the left-hand interface. In cases where reactant and product are of different colour, it may be possible, in carefully designed experiments, to observe visually the displacement of the interfaces; this is known as the *Kirkendall effect*.

The spinel formation reaction is particularly slow because ions such as Mg^{2+} and Al^{3+} diffuse very slowly; typically, heating for 1 week at 1500 °C would be required to form a fairly pure spinel product. Defects in both reactants and product are required, particularly, vacant sites for adjacent ions to hop into. High temperatures are therefore required so that ions have sufficient thermal energy to, occasionally, hop out of one site into an adjacent vacancy or interstitial site. Consequently, it can be difficult for solid state reactions to proceed to completion once the remaining reactants are well separated from each other. One important way to accelerate reactions is to grind the partially reacted mixtures, so as to break up reactant/product interfaces and bring fresh reactant surfaces into contact. Another way is if gas- or liquid-phase assisted transport of matter can occur, and bring reactants together without the need for long-range solid state diffusion. A small amount of liquid or gaseous *transporting agent* may be very effective in enhancing reaction rates.

4.2.2 Practical considerations and some examples of solid state reactions

The problems associated with spinel synthesis by solid state reaction are particularly difficult since both reagents, MgO and Al_2O_3, are very stable, inert, non-reactive solids. Solid state reactions may be easier if one or more of the starting materials is chemically reactive and/or contains ions that can diffuse easily. Other problems may arise, however, such as potential loss of reactants by evaporation (e.g. alkali metal oxides, Tl_2O, PbO, Bi_2O_3, HgO), or reactivity towards the container (e.g. transition metal-containing materials). With care, and attention to synthesis procedures, these problems can usually be avoided. There are four main issues for consideration in planning a solid state reaction: choice of starting materials, mixing method, container and heat treatment conditions.

Starting materials

Ideal starting materials should be of accurately known stoichiometry, pure and reactive. Problems with stoichiometry and purity can arise if the reagents are sensitive to water and/or CO_2 in the atmosphere or contain transition elements in uncertain or mixed valence states. Reagents may need to be dried, at a temperature found by trial and error, and subsequently kept in a desiccator. For oxide synthesis, it can be useful to use oxy salt reagents, such as carbonates, acetates or nitrates (but not sulfates, which are very

stable thermally), since these decompose during the initial stages of reaction on heating. This decomposition step may decrease greatly the particle size of the reagent, increase its surface area and therefore increase its potential reactivity. In addition, evolution of gases during the decomposition can help to mix the solid reactants.

Mixing of reactants

For solid state reaction, it is essential to bring particles of reactants into contact; this is facilitated by increasing the surface area of the reactants. Samples are therefore mixed and milled. This can be done manually using a mortar and pestle (agate mortars and pestles are useful since they are non-porous, readily cleaned and should not contaminate samples); there are also various labour-saving mechanical mixing techniques such as ball milling, in which the mixture of reactants is placed inside a rotating container together with a number of balls of, for instance, agate. The container is then rotated for a period of time, e.g. 3–24 h, and the effect of the tumbling motion with the agate balls is to reduce the average particle size of the reactants in addition to achieving an intimate mixture. High-energy milling is possible in planetary ball mills, which are rotated at very high speed. While the mixing and milling processes are carried out rapidly and effectively, there is a danger of contamination from the milling media. To facilitate mixing, by whatever method, a liquid such as water or an organic liquid is often added and then needs to be removed by drying at the end of the mixing stage.

Container

If atmospheric sensitivity of the desired product phase is not a problem, the reaction mixture can simply be heated in air in a furnace in a suitable container. Prime consideration for the container is that it should not react with the sample. Frequently-used inert containers for oxides are boats or crucibles of

- Pt (but Pt reacts with Li_2O, BaO and many transition metal oxides).
- Au (but its melting temperature, 1063 °C, places an upper limit on the reaction temperatures that may be used; however, Au is generally more inert than Pt).
- Al_2O_3 (high-purity alumina is inert and high melting, but nevertheless contamination from the reactants may occur).
- SiO_2 glass (crucibles of pure SiO_2 glass can withstand temperatures up to 1200 °C before the glass softens and starts to devitrify; oxides of alkali metals, in particular, are reactive towards SiO_2).
- Graphite crucibles are commonly used as containers for the synthesis of sulfides, other chalcogenides and nitrides.

Firing schedule

The heating schedule should be designed to (a) cause smooth decomposition of any oxy salt reagents without excessive frothing, melting or leakage of reagent from the container, (b) avoid melting and in particular volatilisation of one or more of the reagents and (c) apply temperatures at which the reagents react together on a reasonable timescale (e.g. 12–24 h). Reactions may be carried out in air or, if a tube furnace is available, a range of controlled atmospheres may be used. Alternatively, sealed ampoules of silica glass or precious metals may be used to prevent loss of volatile reagents or atmospheric attack. For known materials, there may be recipes in the literature detailing appropriate reaction conditions, but for the attempted synthesis of new materials, a trial and error approach is usually needed.

Many solids can be prepared perfectly satisfactorily by solid state reaction; below, three illustrative examples are presented, each of which requires some special care.

4.2.2.1 Li_4SiO_4

Li_4SiO_4 is the parent phase for a family of Li^+ ion conductors that can be prepared by the reaction:

$$2Li_2CO_3 + SiO_2 \xrightarrow[24\ h]{\sim 800\,^\circ C} Li_4SiO_4 + 2CO_2$$

Problem: Li_2CO_3 melts and decomposes readily above ~720 °C; it is reactive towards most container materials, including Pt and silica glass.

Solution: Use Au containers; carry out decomposition and pre-reaction of Li_2CO_3 at ~650 °C for a few hours before final firing at 800–900 °C overnight.

4.2.2.2 $YBa_2Cu_3O_{7-\delta}$

$YBa_2Cu_3O_7$, YBCO, is the classic 90 K superconductor. It can be prepared by the reaction

$$Y_2O_3 + 4BaCO_3 + 6CuO + \frac{1}{2}O_2 \xrightarrow{950\,^\circ C, 350\,^\circ C} 2YBa_2Cu_3O_7 + 4CO_2$$

Problems: (i) $BaCO_3$ is particularly stable on heating and it can be difficult to remove the last traces of CO_2 during reaction. Also, many materials, such as YBCO, may react slowly with atmospheric CO_2 causing a partial reversal of the reaction used in the synthesis. (ii) CuO is reactive to most container materials at high temperatures. (iii) The oxygen content $7 - \delta$ of the YBCO product is variable and must be controlled to optimise T_c.

Solutions: (i) React in a CO_2-free atmosphere, with $Ba(NO_3)_2$ as a source of BaO. (ii) Make pellets of reaction mixture [after decomposition of $Ba(NO_3)_2$] and react these on a bed of pre-prepared YBCO. (iii) After reaction at ~950 °C, carry out final heating at ~350 °C to allow O_2 uptake to occur and achieve the desired stoichiometry, $YBa_2Cu_3O_7$.

4.2.2.3 Na β/β'' alumina

Na β/β''-alumina is the classic Na^+ ion-conducting solid electrolyte. In reality, there are two Na aluminate phases whose formulae are written ideally as $NaAl_{11}O_{17}$ (β) and $NaAl_5O_8$ (β''), although both phases form solid solutions with variable Na:Al ratios. They can be prepared as follows:

$$Na_2CO_3 + x\,Al_2O_3 \xrightarrow{\sim 1500\,^\circ C} \text{Na } \beta/\beta''\text{-aluminas (variable Na content)}$$

Problem: Na_2O is volatile at the required reaction temperatures; lower temperatures cannot be used because of the inertness and unreactivity of Al_2O_3.

Solution: (i) Carry out pre-reaction at 700–800 °C and expel CO_2; fabricate 'green' pellets or tubes of the partially-reacted mixture; cover the pellets or tubes with pre-reacted β/β''-alumina and fire at 1400–1500 °C. This 'buffering' prevents significant loss of Na_2O from the samples at high temperatures. (ii) Use high surface area, reactive alumina starting materials such as boehmite or γ-Al_2O_3. These have a defect spinel structure (similar to that in the β/β''-alumina product) and the first stage of reaction involves intercalation of Na^+ into

the particles of alumina. This greatly assists the nucleation and growth of the product since the reaction is partly topochemical.

The above examples are all oxides. For the synthesis of non-oxides by solid state reactions, similar criteria apply but the methodologies are often completely different. A versatile method for the synthesis of carbides, borides and nitrides is *carbothermal reduction,* which was developed, as the *Acheson process*, to fabricate silicon carbide (an important ceramic used as an abrasive, *carborundum*, and in furnace heating elements) from mixtures of SiO_2 and carbon using the following overall reaction:

$$SiO_2 + 3C \rightarrow SiC + 2CO$$

This is a high-temperature reaction in which carbon is both a reagent and reducing agent for SiO_2. The reaction mechanism is complex and involves silicon monoxide, SiO, as a gaseous intermediate which is very effective in bringing the Si and C reactants into contact. Various other carbides can be made by carbothermal reduction, including Al_4C_3, B_4C, TiC, WC and Mo_2C.

Borides can be prepared by a similar process but the boron reagent is B_2O_3, which needs to be reduced to elemental boron. Carbon is therefore added to the reaction mixture to achieve reduction of both B_2O_3 and the metal oxide, although it should not feature in the boride product. Examples of borides synthesised in this way include AlB_{12}, VB, VB_2 and TiB_2. Nitrides such as BN, AlN, Si_3N_4 and TiN can be prepared similarly using as reagents the metal oxide, N_2 gas and carbon.

4.2.3 Combustion synthesis

In complete contrast to solid state reactions that often require heating for several days at high temperatures to complete a synthesis, controlled explosions in reactions known variously as *combustion synthesis*, *self-propagating high temperature synthesis*, SHS, and *solid state metathesis*, SSM, can give reaction products in minutes or even seconds.

In the SHS technique pioneered by Merzhanov and co-workers in Russia, starting materials are chosen that react together exothermically. Consequently, once a reaction is initiated, enough heat is generated for high temperatures to be attained and complete reaction occurs rapidly. The starting materials may be classified as either 'fuels' or 'oxidants'. An example of an SHS reaction is the *'thermite' process* for the extraction of iron:

$$Fe_2O_3 + 2Al \rightarrow 2Fe + Al_2O_3$$

in which the fuel, Al, and the oxidant, Fe_2O_3, combine to generate temperatures of 3000 °C for a brief period during reaction. In other combustion syntheses, fuels and oxidants are additional reagents that are consumed in the reaction and do not feature in the final product; typical fuels are hydrazine, glycine and urea; common oxidants are nitrates of one of the reacting species.

Many complex oxides can be prepared by combustion synthesis, as shown by the synthesis of nickel zinc ferrite spinel, $Ni_{0.5}Zn_{0.5}Fe_2O_4$. Mixtures of the metal nitrates and urea, in stoichiometric amounts to maximise release of energy, are heated in a vitreous silica crucible to 600 °C, at which temperature self-ignition occurs to give the spinel product in a very short time. The resulting material is a fine powder since gas evolution during reaction prevents agglomeration. The powders may then be pressed into pellets and heated at high temperature to give dense ceramics. Temperature control is important since heating below 1100 °C leads to incomplete densification, whereas above ~1300 °C exaggerated grain growth, EGG, occurs to give a ceramic with an inhomogeneous texture.

In closely-related SSM reactions, the reactants are chosen such that one of the products (usually the side-product) has a very high lattice energy and its formation provides the driving force for the reaction. An example is

$$5Li_3N + 3TaCl_5 \rightarrow 3TaN + 15\,LiCl + N_2$$

in which covalent $TaCl_5$ and highly reactive Li_3N form the desired product, TaN, together with LiCl; LiCl may be dissolved readily to leave pure TaN. SSM methods are usually better controlled than SHS methods since boiling/volatilisation of the side-product (LiCl in the above case) reduces the driving force for the reaction (by the loss of free energy of formation of the crystalline side-product) and therefore limits the temperature attained during synthesis to that of the boiling point of the side-product.

A wide range of materials may be prepared by SHS and SSM methods. Various carbon-related materials including carbon nanotubes can be prepared by SSM of lithium acetylide and carbon halides in the presence of a catalyst:

$$3Li_2C_2 + C_2Cl_6 \rightarrow 8C + 6LiCl$$

The carbon nanotubes synthesised in this way consist of single-walled and multi-walled bamboo-type nanotubes and are generated in a few seconds at temperatures that exceed 2000 °C.

Many carbides, silicides, nitrides, phosphides and sulfides, and also many oxides, can be prepared by SSM with reactions such as:

$$3ZrCl_4 + Al_4C_3 \rightarrow 3ZrC + 4AlCl_3$$

$$V_2O_5 + 2Mg_2Si + CaSi_2 \rightarrow 2VSi_2 + 4MgO + CaO$$

$$2AgF + Na_2S \rightarrow Ag_2S + 2NaF$$

$$HfCl_4 + 2Na_2O \rightarrow HfO_2 + 4NaCl$$

In each case, it is necessary to carry out preliminary tests in order to optimise the conditions so as to obtain phase-pure products.

4.2.4 Mechanosynthesis

Solid state reactions are generally high-temperature methods, but in favourable cases reactions can be carried out simply by grinding together the reagents. High-energy milling is necessary, often for periods of several hours, and there can be a danger of contamination from the milling media. Nevertheless, the method is finding increasing applications as a synthetic method. Details of just how solids react are still rather unclear; the reagents are reduced to nanometre size with the introduction of mechanically induced crystal defects. Although no heat is applied, high temperatures may be generated locally as the mechanical energy is transformed into heat and this acts to speed up interdiffusion rates and formation of the product phase on a very fine scale.

The products of mechanosynthesis may be different from those of the same materials prepared by solid state reaction. They may be in a metastable state, with compositional inhomogeneity, structural disorder and high vacancy concentration. In a recent example of Zn_2SnO_4 nano-sized grains synthesised by mechanochemistry, and characterised by a wide range of diffraction, microscopic and spectroscopic techniques, grain cores were found to have the inverse spinel structure, as expected. However, grain outer shell regions and surfaces showed Zn/Sn disorder and deformed polyhedra, indicating a highly non-equilibrium state. Given the structural

differences between materials made by mechanosynthesis and conventional reaction, it is highly likely that properties may also differ. Mechanosynthesis may become an important method for the synthesis of new nanomaterials and nanocomposite materials, which are relatively cheap and energy-efficient to make as well as for the synthesis of stable, bulk materials that would otherwise require high temperatures for long time periods.

4.3 Low Temperature or Chimie Douce Methods

The greatest drawback of solid state reactions, including mechanosynthesis, is that reactants are not mixed on an atomic scale. There are various methods to achieve atomic scale mixing of reactants, in gas, liquid or even solid phases. Most are low-temperature methods, often referred to as *chimie douce* or *soft chemistry* methods, although final firing at high temperature may be needed, especially for ceramic products. Consequently, greater purity may be attainable, especially if prolonged heating at high temperature can be avoided, since this is a frequent source of contamination from container materials and furnace atmosphere. More importantly, products with high chemical homogeneity are usually obtained, which often leads to improved properties or better understanding of the dependence of properties on structure, composition and dopants. It is also possible to use non-powder feedstocks to produce materials in a variety of forms, such as fibres or coatings.

Disadvantages of these alternative methods are that often the reagents are costly and difficult to handle on a large scale, considerable research may be required to optimise the synthesis of a particular material and, once suitable conditions have been found, may not be readily applicable to the synthesis of related materials. For these reasons, solid state reactions, which are relatively quick, easy and versatile, are usually tried first. Next, some chimie douce methods and their applications are described.

4.3.1 Alkoxide sol–gel method

The first stage is to prepare a homogeneous solution containing all the cationic ingredients in the desired ratio. The solution is gradually dried and, depending on the species present, should transform to a viscous *sol* containing particles of colloidal dimensions and finally to a transparent, homogeneous, amorphous solid known as a *gel*, without precipitation of any crystalline phases. The gel is then fired at high temperatures to remove volatile components trapped in the pores of the gel or chemically bonded hydroxyl and organic side-groups and to crystallise the final product.

There are two main types of sol–gel methods, materials and applications. First, organometallic precursors, particularly alkoxides, are widely used for the small-scale synthesis of known or new materials, often containing several different cations. The reagents for alkoxide-based sol–gel syntheses are metal–organic compounds, such as tetraethyl orthosilicate (TEOS), $Si(OCH_2CH_3)_4$, as a source of SiO_2, titanium isopropoxide, $Ti(O^iPr)_4$, as a source of TiO_2, and aluminium butoxide, $Al(OBu)_3$, as a source of Al_2O_3. These covalent liquids are mixed in the appropriate ratios, often with an alcohol to promote miscibility of the alkoxide and H_2O. Water is a key reagent since it hydrolyses the alkoxides, usually in the presence of either acid or base as a catalyst to speed up reaction.

Hydrolysis occurs in two steps:

(i) replacement of –OR groups by –OH, e.g.

$$Si(OCH_2CH_3)_4 \rightarrow Si(OCH_2CH_3)_3OH + Si(OCH_2CH_3)_2(OH)_2 + \text{etc.}$$

(ii) condensation polymerisation with elimination of H_2O, e.g.

$$(RO)_3Si\text{–}OH + HO\text{–}Si(OR)_3 \rightarrow (RO)_3Si\text{–}O\text{–}Si(OR)_3$$

The composition, structure and viscosity of the reaction products depend very much on the degree of hydrolysis/condensation; careful control of the reaction variables is required to achieve the desired product. For the synthesis of complex oxides containing more than one cation, M, M′, cross-condensation is required:

$$\{\text{–M–OH} + \text{HO–M}'\text{–}\} \rightarrow \{\text{–M–O–M}'\text{–}\}$$

and it is essential to achieve this instead of condensation of the components separately.

Hydrolysis may be base catalysed, with nucleophilic substitution of OH^-:

$$HO^- + Si(OR)_4 \rightarrow (HO)Si(OR)_3 + RO^-$$

or acid catalysed, with electrophilic attack by H^+ (or H_3O^+):

$$Cl^- + H^+ + ROSi(OR)_3 \rightarrow HOSi(OR)_3 + RCl$$

The final stage of synthesis is to heat, or *calcine*, the gel, to decompose organic matter and leave an oxide product.

The alkoxide-based sol–gel method is extremely versatile and can incorporate most elements of the Periodic Table. A selection of examples follows.

4.3.1.1 Synthesis of $MgAl_2O_4$

Appropriate reagents are $Mg(OCH_3)_2$ and $Al(O^sBu)_3$. On mixing, hydrolysis, condensation and drying, an amorphous gel is obtained that decomposes to form very fine particles of spinel after final heating at 250 °C. This compares very favourably, in energy terms, with solid state reaction, which typically requires heating for several days at 1500 °C, but the alkoxide reagents are expensive, hygroscopic and difficult to use.

4.3.1.2 Synthesis of silica glass

Pure silica glass is difficult, and expensive, to prepare by traditional melting since molten SiO_2 is extremely viscous, even at 2000 °C (the melting point of SiO_2 is ∼1700 °C). Starting from TEOS, it is possible by the sol–gel route to prepare an amorphous product which, to all intents and purposes, resembles silica glass but has never been heated above 1200 °C. Both silica glass and this amorphous silica look-alike are metastable and care is required to avoid some crystallisation at the final processing temperature of 1200 °C.

4.3.1.3 Spinning of alumina fibres

We normally encounter Al_2O_3 as a powder or a ceramic. By the sol–gel route, it can be prepared as fibres which, under the tradename Saffil, are used as a thermal insulation to replace asbestos, Table 4.1. The starting material is $Al(OBu)_3$. To produce fibres, it is crucial to control the viscosity at the stage of drawing fibres and the polymer intermediates should be linear rather than branched chain. Acid catalysis is used since base catalysis tends to give 3D polymers. With base catalysis, there is an increase in the rate of substitution of OR by OH groups; with each additional OH group, which is electron withdrawing, the Si carries an increasing $\delta+$ charge and is more susceptible to further nucleophilic attack by OH^- groups. This leads to a high concentration of Si atoms containing several OH groups and, therefore, a high concentration of silanol (–SiOH) groups that lead to cross-linked structures on condensation polymerisation.

4.3.1.4 Preparation of indium tin oxide (ITO) and other coatings

Semiconducting ITO is widely used as an electrode, especially as a thin film coating. It is used as a coating on glass for energy control of buildings since it transmits visible light and reflects infrared radiation. The coatings are prepared by the sol–gel route; starting with In and Sn alkoxides, a thin film is poured onto the substrate or, alternatively, the substrate is dipped into the liquid precursor mixture. The coating is then processed *in situ* giving finally an amorphous, thin, transparent, adherent layer. Other coatings prepared in this way include:

- Ta_2O_5 [from $Ta(OCH_2CH_3)_5$] on Si: Ta_2O_5 is an insulator and is used in Ta_2O_5/Si capacitors;
- SiO_2/TiO_2 on Si wafers: the refractive index depends on the TiO_2 content and SiO_2/TiO_2 anti-reflective coatings are used to increase the efficiency of Si solar cells.

A major concern in preparing coatings is to control the drying conditions such that the coatings remain crack-free and adherent. Generally, a large reduction in volume occurs on drying/decomposition of a gel, because of its high content of water and other volatiles. Thin coatings are better able to sustain these losses and retain their integrity than are thick coatings or bulk samples which can experience large tensile stresses as a result of differential drying and shrinkage rates in the bulk and at the surface.

4.3.1.5 Fabrication of YSZ ceramics

YSZ (yttria-stabilised zirconia) is an important oxide ion conducting ceramic used in gas sensors and solid oxide fuel cells (see Chapter 8). It can be prepared by the sol–gel route starting from $Y(OCH_2CH_2CH_3)_3$ and $Zr(OCH_2CH_2CH_3)_4$.

4.3.2 Sol–gel method using oxyhydroxides and colloid chemistry

Sols and gels can be made from inorganic oxides and hydroxides by adjusting the pH to form polymeric, multi-cation oxyhydroxyl species at intermediate pH values. For example, the aluminium species in aqueous solution range from simple hydrated cations such as $Al(H_2O)_6^{3+}$ in acidic solution to anionic species such as $Al(OH)_4^-$ in basic solution. The reason why Al is regarded as an amphoteric element is that Al_2O_3 is able to dissolve in both acids and bases but is insoluble at intermediate pH values. Figure 4.5 shows some of the species that occur in the $Al^{3+} - H_2O - H^+/OH^-$ system. Starting with hydrated Al^{3+} ions in acidic solution and gradually adding base to increase the pH, a series of hydrolysis reactions occur. At the same time, large polymeric multi-cation species such as $[Al_{13}O_4(OH)_{24}(H_2O)_{12}]^{7+}$ form, leading to a colloidal system, or a sol, with nanoparticulate species in the nanometre size range. These then coalesce to form a gelatinous solid of uncertain formula and structure that is often referred to as a hydrous oxide, but more correctly is probably a hydrated hydroxide. With further increase in pH, this gelatinous precipitate redissolves to give complex anionic species.

Numerous other elements of the Periodic Table with intermediate electronegativity also show complex aqueous solution chemistry. Amphoteric elements such as Zn show the full range of cationic, colloidal and anionic behaviours similar to Al. Slightly more electronegative elements show a wide range of polymeric multi-element anionic species, such as the vast range of complex silicate anions observed with Si. Other examples which form a wide range of discrete polymeric anions are borates and phosphates. Identification of these polymeric species in aqueous solution can be difficult, but in the case of silicates and phosphates they can be separated and characterised using chromatography.

Oxides such as Al_2O_3, SiO_2, Fe_2O_3, ZrO_2 and Y_2O_3 are ideally suited as reagents in sol–gel chemistry because of their wide-ranging solution equilibria and pH-dependent solubilities. They can be used for the

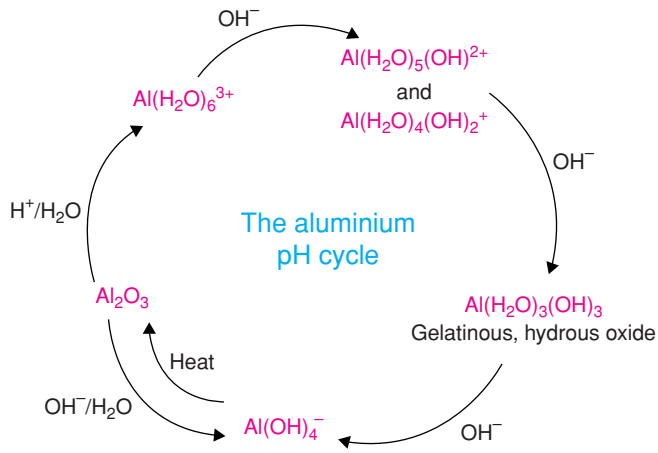


Figure 4.5 ***Solution chemistry of Al showing amphoteric behaviour and effect of pH.***

synthesis of complex oxides by mixing sols of different materials prior to gelation or by adding anionic solutions or even oxide powders to one of the sol liquids. However, this method is particularly suited to preparing films, fibres and ceramic shapes. Micron-thick coatings can be prepared either by dip-coating or spinning a sol onto a substrate such as for the production of silica coatings on stainless-steel wire mesh to increase its oxidation resistance. Some other examples of materials prepared by this route are outlined below

4.3.2.1 Synthesis of zeolites

Zeolites are aluminosilicate framework structures containing large, interconnected voids and channels that permit either ready ion-exchange processes or, in empty-channel structures, small organic molecules to enter where they can be induced to undergo various transformations or reactions. Recently, related framework materials such as ALPOs (aluminophosphates) have been synthesised, and methods to increase the size of the cavities, in the so-called MCM phases, have been developed. Zeolites, or more generally *microporous solids*, are synthesised by sol–gel methods, often with final treatment under hydrothermal conditions (heat in high-pressure water). A typical reaction (idealised) is

$$\begin{aligned} &\mathrm{NaAl(OH)_4(aq) + Na_2Si_3O_7(aq) + NaOH(aq)} \\ &\qquad\qquad \downarrow 25\,^{\circ}\mathrm{C} \\ &\qquad\quad \mathrm{Gel} \\ &\qquad\qquad \downarrow 25\text{–}175\,^{\circ}\mathrm{C}\text{, hydrothermal conditions} \\ &\mathrm{Na}_w\mathrm{Al}_x\mathrm{Si}_y\mathrm{O}_z \text{ zeolite} \end{aligned}$$

The particular zeolite that forms depends on the starting composition, temperature, pressure and also on the *template* that is added. The template is usually an alkylammonium cation, $(R_4N)^+$, around which the aluminosilicate structure crystallises; this controls the size and shape of the aluminosilicate building blocks that coalesce to form the zeolite structure.

After crystallisation of the zeolite, it is necessary to remove the template ions by heating to 300–400 °C, where the organic material burns off, leaving the inorganic framework intact. Although these open aluminosilicate frameworks are not usually thermodynamically stable, they are kinetically stable up to 300–400 °C and are used in a range of applications as 'molecular sieves' and catalysts.

4.3.2.2 Preparation of alumina-based abrasives and films

Fine powders of doped Al_2O_3 are used as abrasive grit, but to control the grade or performance, accurate control of particle size is necessary. A sol–gel fabrication method is used that starts from AlOOH, which is *peptised* or dispersed in dilute acid to give a sol which is then mixed with other cation-containing solutions of, e.g., zirconyl acetate or $Mg(NO_3)_2$. The resulting mixtures polymerise to give a gel which is dried, crushed to give fine particles with sharp edges caused by their fracture and then fired to convert the particles into their oxides.

Sol–gel methods are used to fabricate Al_2O_3 films and coatings. Advantages of this processing method are that the required temperatures are low and the product is homogeneous and cheap and can be used to coat large-area substrates. In addition, high-vacuum kit, such as is needed for other film coating techniques, is not needed. The resulting films are hard, transparent, insulating and resistant to chemicals, abrasion, and corrosion. In a typical method, Al *sec*-butoxide dissolved in isopropanol is mixed with acetylacetone complexing/chelating agent in the ratio 20:0.5. The mixture is stirred at room temperature for 6 h, nitric acid is added dropwise, then water for stabilisation and hydrolysis. The solution is refluxed at 80 °C for 2 h and is then ready for coating on to a substrate, followed by firing to give an adherent Al_2O_3 film. The films can be crystallised either by heating in N_2 at temperatures in the range 200–800 °C for a few hours or in a microwave oven for a few minutes.

4.3.3 Citrate gel and Pechini processes

A third family of sol–gel processes, often with a combustion step to decompose the gel product, utilises carboxylic acids such as citric acid and polyhydroxy alcohols such as ethylene glycol, Fig. 4.6. Citric acid is a chelating agent in which the carboxylate groups act as ligands to form complexes with metal cations. Citric acid is a very useful complexing agent, with three carboxylic ligands on each acid molecule.

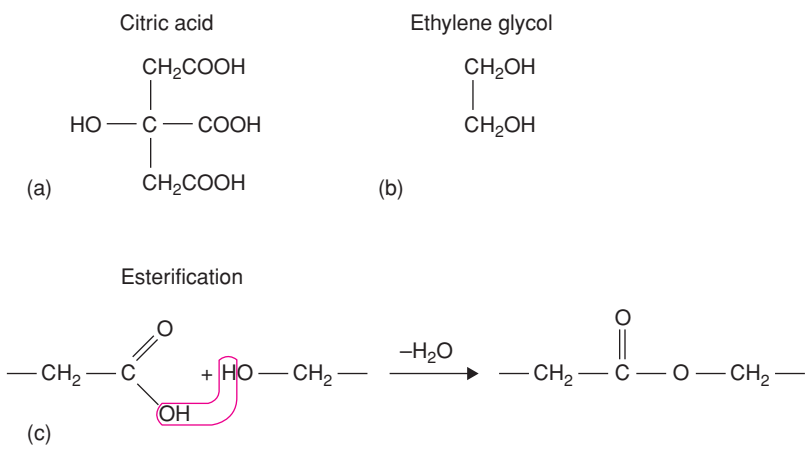


Figure 4.6 ***Reagents and esterification mechanism in the Pechini process.***

The carboxylic acid groups can also take part in esterification reactions, Fig. 4.6, by condensation with hydroxyl groups on ethylene glycol. There is much scope for varying the detailed chemistry of this reaction; other organic acids such as tartaric acid and poly(acrylic acid) can be used; also, other polyhydroxy alcohols ranging from small molecules such as glycerol to larger polymers such as poly(vinyl alcohol) and poly(ethylene glycol) can be used.

To facilitate decomposition of the polymeric gel products, combustion synthesis aids such as urea and nitric acid may be added to the reaction mixture. The combustion synthesis step can give highly porous foam products which may be beneficial either for subsequent ceramic processing associated with the high surface area of crushed foam powders or for applications requiring high surface area materials. The nitric acid also acts to control the pH during reaction and in particular to avoid precipitation of insoluble citrates, which can lead to loss of homogeneity during the intermediate stages of reaction or in the final product.

The Pechini process is chemically complex and involves both acids and alcohols together with esterification, whereas in the citrate gel method, the alcohols are not utilised and, instead, citric acid is simply added to metal nitrate aqueous solutions. The citrate gel method has lower organic content in the gel product with less weight loss and shrinkage during firing and is often preferred for film fabrication.

Many materials have been made by Pechini and citrate gel processes with much scope to modify the nature of the final product by attention to the composition of the reaction mixtures. Examples include synthesis of $YBa_2Cu_3O_7$ superconductor, $Sr(Co,Fe)O_{3-\delta}$ for use as solid oxide fuel cell cathode and $LiMn_2O_4$ lithium battery cathode.

4.3.4 Use of homogeneous, single-source precursors

A solid state analogue of the sol–gel method is to prepare a crystalline, homogeneous, single-phase precursor material that contains all the required cations in the correct ratio. This precursor should not be particularly stable and on heating decomposes to give the desired product. This method suffers from the limitation that it is necessary to find a single-source precursor that has the same cation ratio as the desired product phase. Some examples of crystalline precursors which have the desired properties of limited thermal stability and accurately controlled stoichiometry with cations present in the desired ratio are as follows:

(a) $NH_4Fe(CrO_4)_2$ precursor to $FeCr_2O_4$:
The precursor is synthesised by precipitation from aqueous solution:

$$Fe^{3+}(aq) + 2{CrO_4}^{2-}(aq) + {NH_4}^{+}(aq) \rightarrow NH_4Fe(CrO_4)_2$$

It decomposes on heating:

$$2NH_4Fe(CrO_4)_2 \xrightarrow{1150\,^{\circ}C} 2FeCr_2O_4 + 2NH_3 + H_2O + 7/2O_2$$

(b) $MnCr_2O_7{\cdot}5C_5H_5N$ precursor to $MnCr_2O_4$:
The pyridinate precursor is precipitated from aqueous solution and decomposed in H_2:

$$Mn^{2+}(aq) + {Cr_2O_7}^{2-}(aq) + 5C_5H_5N \rightarrow MnCr_2O_7 \cdot 5C_5H_5$$

$$MnCr_2O_7{\cdot}5C_5H_5N \xrightarrow[H_2]{1100\,^{\circ}C} MnCr_2O_4$$

(c) $Ba(TiO)[(COO)_2]_2$ precursor to $BaTiO_3$:
Barium titanyl oxalate is prepared by a multistep route and decomposed by heating:

$$Ti(OCH_2CH_2CH_2CH_3)_4(aq) + H_2O \rightarrow Ti(OH)_4(s)$$
$$Ti(OH)_4(s) + (COO)_2{}^{2-}(aq) \rightarrow TiO(COO)_2(aq)$$
$$TiO(COO)_2(aq) + Ba^{2+}(aq) + (COO)_2{}^{2-}(aq) \rightarrow BaTiO[(COO)_2]_2$$
$$BaTiO[(COO)_2]_2 \xrightarrow{650\,^\circ C} BaTiO_3$$

(d) BaTi catecholate precursor to $BaTiO_3$:
Again, a multistep route is used:

$$C_6H_4(OH)_2 \text{ (catechol, cat)} + TiCl_4 \rightarrow H_2Ti(cat)_3$$
$$H_2Ti(cat)_3 + BaCO_3 \rightarrow BaTi(cat)_3$$
$$BaTi(cat)_3 \xrightarrow{650\,^\circ C} BaTiO_3$$

4.3.5 Hydrothermal and solvothermal synthesis

Hydrothermal synthesis involves heating reactants in water/steam at high pressures and temperatures. The water has two functions, as a pressure-transmitting medium and as a solvent, in which the solubility of the reactants is *P*,*T*-dependent. The method is simple; reactants and water are placed inside a PTFE-lined cylinder or 'bomb' which is either sealed or connected to an external pressure control. The bomb is placed in an oven, usually at a temperature in the range 100–500 °C. Pressure is controlled either externally or by the degree of filling in a sealed bomb, by making use of the *P*–*T* 'phase diagram', Fig. 4.7(a); curve AB is the *saturated steam curve* and separates water (above) from steam (below); at temperatures above 374 °C, point B, water is in the *supercritical* condition and there is no distinction between liquid and vapour states.

Hydrothermal methods have a wide range of applications:

1. For zeolite synthesis, Section 4.3.2.1; the final stage is to heat aluminosilicate gels hydrothermally to crystallise the zeolites.
2. Single crystals of quartz, SiO_2, which are used as piezoelectrics in a range of applications, are grown hydrothermally in a temperature gradient using the arrangement shown in Fig. 4.7(b). The solubility of SiO_2 in NaOH solution increases with increase in temperature; effectively, therefore, SiO_2 dissolves at the hot end, is transported to the cooler end by convection and crystallises on the suspended seeds. In this reaction, NaOH acts as a *mineraliser*. SiO_2 is only sparingly soluble in water, but solubility increases greatly when NaOH is dissolved in the water and speeds up the growth of quartz crystals by hydrothermal treatment.
3. Many phases can be prepared hydrothermally at much lower temperatures than would be required for solid state reaction. For example, to produce the magnetoplumbite phase, $BaFe_{12}O_{19}$, would require heating of the oxide components, $BaCO_3$ and α-Fe_2O_3, at e.g. 1250 °C but hydrothermally can be carried out at a much lower temperature:

$$Ba(OH)_2 + \alpha\text{-FeOOH} \xrightarrow{325\,^\circ C} BaFe_{12}O_{19}$$

$SrFe_{12}O_{19}$ can be prepared similarly.

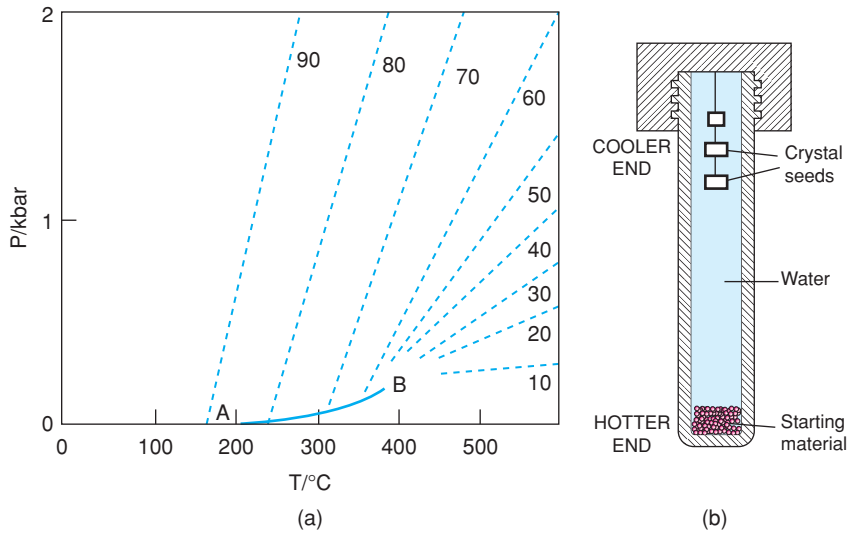


Figure 4.7 *(a) Pressure–temperature relations for water at constant volume. Dashed curves represent pressures developed inside a closed vessel; numbers represent the percentage filling of the vessel by water at ordinary* **P,T.** *Adapted from G. C. Kennedy (1950), Amer. J. Sci., 248 540. (b) Schematic of hydrothermal bomb used for crystal growth.*

4. In the Bayer process for the commercial extraction of Al_2O_3 prior to production of Al metal, bauxite is dissolved hydrothermally in NaOH solution from which gibbsite, α-$Al_2O_3{\cdot}3H_2O$, is precipitated and heated to produce anhydrous α-Al_2O_3. The process is summarised in Fig. 4.8.

Solvothermal synthesis is similar to hydrothermal synthesis but uses supercritical solvents or solvent mixtures and is finding an increasing number of applications. For example, Li_2MnSiO_4, of current interest as

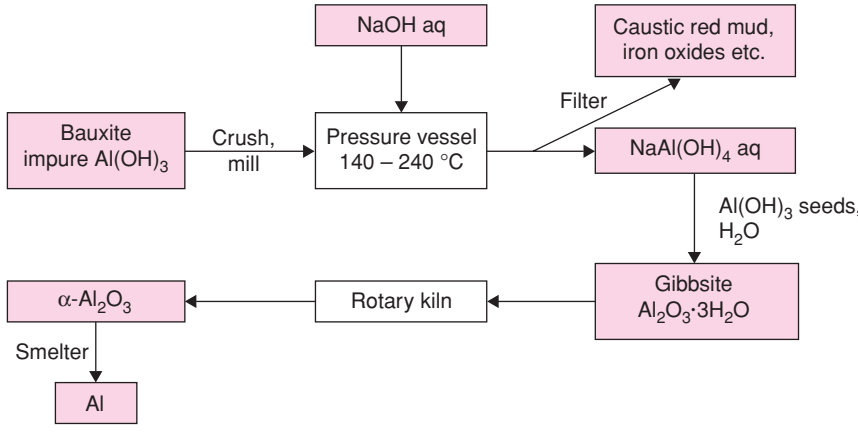


Figure 4.8 *The Bayer process for the extraction of α-Al_2O_3 from bauxite, which is impure alumina containing other oxides such as SiO_2, TiO_2 and iron oxides.*

a lithium battery cathode material, was synthesised, phase pure, at 300 °C and 38 MPa in 5 min in supercritical water–ethanol mixtures. This compound is particularly difficult to prepare by conventional solid state routes because high reaction temperatures are required and Mn^{2+} is very easily oxidised; in addition, Mn^{2+} is a rather large divalent cation and is not expected to occupy tetrahedral sites easily in the Li_2MnSiO_4 crystal structure.

4.3.6 Microwave synthesis

The use of microwave heating is well established in organic chemistry and is increasingly finding use in the synthesis of inorganic materials, especially at the nanolevel. Reaction times are orders of magnitude less than required for solid state reaction and side reactions are less problematic, leading to improved yield and reproducibility of microwave synthesis. The method can, with care, use domestic microwave ovens, but specially designed commercial systems allowing control of power input, temperature and pressure are increasingly available.

The microwave region covers the range 0.3 GHz (1 m)–300 GHz (1 mm), but most ovens are restricted to the frequency 2.45 GHz (12.25 cm). This corresponds to an energy of 1×10^{-5} eV or 1 J mol^{-1}. Absorption of radiation of this energy is able to affect molecular rotations, which effectively leads to an increase in temperature. The energies involved are far too small to affect chemical structures directly and to break bonds, but this may happen at high temperatures as a consequence of heating.

Microwave heating is different from conventional heating since it is a volumetric, or bulk, effect. Microwaves are absorbed and the sample is heated, to a certain penetration depth. Conventional heating, by contrast, requires in-diffusion of heat. The thermal conductivity of the sample, κ, controls the coefficient of thermal diffusivity, A, according to

$$A = \kappa \rho_{\mathrm{d}}^{-1} C_{\mathrm{p}}^{-1} \tag{4.1}$$

where ρ_{d} is the sample density and C_{p} its specific heat.

Microwave radiation produces an oscillating electric field, which couples to dipoles or ions in a material whose motions continuously attempt to follow the alternating field. The success of microwave heating depends on the microwave frequency being comparable to that of dipolar and ionic motions; consequently, absorption of energy and generation of heat occur. The important material properties for microwave absorption are the dielectric loss, ε'', and permittivity, ε'; ε'' is related to the sample conductivity, σ, by

$$\sigma = \omega e_0 \varepsilon'' \tag{4.2}$$

where ω is the angular frequency $2\pi f$ and e_0 is the permittivity of free space, 8.854×10^{-14} F cm^{-1}; ε' is a measure of the polarisation, P, that arises under the action of the electric field, E. The dielectric loss and permittivity are combined in the parameter $\tan\delta$, where

$$\tan\delta = \varepsilon''/\varepsilon' \tag{4.3}$$

A medium with high $\tan\delta$ is needed for efficient microwave absorption and consequent rapid heating. The penetration depth, defined as the depth at which the initial power has dropped to 50% of its original value, depends inversely on $\tan\delta$, which is why high loss materials have low penetration depth, why meals cooked in microwave ovens may still be cold at the centre and why metallic objects should not be placed in microwave ovens since their high conductivity leads to a very small penetration depth and sparking may occur at the metal surface.

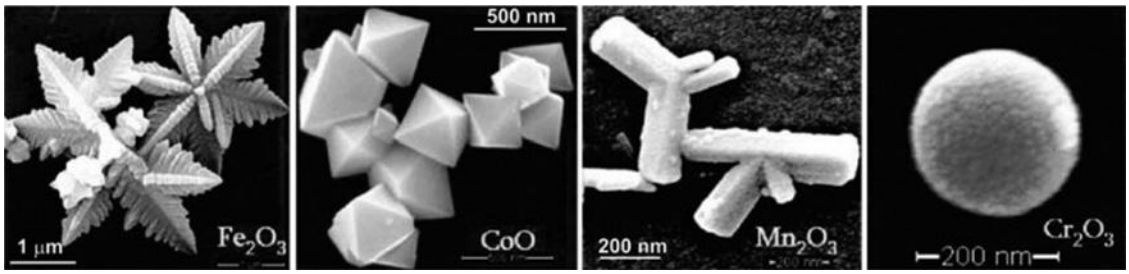


Figure 4.9 ***SEM images of various metal oxide nanostructures. Reproduced with permission from I. Bilecka and M. Niederberger,*** **Nanoscale**, 2, *1358, © 2010 Royal Society of Chemistry.*

Microwave heating finds increasing applications in materials synthesis, either as a single heating method or in combination with hydrothermal methods. Microwave heating is rapid, convenient, cheap and gives uniform nucleation of products, often with a narrow particle size distribution. A good example was the synthesis of the perovskite phases $LaMO_3$, M = Al, Cr, Mn, Fe, Co. Starting materials were metal nitrates to which 5% carbon black was added to promote microwave absorption. Heating samples in a microwave oven was found to be far more effective in terms of speed of reaction than heating in a conventional furnace.

Microwave heating has been used for the synthesis of a wide range of complex metal oxides and chalcogenides and for the crystallisation of Al_2O_3 in films and coatings produced by sol–gel methods. Such films are important for their corrosion, scratch and abrasion resistance. It is becoming an important method for controlling the size and shape of nanoparticles by varying conditions such as precursor concentration and nature/concentration of templating additives and also controlling the temperature. Some examples of various metal oxide nanostructures produced by microwave methods are shown in Fig. 4.9.

4.3.7 Intercalation and deintercalation

A very elegant way to synthesise new materials is to take an existing crystalline solid and either introduce new atoms into empty sites or selectively remove certain atoms, while keeping the overall original structure intact. These are examples of *topotactic* or *topochemical reactions* in which there is a strong 3D similarity in the structures of the starting phase (the host structure) and the product.

Intercalation species (guests) can be either ions or molecules. For intercalation and deintercalation reactions that involve addition or removal of ions (commonly Li^+, Na^+, H^+, O^{2-}), electrons must also be added or removed to preserve charge balance. These are, therefore, solid state redox processes and an automatic requirement is that the host material should be a mixed conductor so as to allow the migration of both ions and electrons. With certain host structures, it is also possible to intercalate organic molecules and even fairly large polymers.

Most intercalation hosts are layered structures with strong intralayer bonds but weak interlayer bonds. Guest species intercalate the empty space between the layers and push them apart. There are examples of 3D intercalation hosts, but in this case there is very little volume change on intercalation, which places a strict limit on the size, often small, of the species that may be intercalation guests. Some examples, Table 4.2, are as follows:

Alkali metal intercalation can be carried out:

(a) electrochemically using cells such as
Li/poly(ethylene oxide) + $LiClO_4/TiS_2$
Na/propylene carbonate + $NaI/NbSe_2$

Table 4.2 ***Examples of intercalation compounds***

Host	Guest	Comments
Graphite	K, Br_2, $FeCl_3$	Staging occurs to give KC_8, KC_{24}, etc.
TiS_2, $ZrSe_2$	Li, NH_3, amines	Li_xTiS_2 was an early prototype lithium battery cathode. Most lithium cathodes are intercalation compounds
$NiPS_3$, $VPSe_3$	Cobaltocene, $CoCp_2$	
MoO_3, V_2O_5	H, Na	Mo and V bronzes; colour changes find electrochromic applications
FeOCl	Ferrocene, $FeCp_2$	
$Zr(HPO_4)_2$	Organic molecules	
Kaolin	Organic molecules, aluminate anions	Example of a pillared clay on heating
TiO_2 (anatase)	Li	Transforms to $LiTi_2O_4$ spinel superconductor on heating
Mo_6X, X = S, Se, Te	Li	Chevrel phases; some are superconducting
$YBa_2Cu_3O_6$	O_2	$YBa_2Cu_3O_7$: 90 K superconductor

Source: Adapted with permission from CNR Rao. *Mat. Sci. Eng.* B18, 1–21, © 1993 Elsevier.

In these cases the electrolytes are non-aqueous solvents of either organic liquids such as propylene carbonate or polymers containing polar functional groups, such as poly(ethylene oxide), into which are dissolved alkali metal salts.

(b) chemically:

$$xC_4H_9Li + TiS_2 \rightarrow Li_xTiS_2 + \frac{x}{2}C_8H_{18}$$

An example of an intercalation reaction that has to be mentioned because of the incredible change in properties that occurs is the topotactic conversion of the anatase polymorph of TiO_2, which is a white insulating solid, to a lithium titanate superconductor. The first step is to react anatase with *n*-butyllithium dissolved in hexane. Li^+ ions and electrons intercalate into the anatase structure, which has open, 1D channels not unlike those in rutile, Fig. 1.37; no significant structural changes occur on intercalation of Li. The second step is to heat the lithiated anatase, Li_xTiO_2, to 500 °C; this causes structural reorganisation and gives a material with the same composition but with the spinel structure and which is superconducting with a T_c of 13 K. This insulator to superconductor transformation is almost as miraculous as the 'fishes to birds' transformation, Fig. 4.3, crafted by Escher! The reactions that take place may be summarised as

$$TiO_2 \text{ (anatase)} + xC_4H_9Li \rightarrow Li_xTiO_2 + \frac{x}{2}C_8H_{18}\text{(octane)}$$

$$Li_xTiO_2 \text{ (lithiated anatase)} \xrightarrow{500\,^\circ C} Li_xTiO_2 \text{ (spinel)}$$

n-Butyllithium is a powerful lithiating agent; effectively, it is a source of atomic Li.

A useful delithiating agent is I_2 dissolved in acetonitrile, CH_3CN, which gives reactions such as

$$LiCoO_2 + I_2/CH_3CN \rightarrow Li_xCoO_2 + LiI/CH_3CN$$

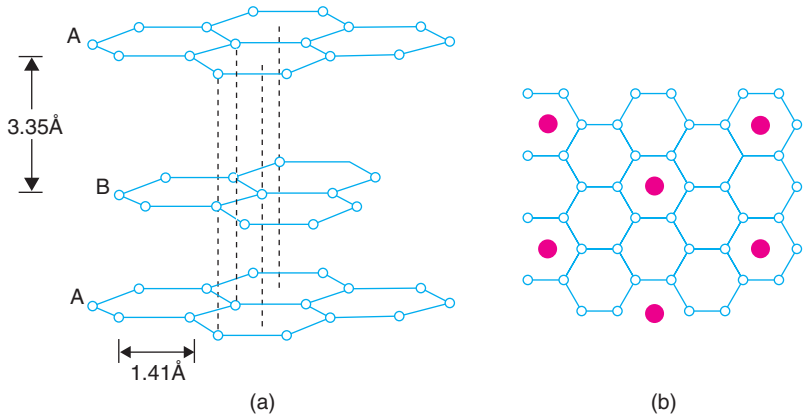


Figure 4.10 ***Structures of (a) graphite, in oblique projection showing the two-layer stacking sequence, and (b) graphite potassium, C_8K. In (b) the graphite layers are superposed in projection but the K atoms are not and various stacking sequences are possible. The structure in (b) is typical of donor and acceptor complexes of graphite.***

Products with a range of x values are obtained by controlling the amount of I_2 used. The driving force is formation of LiI which is soluble in CH_3CN.

4.3.7.1 Graphite intercalation compounds

Graphite is the classic example of a host crystal capable of intercalating a wide variety of atoms, ions and molecules. Graphite has a planar ring structure, Fig. 4.10(a); it is possible to intercalate alkali metal cations, halide anions, ammonia and amines, oxy salts and metal halides between the carbon layers. Some typical reactions and conditions for formation of graphite intercalation compounds are as follows:

$$\text{Graphite} \xrightarrow[25\,^\circ\text{C}]{\text{HF/F}_2} \text{graphite fluoride, } C_{3.6}F \text{ to } C_{4.0}F \quad \text{(black)}$$

$$\text{Graphite} \xrightarrow[450\,^\circ\text{C}]{\text{HF/F}_2} \text{graphite fluoride, } CF_{0.68} \text{ to } CF \quad \text{(white)}$$

$$\text{Graphite} + \text{K (melt or vapour)} \rightarrow \underset{\text{(bronze)}}{C_8K}$$

$$C_8K \xrightarrow[\text{vacuum}]{\text{partial}} \underset{\text{(steel blue)}}{C_{24}K} \rightarrow C_{36}K \rightarrow C_{48}K \rightarrow C_{60}K$$

$$\text{Graphite} + H_2SO_4 \text{ (conc.)} \rightarrow C_{24}{}^{+}(HSO_4)^{-}\cdot 2H_2SO_4 + H_2$$

$$\text{Graphite} + FeCl_3 \rightarrow \text{graphite}/FeCl_3 \text{ intercalate}$$

$$\text{Graphite} + Br_2 \rightarrow C_8Br$$

Most of these reactions are reversible; thus C_8K forms on exposure of graphite to molten potassium, which may be subsequently removed, in stages, under vacuum. The reactions are reversible because the structure and planarity of the carbon layers are essentially unaffected by intercalation.

The structure and electronic properties of graphite are described in Section 3.4.4.3. Here, we note that bonding within the graphite layers is strong, with an average C–C bond strength of 1.5, but adjacent layers are held together by weak van der Waals bonds, as evidenced by the relatively large inter-layer spacing of 3.35 Å. This weak bonding permits foreign species to intercalate between the layers and push them apart to, for example, 5.5 Å in C_4F, ~6.6 Å in CF and 5.41 Å in C_8K.

The structures of the intercalation compounds are often not known with certainty; the probable structure of one, C_8K, is shown in 4.10(b). The relative positions of the carbon layers are different to those in pure graphite as they form an ... AAA ... stacking sequence. K^+ ions are sandwiched between pairs of carbon rings, giving K^+ a coordination number of 12. If all such sites were occupied, the stoichiometry C_2M would result, but in C_8K only one-quarter are occupied, in an ordered fashion. The electronic structure of graphite is modified on intercalation of K since partial electron transfer from K to graphite occurs, resulting in a polar structure which may be represented as $C_8^-K^+$, Fig. 3.33.

4.3.7.2 Pillared clays and layered double hydroxides

Many other layered structures undergo intercalation reactions to form new materials, often with interesting and novel properties. Such reactions have been known for many years in clay mineralogy. Clays such as kaolinite, $Al_2(OH)_4(Si_2O_5)$, have layered silicate structures in which the layers are either electrically neutral, as in kaolinite, or charged, as in e.g. muscovite, $K[Al_2(OH)_2(AlSi_3O_{10})]$, in which the layers have a net negative charge per formula unit due to replacement of one Si^{4+} by Al^{3+}. Muscovite already contains interlayer K^+ ions to give the structure some 3D rigidity, whereas the kaolinite layers are held together by hydrogen bonding giving a mineral which is much softer (as in talcum powder).

Various ionic and molecular species can intercalate between the sheets of clay mineral structures. In *pillared clays*, PILCs, the K^+ ions of muscovite can be ion exchanged for large complex Al cations such as $[Al_{13}O_4(OH)_{24}(H_2O)_{12}]^{7+}$ and, because of their size, have the effect of pushing apart the silicate sheets. On heating the product material, the Al complex cation dehydrates and the residue forms strong covalent bonds with the silicate sheets, producing a pillared structure with a porous 2D interconnected network.

Layered materials that behave similarly to clay minerals are *layered double hydroxides*, LDHs, and zirconium phosphonates. LDHs have positively-charged metal hydroxide layers of general formula $[M^{2+}_{1-x}M^{3+}_x(OH)_2]^{x+}$: M^{2+} = Mg, Ni, etc., M^{3+} = Al, Fe, etc., and x in the range 0.2–0.33 for single-phase solid solutions. In these, similar-sized M^{2+} and M^{3+} cations are disordered over the octahedral sites in the layers of edge-sharing octahedra, Fig. 4.11. These LDHs contain a mixture of water molecules and anions, such as carbonate, nitrate, halide, $Fe(CN)_6^{3-}$ and oxalate, between the sheets.

LDHs find applications as ion exchangers, catalysts and sorbents and increasingly are of interest for the intercalation and release of organic and biologically active molecules. Clay minerals are well known for their ability to intercalate organic molecules and LDHs behave similarly under appropriate conditions. Examples are shown schematically in Fig. 4.12 of the intercalation of either organic molecules which can be subsequently polymerised *in situ*, or of the direct intercalation of the same polymer molecules.

Clay minerals and LDHs with low charge density on the octahedral sheets are readily able to exfoliate and separate the sheets, especially if they are only held together by weak hydrogen bonds, and subsequently restack the sheets. With LDHs, this process is not so easy given the charged nature of the layers but is possible in certain cases, as shown schematically in Fig. 4.12. Recent research shows that is it possible to intercalate a range of biologically-active molecules such as vitamins and amino acids and, in the future, LDHs may find applications in medicine as drug and gene delivery systems.

LDH materials can also take part in anion-exchange reactions with polyoxometallate anions such as $PW_{12}O_{40}^{3-}$, $SiV_3W_9O_{40}^{7-}$ and the Keggin ion, $H_2W_{12}O_{40}^{6-}$, to form pillared products. In some cases, it

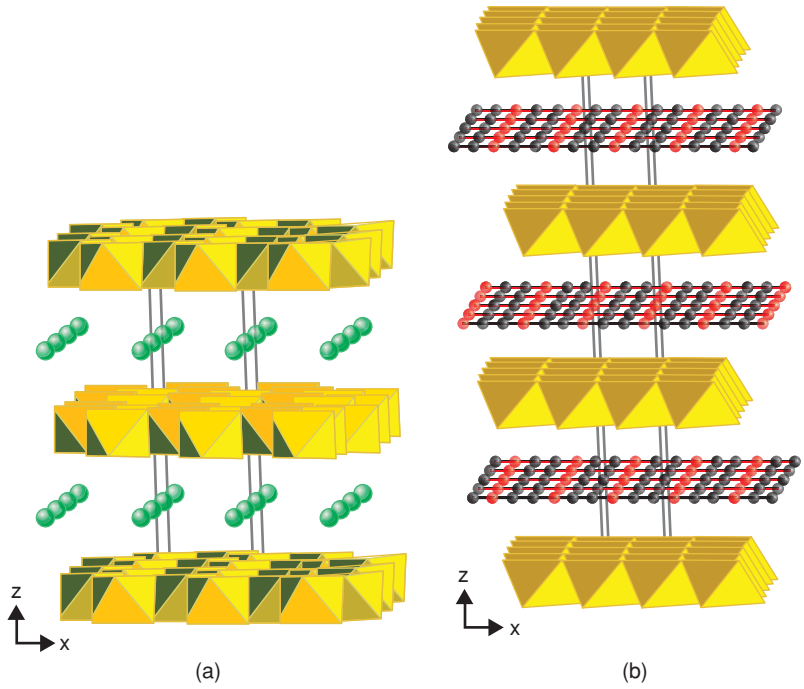


Figure 4.11 ***Stacking of octahedral layers in LDHs showing (a) two-layer repeat in a hexagonal polymorph and (b) three-layer repeat in a rhombohedral polymorph. Reproduced with permission from A. I. Khan and D. O'Hare,*** **J. Mat. Chem.,** *12, 3191,* ***© 2002 Royal Society of Chemistry.***

is not possible to introduce large anions in a single step but a two-stage reaction, introducing anions of increasing size, gradually pushes apart the layers of the host structure.

4.3.7.3 Synthesis of graphene

Graphene is a single sheet of carbon atoms that form hexagonal rings, as in graphite. It had been thought that a single, flat graphene layer would be intrinsically unstable until the amazingly simple discovery by Geim and Novoselov at Manchester University, for which they were awarded the Nobel Prize in Physics in 2009, that individual graphene layers could be peeled off a graphite block using Scotch tape. This method, and subsequent discovery of several unusual physical properties exhibited by graphene, is the latest stage in a remarkable series of fundamental advances in the chemistry and physics of carbon, as shown by the following timeline:

1954	Synthesis of artificial diamonds by HPHT
1981	Synthesis of diamond films by low-pressure gas-phase deposition
1985	Discovery of C_{60} buckyballs
1992	Synthesis of carbon nanotubes
2004	Synthesis and characterisation of graphene

This rich variety of carbon structures arises for two reasons: first, because of the ability of C to exhibit three hybridised bonding states, i.e. *sp* hybridised carbon characterised by linear coordination and including

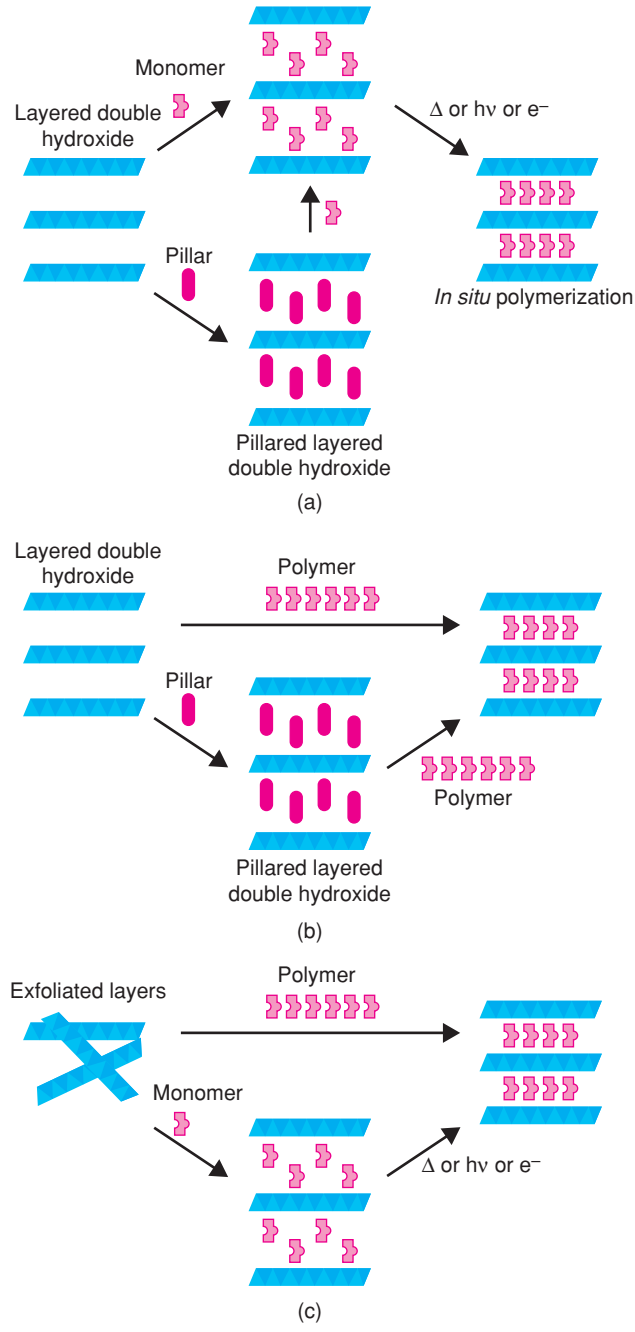


Figure 4.12 ***Synthesis of polymer-intercalated LDH showing (a)* in situ *polymerisation of an intercalated monomer, (b) direct intercalation of a polymer and (c) restacking of the LDH layers after exfoliation. Reproduced with permission from A. I. Khan and D. O'Hare,* J. Mat. Chem., 12*, 3191, © 2002 Royal Society of Chemistry.***

one triple bond, sp^2 hybridisation characterised by triangular coordination and one double bond and sp^3 hybridisation with tetrahedral coordination and single bonds to four neighbours; and second, because of the ease of *catenation* in which carbon atoms form bonds to other C atoms leading to a large variety of 2D and 3D structures with a carbon backbone.

Graphene was first synthesised by the simple mechanical route of peeling layers from a piece of graphite, which were then transferred to a silicon substrate. High-quality graphene layers can be obtained in this way and have been much used for property measurements. Since then, numerous chemical routes have been devised for exfoliating or separating graphene sheets, or multiple sheets, from graphite. This is an ongoing topic of research to establish routes for preparing high-quality graphenes which are free from impurities and which are either single-layer graphene or graphenes with a controlled number of layers.

One method is to first prepare single-layer graphene oxide, GO, by a sequential chemical treatment of graphite powder with H_2SO_4, $NaNO_3$, $KMnO_4$ and H_2O_2, followed by ultrasonication to separate individual layers of GO and form a colloidal suspension in water. The suspension is then reduced with hydrazine hydrate and most, or all, of the oxygen is removed. Other reducing agents can also be used. By varying the conditions, single-layer or multilayer graphene may be obtained. Other methods involve: solvothermal treatment with Na in ethanol; ultrasonication of GO in surfactant–water solution or *N*-methylpyrrolidone; microwave-assisted exfoliation and reduction of GO with hydrazine hydrate.

Numerous gas-phase routes to graphenes have also been devised. These include:

- thermal exfoliation of GO by briefly heating at 1050 °C in Ar;
- heating nanodiamond in an inert atmosphere;
- arc evaporation of graphite in the presence of H_2;
- epitaxial thermal decomposition of SiC in which the outer layer of an oriented single crystal of SiC in the (001) orientation is heated to remove the outer layer of Si and leave a graphene surface layer;
- Chemical vapour deposition, CVD, of various hydrocarbons on metal (Ni, Cu) substrates.

4.3.8 Example of a difficult synthesis made possible by chimie douce methods: $BiFeO_3$

$BiFeO_3$ is a material of great current interest since at room temperature it is both ferroelectric and antiferromagnetic and is a rare example, as such, of a *multiferroic* material offering numerous possible applications involving coupling between the effects of electric and magnetic fields. Synthesis of high-purity samples of $BiFeO_3$ by solid state reaction is difficult since small amounts of impurity phases, $Bi_2Fe_4O_9$ and $Bi_{25}FeO_{39}$, are often present even after many heat treatment cycles.

The difficulty in solid state synthesis can be understood from the Bi_2O_3–Fe_2O_3 phase diagram (see Fig. 7.29). Although $BiFeO_3$ appears on the phase diagram as a thermodynamically stable phase, it is not particularly stable and melts incongruently at 933 °C to give $Bi_2Fe_4O_9$ and liquid. The phase diagram shows that of the three bismuth ferrite phases, $Bi_2Fe_4O_9$ has the highest melting temperature and is the most stable. The driving force to form pure $BiFeO_3$ in mixtures is rather small and explains why phases such as $Bi_2Fe_4O_9$ are difficult to eliminate once they have formed. Additional problems with solid state synthesis of $BiFeO_3$ concern the volatility of Bi_2O_3 on heating at temperatures above ~800 °C and the variable valence state of Fe leading to other possible unwanted side-products.

Many chimie douce methods have been used to synthesise $BiFeO_3$ and all make use of either fine-scale reaction mixtures or single-source homogeneous precursors. For example, fine-scale mixed hydroxide precursor is obtained by hydroxide precipitation from aqueous nitrate solutions:

$$Bi^{3+} + Fe^{3+} + 6OH^- \rightarrow Bi(OH)_3 + Fe(OH)_3$$

Table 4.3 ***Synthesis of $BiFeO_3$ by chimie douce methods***

Synthesis method	Product size range (nm)	Shape control	Reaction temperature/°C	Reaction time	Subsequent anneal/°C
Coprecipitation	30–100	Poor	20–90	Minutes	550–700
Microemulsion	15–40	Good	20–70	Hours	400
Hydrothermal	>100	Very good: powders, spindles, nanoflakes	150–200	Hours/day	–
Sol–gel	15–150	Good: powders, fibres, nanotubes	20–90	Hours/day	400–600
Microwave-assisted hydrothermal	>100	Good: powders, nanocubes	100–200	Minutes	–
Mechanosynthesis	>100	Fair: powers	~25	Hours	–

Source: Adapted with permission from R. Safi and H. Shokrollahi, *Prog. Solid State Chem.*, **40**, 6–15, © 2012 Elsevier.

The hydroxide precipitate is filtered, dried and decomposed by heating. Coprecipitation is a simple method but it can be difficult to achieve precipitation of both components simultaneously as fine-scale mixed precipitates. Thus, if the component hydroxides have very different solubilities and precipitate sequentially on addition of base, a very inhomogeneous hydroxide may precipitate, which is the opposite of the desired effect. Control of the precipitation process may be possible through careful control of pH.

$BiFeO_3$ can be synthesised by sol–gel methods, both as fine powders and, by control of sol viscosity, as ultra-fine fibres. It can also be prepared by hydrothermal synthesis and by combined microwave-assisted hydrothermal synthesis at e.g. 200 °C for 30 min. Homogeneous liquid precursors can be used to produce powders by solution combustion synthesis (SCS), in which the metal components, fuel and oxidiser are all present in the correct amounts in the precursor solution prior to ignition of a self-sustained exothermic reaction to form $BiFeO_3$ product.

Crystallisation of $BiFeO_3$ from aqueous solution can be facilitated by the application of high-power ultrasound waves with frequencies between 20 kHz and 10 MHz. This technique, known as *sonochemistry*, is based on the acoustic cavitation resulting from formation, growth and implosion of nanometre-sized solvent bubbles in the precursor solution. Very high temperatures on a nanosecond timescale are generated during bubble collapse and these are sufficient to break chemical bonds within the precursor components, leading to subsequent rapid crystallisation of $BiFeO_3$.

Another method involving the formation of droplets within a matrix liquid phase is the preparation of *microemulsions* or *micelles*. In the *reverse microemulsion method* for $BiFeO_3$ synthesis, nano-sized water droplets are dispersed in a continuous oil phase together with a surfactant. The water droplets containing precursor species act as a microreactor for synthesis of $BiFeO_3$.

Some of the methods used to synthesise $BiFeO_3$, the experimental conditions and the nature of the products are summarised in Table 4.3.

4.3.9 Molten salt synthesis, MSS

Molten salt synthesis uses low-melting, water-soluble salts such as alkali halides, carbonates or sulfates as a liquid medium in which one or more of the reactants are soluble but the desired product is insoluble. The liquid

assists rapid diffusion of reactant species at low temperatures, 300–800 °C depending on the salt composition, and can give a high-quality product in a shorter time and at much lower temperature than achieved by solid state synthesis. At the end of the reaction, the molten salts can be dissolved, leaving behind the solid product.

There are two generally accepted mechanisms of MSS, both of which are based on relative solubilities of the reactants. In dissolution–precipitation, all the reactants dissolve initially and the product then precipitates. In template-controlled growth, one reactant is insoluble and serves as a template upon which the product can form and grow; this is a more controlled mechanism and can allow specific particle sizes and morphologies to be produced. Numerous mixed oxide materials have been produced by MSS. These include $LaAlO_3$ formed after 3 h at 630 °C in KCl–KF liquid mixtures and $MgAl_2O_4$ formed after 3 h at 1100 °C in mixed chloride melts; other examples include $CaZrO_3$, $Yb_2Ti_2O_7$ and TiC coatings.

4.4 Gas-Phase Methods

4.4.1 Vapour-phase transport

The essential feature of this method is the formation of a volatile, unstable intermediate that contains at least one of the elements in the desired final product. The method may be used to synthesise new compounds, to grow single crystals or to purify a compound. It involves chemical transport along a temperature gradient via the vapour phase. The method consists of a tube, usually of silica glass, with the reactants, A, at one end, which is sealed either under vacuum or with a small amount of the transporting agent, B, Fig. 4.13(a). The tube is placed in a furnace such that a temperature gradient of e.g. 50–100 °C exists inside the tube. In the simplest case, A and B react to form gaseous intermediate AB, which permeates the whole tube and decomposes somewhere in the tube, at either higher or lower temperature to redeposit A, Fig. 4.13(b).

The method depends on the existence of a reversible equilibrium between A, B and AB:

$$A + B \overset{K}{\rightleftharpoons} AB$$

The equilibrium constant, K, should be small; if K is zero then no AB would form and vapour-phase transport could not occur; if K is very large, AB would be stable and would not subsequently decompose. If the formation of AB is endothermic, it forms at higher temperatures and decomposes at lower temperatures, i.e. equilibrium is displaced to the right with increasing temperature and to the left with decreasing temperature. The exact opposite happens if AB forms exothermically. Hence the tube is arranged for the reactants A to be at either the hot end or cold end, respectively. Since K is temperature dependent, a concentration gradient of gaseous AB builds up along the length of the tube, which provides the driving force for gaseous diffusion and, therefore, for chemical transport.

An example of an endothermic reaction which can be used for chemical transport is the formation of gaseous PtO_2 at 1200 °C or higher:

$$Pt(s) + O_2 \rightleftharpoons PtO_2(g)$$

The PtO_2 diffuses to lower temperatures where it may deposit well-formed crystals of platinum metal. In furnaces that contain heating elements of platinum, it is common to find platinum crystals deposited on cooler parts of the furnace wall; these form by a process of vapour-phase transport. A nice (but expensive!) example from the author's laboratory is shown in Fig. 4.13(c).

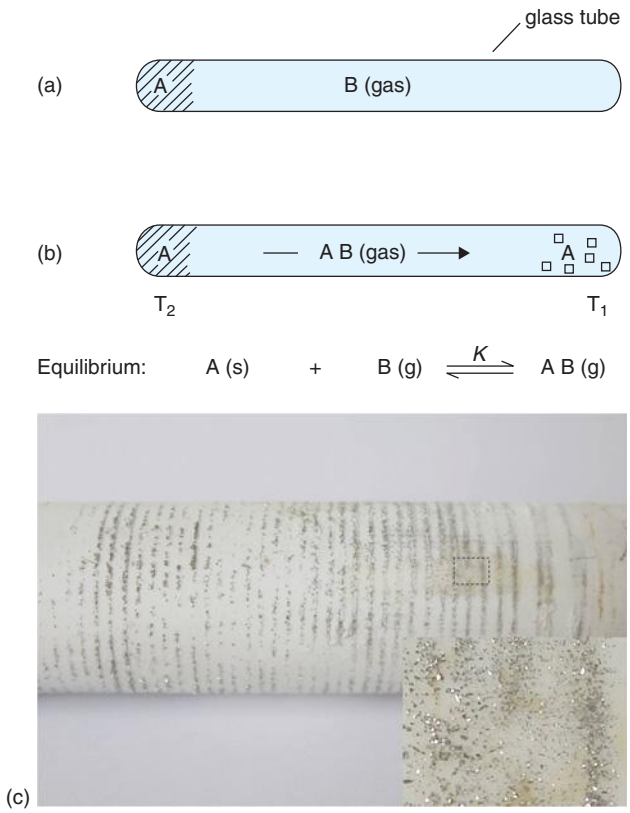


Figure 4.13 ***(a, b) Simple vapour-phase transport experiment for the transport and growth of crystals of A via a volatile intermediate AB. (c) Part of a ceramic furnace tube in which the Pt element coil had transported to the tube surface and deposited as single crystals of Pt.***

The *van Arkel method* for the purification of certain metals makes use of an exothermic reaction between metal and iodine to form a gaseous iodide, e.g.

$$Cr + I_2 \rightleftharpoons CrI_2(g)$$

Since formation of CrI_2 is exothermic, Cr metal is redeposited at a higher temperature. Other metals which may be purified by this method include Ti, Hf, V, Nb, Cu, Ta, Fe and Th. By this method, the metals may be extracted from their carbides, nitrides and oxides.

The *Mond process* for the purification of nickel involves an exothermic reaction of Ni with CO to form gaseous nickel carbonyl:

$$Ni + 4CO \underset{230\,^\circ C}{\overset{50\,^\circ C}{\rightleftharpoons}} Ni(CO)_4$$

This is an old industrial process for extracting high-purity Ni from impure Ni obtained by reduction of NiO, but is falling out of favour because of environmental issues concerning Ni contamination and also human

health concerns over the toxicity of CO and $Ni(CO)_4$. Nevertheless, this is a good example of chemical vapour transport.

An elegant variation, to transport *two* substances in opposite directions along a temperature gradient, is possible if one reaction is exothermic and the other endothermic. For instance, WO_2 and W may be separated using I_2 and H_2O as the gas phase: W is deposited at 1000 °C and WO_2 at 800 °C. The reactions involved are

$$WO_2(s) + I_2(g) \underset{800\,^\circ C}{\overset{1000\,^\circ C}{\rightleftharpoons}} WO_2I_2(g)$$

and

$$W(s) + 2H_2O(g) + 3I_2(g) \underset{800\,^\circ C}{\overset{1000\,^\circ C}{\rightleftharpoons}} WO_2I_2(g) + 4HI(g)$$

Another example is the separation of Cu and Cu_2O using HCl transporting agent. The reactions involved are

$$Cu_2O(s) + 2HCl(g) \underset{900\,^\circ C}{\overset{500\,^\circ C}{\rightleftharpoons}} 2CuCl(g) + H_2O(g)$$

and

$$Cu(s) + HCl(g) \underset{500\,^\circ C}{\overset{600\,^\circ C}{\rightleftharpoons}} CuCl(g) + 1/2H_2(g)$$

Since CuCl forms exothermically from Cu_2O and endothermically from Cu, Cu_2O is redeposited at a higher temperature and Cu at a lower temperature.

The above examples are simple cases of transport in which reactant and product are the same. For use as a preparative method, transport is coupled with a subsequent reaction, such as:

At T_2:	$A(s) + B(g) \rightarrow AB(g)$
At T_1:	$AB(g) + C(s) \rightarrow AC(s) + B(g)$
Overall:	$A(s) + C(s) \rightarrow AC(s)$

There are many examples of using this method to prepare binary, ternary and even quaternary compounds. Some examples (after Schäfer, 1971, who was the pioneer of vapour phase transport methods of chemical synthesis) are as follows:

(a) *Preparation of Ca_2SnO_4*. CaO and SnO_2 react slowly according to

$$2CaO + SnO_2 \rightarrow Ca_2SnO_4$$

In the presence of gaseous CO, SnO_2 converts to gaseous SnO which is chemically transported:

$$SnO_2(s) + CO \rightarrow SnO(g) + CO_2$$

and subsequently reacts with CaO and CO_2 to give Ca_2SnO_4 and CO.

(b) *Preparation of nickel chromite*, $NiCr_2O_4$. Reaction of NiO and Cr_2O_3 speeds up in the presence of oxygen since gaseous CrO_3 forms which migrates to the NiO:

$$Cr_2O_3(s) + 3/22O_2 \rightarrow 2CrO_3(g)$$
$$2CrO_3(g) + NiO(s) \rightarrow NiCr_2O_4(s) + 3/2O_2$$

(c) *Preparation of niobium silicide, Nb_5Si_3*. Metallic Nb and SiO_2 do not react under vacuum at, e.g., 1100 °C but, in the presence of traces of H_2, gaseous SiO forms which migrates to the Nb:

$$SiO_2(s) + H_2 \rightarrow SiO(g) + H_2O$$
$$3SiO(g) + 8Nb \rightarrow Nb_5Si_3 + 3NbO$$

Alternatively, in the presence of I_2, gaseous NbI_4 is formed and is transported to the SiO_2:

$$Nb(s) + 2I_2 \rightarrow NbI_4(g)$$
$$11NbI_4 + 3SiO_2 \rightarrow Nb_5Si_3 + 22I_2 + 6NbO$$

(d) *Preparation of aluminium sulfide, Al_2S_3*. Al and S react slowly at, e.g., 800 °C, since liquid Al becomes coated with a skin of Al_2S_3, which acts as a diffusion barrier to further reaction. In the presence of I_2, however, and with a temperature gradient of 100 °C, Al_2S_3 separates as large, colourless crystals at the cool, 700 °C, end. The product, Al_2S_3, is effectively transported via formation of gaseous AlI_3:

$$2Al + 3S \rightarrow Al_2S_3$$
$$Al_2S_3(s) + 3I_2 \rightarrow 2AlI_3(g) + 3/2S_2(g)$$

(e) *Preparation of Cu_3TaSe_4*. Cu, Ta and Se, are heated at 800 °C in the presence of I_2; transport occurs to give Cu_3TaSe_4 at 750 °C, presumably by means of a gaseous complex iodide intermediate.

(f) *Preparation of zinc tungstate, $ZnWO_4$*. ZnO and WO_3 are heated at 1060 °C in Cl_2 gas. Intermediate volatile chlorides form which are transported and crystals of $ZnWO_4$ are deposited at 980 °C.

These examples illustrate the importance of the gas phase as a transporting agent and its subsequent influence on reaction rates. Gases react more quickly than solids because mobilities are increased. In reality, the gaseous phase is often important in normal 'solid state' reactions under isothermal conditions, where it may act as a rapid means of transporting matter from one region of the reaction mixture to another.

4.4.2 Chemical vapour deposition, CVD

CVD is an extremely important way of making high-purity thin films and coatings for industrial applications, especially in electronics, and also for fundamental scientific research. Conceptually, it is simple: precursor molecules containing the elements of interest are decomposed in the gas phase and the products deposit as thin films on every available object in the vicinity.

Early work focused on simple precursor molecules such as volatile, unstable hydrides, e.g. SiH_4, and metal alkyls, e.g. $Al(CH_3)_3$. For the deposition of elements, only one precursor is necessary, but to deposit compounds, e.g. GaAs, a mixture of precursors is needed. An alternative is to use single-source precursors which contain all the elements of interest and in the correct ratio. This necessitates the design of organometallic molecules which decompose cleanly to give the desired product.

Various acronyms are used to describe the CVD technique; thus, MOCVD refers to the *m*etal–*o*rganic nature of the precursors. The precursors can be decomposed in various ways, by the action of heat or by irradiation. Some examples are:

$$Si(CH_2CH_3)_4 + 14O_2 \longrightarrow SiO_2 + 8CO_2 + 10H_2O$$

$$GeH_4 \xrightarrow{h\nu} Ge + 2H_2$$

$$CrCl_2 + H_2 \xrightarrow{600\,^\circ C} Cr + 2HCl$$

$$W(CO)_6 \rightarrow W + 6CO$$

$$Fe(C_5H_5)_2 + 7/2O_2 \rightarrow Fe + 10CO + 5H_2O$$

$$Ga(CH_3)_3 + AsH_3 \xrightarrow[h\nu]{\text{heat}} GaAs + 3CH_4$$

In many applications, such as the fabrication of multilayer semiconductor devices, it is necessary for the films to have the correct structural orientation and be coherent with the underlying layer, in which case *vapour-phase epitaxy*, MOVPE, is essential in the growth mechanism.

General problems associated with precursor molecules include: (i) the pyrophoric nature of alkyls; (ii) the water sensitivity of alkyls, leading to possible oxygen contamination of the thin-film product; (iii) toxicity of hydrides; (iv) nucleation of reaction products inside the reaction chamber prior to deposition as a growing film, the so-called *snowing effect*, which occurs because the alkyls, with their coordinative unsaturation, react readily with electron pair donors such as Group V and VI hydrides; (v) the need to control accurately the precursor mix to achieve the desired film stoichiometry and homogeneity; and (vi) variations in volatility or reactivity of the reactants. These problems are not insurmountable, however, and, for instance, high-quality GaAs films have been prepared and used in devices.

New single-source precursors are being continually developed. The attributes required for use in MOCVD are that they should be volatile, have stable ligands (to avoid charring), pyrolyse the metal–ligand bonds cleanly (to retain desired stoichiometry of product and avoid contamination) and be of low toxicity. They also often have other natural advantages such that pre-reaction and snowing are less of a problem and the precursors are less reactive towards O_2 and H_2O since they are less coordinatively-unsaturated than metal alkyls.

Some examples of single-source precursors, their synthesis and decomposition products are given in Fig. 4.14. Next, we deal with two materials prepared by CVD methods that are of great technological importance.

4.4.2.1 Amorphous silicon

Amorphous silicon, a-Si, is the key materials component in a variety of technologies that utilise photoconductivity. The critical advance which made these technologies possible was the discovery by Dundee physicists, Spear and Le Comber, that a-Si, when contaminated by hydrogen impurity, could be doped to make it n- or p-type, in the same way that crystalline Si can be doped. This, coupled with the possibility of making a-Si as large-area thin films, led to applications of a-Si in, for instance, photocopiers and solar panels.

First, let us review the characteristics of crystalline Si. Si has a band gap of $\sim$1.1 eV and can be made n- or p-type by doping with, e.g., P or Al (see Fig. 8.11 and Fig. 8.12). For device applications, high-purity, defect-free crystals are required which are doped selectively to create, for instance, p–n junctions. These are

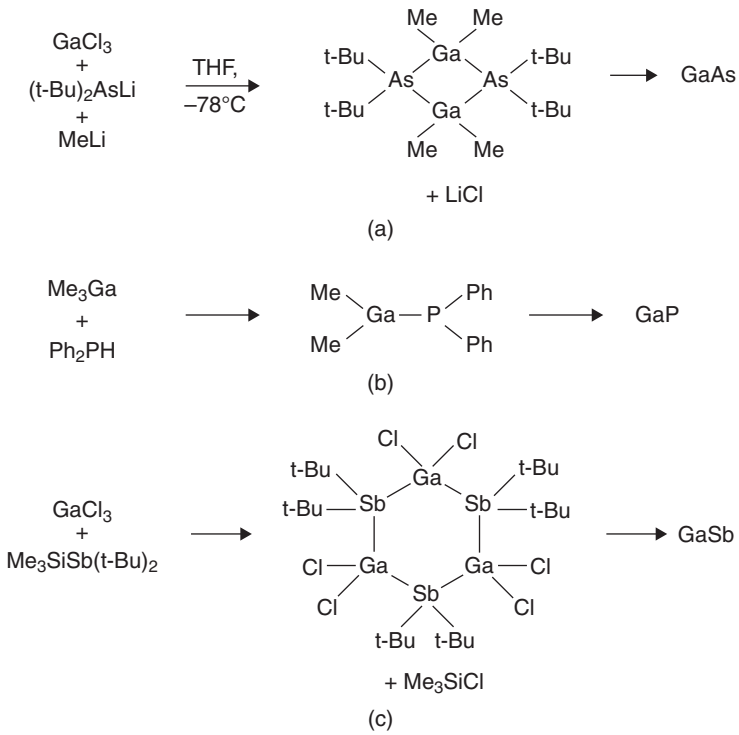


Figure 4.14 ***Single-source precursor molecules for MOCVD.***

either used as thin wafers cut from a single crystal boule or are prepared as epitaxial layers. There is, however, an inherent practical limitation to the size (area) to which single crystals of acceptable quality can be grown.

Large-area thin films of Si can be deposited by CVD techniques but are usually amorphous. Early attempts to add Group III or Group V dopants to induce p- or n-type semiconductivity were unsuccessful since, although the dopants could be added, they had little effect on the electrical properties.

The discovery that triggered the interest in a-Si technology and its applications resulted from the (unintentional) fabrication of impure, hydrogen-doped a-Si. SiH_4 was used as a precursor, which decomposed in a CVD chamber, under the action of an r.f. plasma, to yield hydrogenated amorphous silicon, a-Si:H; unlike pure a-Si, it was found that a-Si:H can be doped to make it n- or p-type.

The differences between a-Si and a-Si:H are shown schematically in Fig. 4.15. Both a-Si and a-Si:H are built of random networks of tetrahedrally coordinated Si atoms, i.e. they are non-crystalline analogues of the diamond structure adopted by crystalline Si in which each bond is an electron pair, covalent single bond. In a-Si, some bonds are broken, leaving three-coordinate Si with an additional, single, unpaired electron (a). These unpaired electrons, often referred to as 'dangling bonds', are readily available for covalent bond formation and, in a-Si:H, most form Si–H bonds (e). The band structures are shown in (b) and (f); valence and conduction bands are present, as with crystalline Si. In a-Si, however, the dangling bonds form localised states, or discrete energy levels at various energies within the forbidden band gap; these can act as trap states on attempting to dope a-Si and make it n- or p-type. Thus, if n-type (Group V) dopants are added, the donor electrons may enter initially the conduction band but subsequently fall into inter-band traps associated with

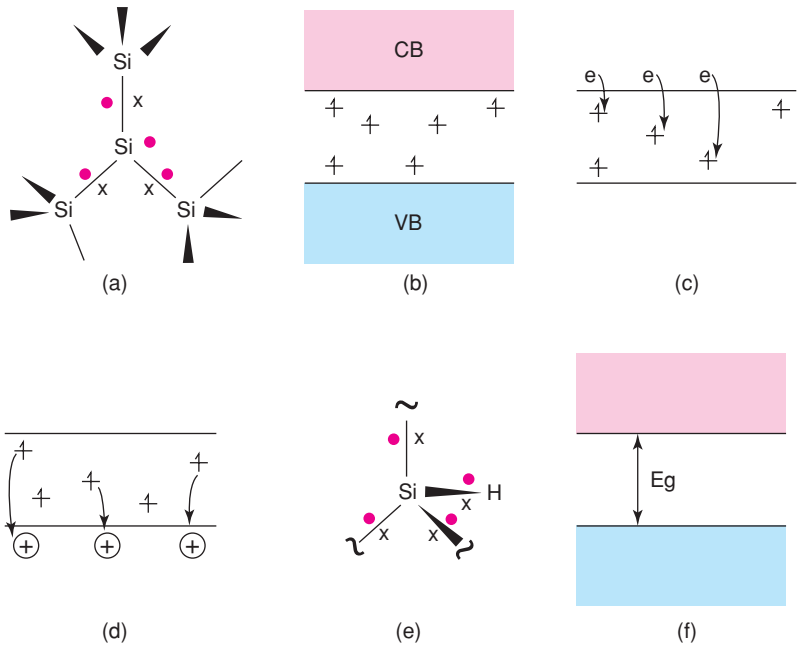


Figure 4.15 ***(a) A three-coordinate Si atom in a-Si; (b) the band structure showing inter-gap states; (c) and (d) inter-gap states can act as traps; (e) an Si-H bond within the structure of a-Si:H; (f) band structure of a-Si:H showing the absence of inter-gap states.***

a dangling bond (c). Similarly, attempts to create p-type a-Si are ruined because electrons in the inter-band levels can drop down into the positive holes that have been generated in the valence band (d).

The band structure of a-Si:H is different because the band gap is 'clean' and there are no inter-band levels associated with dangling bonds (f). Consequently, a-Si:H can be doped to make it n- or p-type.

4.4.2.2 Diamond films

Synthetic diamonds were first prepared in the 1950s by scientists at General Electric in the USA. From the phase diagram for carbon, Fig. 4.16(a), diamond is the equilibrium, thermodynamically stable polymorph only at high pressures and high temperatures (HPHT). In order to prepare synthetic diamonds, extreme conditions, 60 000 atm pressure at 1500 °C, are required. The resulting 'industrial' diamonds find applications in, for instance, machine tool tips.

At the time, it was thought that HPHT conditions were essential for diamond synthesis and there seemed no way in which the constraints imposed by the phase diagram could be avoided. Now the situation has changed totally. It is possible to prepare diamond films under ordinary laboratory conditions using CVD, by burning certain gas mixtures in an oxyacetylene flame. The individual crystals in the films are usually small and not of gem quality, but nevertheless, the films offer many possibilities for industrial applications, Table 4.4. This is because diamond has many unique and useful properties.

The thermal conductivity of diamond is several times higher than that of metals such as Ag, but electrically it is an insulator. It is therefore used as an insulating heat sink. Diamond has the highest transparency in the IR region of the electromagnetic spectrum of any substance prepared to date and has applications in

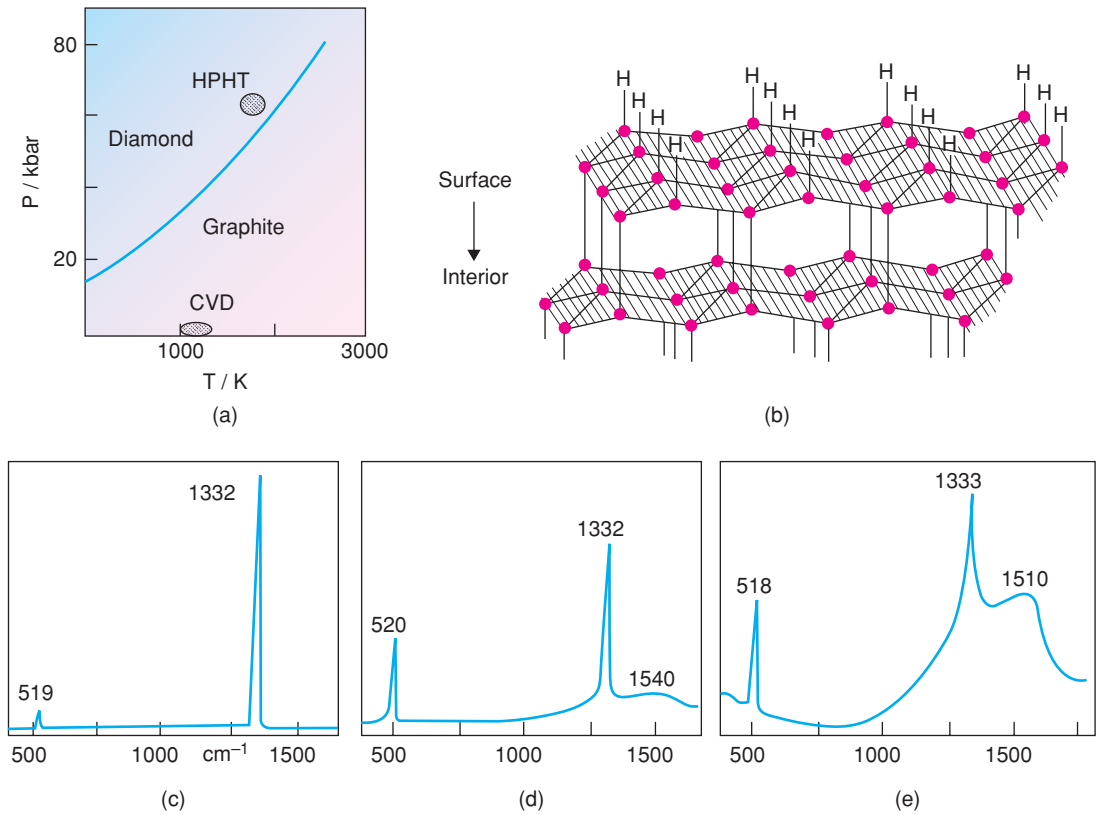


Figure 4.16 ***(a) Phase diagram for carbon; (b) schematic surface of diamond film and growth mechanism; (c)–(e) laser Raman spectra of diamond films of variable quality.***

IR-transmitting windows and lenses. These properties, coupled with its legendary hardness (it is the reference material, with a hardness of 10 on the Mho scale), resistance to chemical attack and radiation damage, make it exceptionally valuable technologically, provided that it can be made cheaply and in a suitable form.

The CVD synthesis of diamond requires a hydrocarbon precursor such as methane, CH_4, or acetylene, C_2H_2, hydrogen and a means of activation such as microwaves, heat or r.f. induction. Typically, the process

Table 4.4 ***Diamond coatings, their properties and possible areas of application***

Property	Application
Hardness	Machine tools
Low thermal expansion	Electronic substrate
High thermal conductivity	Abrasive
Semiconductivity (doped)	Electronic devices
Transparency in visible/IR region	Window/lens
Radiation resistance	Microwave devices
Large refractive index	Electrooptic devices

operates at 600–1000 °C and at a reduced pressure of 0.1 atm. Diamond films are deposited inside the reaction chamber on a substrate, such as a glass slide, with growth rates of 1–10 μm h^{-1}. This synthesis method does not involve transformation of graphite to diamond but instead involves deposition from the gas phase onto a growing surface. Deposition is surface controlled and surface free energies are very different from bulk free energies. Under certain conditions, it is easier, kinetically, for the depositing carbon atoms to form a diamond-like rather than a graphitic structure, thereby avoiding the phase diagram constraints of thermodynamic equilibrium referred to earlier. The resulting diamonds are metastable, rather than in the thermodynamic equilibrium state, as illustrated in Fig. 7.2, but this is true for all diamonds, even gemstones. Ladies, another reason to inspect your diamonds regularly is to check that they have not spontaneously changed into soot, as they should according to thermodynamics!

The mechanism by which diamonds grow from the gas phase is illustrated in Fig. 4.16(b). Carbon is tetrahedrally coordinated, sp^3 hybridised in diamond and the outermost surface of the diamond crystallites in a film is likely to be composed of C–H bonds. An important component of the gas mixture in the CVD chamber is H_2 which, 'under activation', dissociates to give atomic hydrogen or hydrogen radicals, $H^{\bullet}$. These radicals abstract H atoms from the diamond surface, leaving carbon radicals that readily react with species such as acetylene:

$$-C-H + H^{\bullet} \rightarrow \ -C^{\bullet} + H_2$$
$$-C^{\bullet} + C_2H_2 \rightarrow \ -C-CH = CH^{\bullet}$$

In this way, the diamond surface grows by the gradual attachment of hydrocarbon species which subsequently reorganise to form a fresh layer with the diamond structure. Other organic radicals and molecules may also be involved in diamond nucleation and growth; further research is needed for the mechanisms to be fully understood.

A valuable technique for assessing the quality of diamond films is laser Raman spectroscopy; sp^3 carbon has a characteristic peak at 1332 cm^{-1} whereas sp^2 carbon has a broad band at 1500–1600 cm^{-1}. Spectra are shown in Fig. 4.16 for three films of variable quality: (c) is of high quality, (d) has some graphitic impurity but is largely diamond-like and (e) is mixed diamond- and graphite-like. This provides an excellent example of the use of laser Raman spectroscopy as a fingerprint technique for inorganic solids (see Chapter 6).

4.4.3 Sputtering and evaporation

The apparatus used for sputtering is outlined in Fig. 4.17(a). Basically, it consists of a bell jar which contains a reduced pressure, 10^{-1}–10^{-2} Torr, of an inert gas, Ar or Xe. This gas is subjected to a potential drop of several kilovolts, creating a glow discharge from which positive ions are accelerated towards the cathode (target). These high-energy ions remove (sputter) material from the cathode, which then condenses on the surroundings, including the substrates to be coated, which are placed in a suitable position relative to the cathode. The mechanism of sputtering involves the transfer of momentum from the gaseous ions to the cathode in such a way that atoms or ions are ejected from the cathode. Modern sputtering equipment has various refinements, including the means to prevent permanent contamination of the substrate by inert gas atoms or ions.

The simpler vacuum evaporation method is shown in Fig. 4.17(b). The system operates under high vacuum, 10^{-6} Torr or better. Material from the evaporation source is converted into the gaseous phase by heating or electron bombardment, which then deposits on the substrate and its surroundings as a film. Various substrate materials are used, depending on the subsequent application. For electronic applications, the substrate acts both as a mechanical support and as an electrical insulator. Typical substrates are ceramics (Al_2O_3), glass, alkali halides, Si, Ge, etc. Evaporation source materials include metals, alloys, semiconductors, insulators

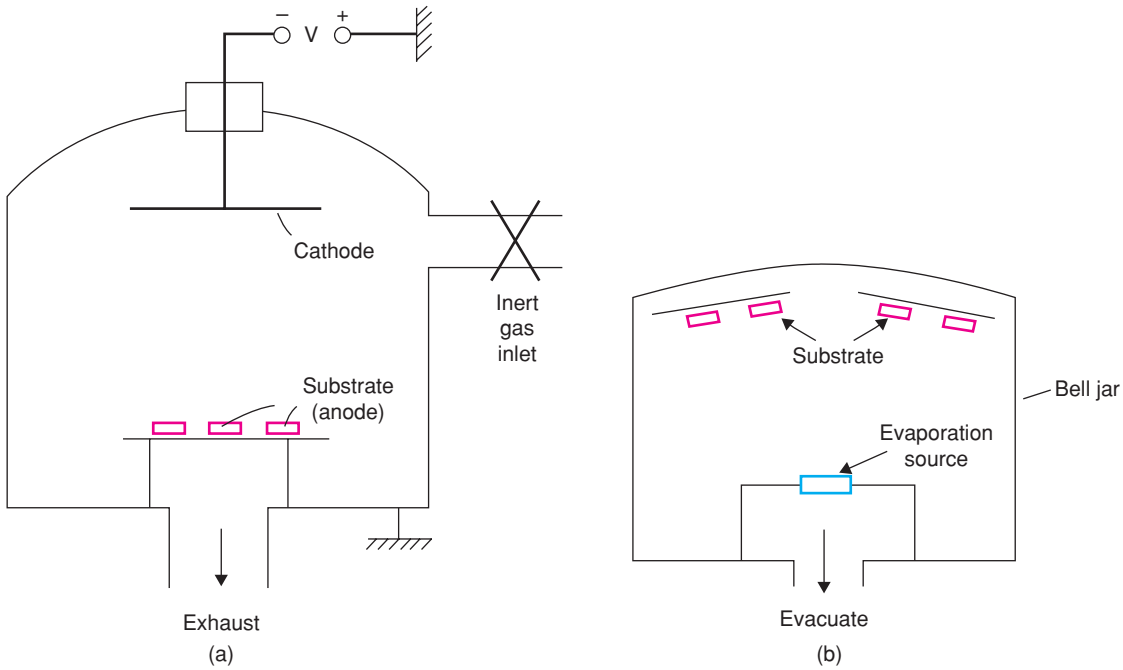


Figure 4.17 ***(a) Cathode sputtering equipment and (b) vacuum evaporation equipment for thin-film deposition.***

and inorganic salts. These are placed in containers made of, for example, W, Ta or Mo, which withstand very high temperatures and are chemically unreactive towards the material being evaporated.

It is often important to clean the substrate surface thoroughly prior to evaporation, and this may be carried out by a sequence of steps involving ultrasonic cleaning in a detergent solution, degreasing in alcohol, degassing under vacuum and, finally, ion bombardment to remove surface layers from the substrate. Such cleaning may be necessary in order to obtain good adhesion of the film to the substrate or simply for reasons of the purity required.

4.4.4 Atomic layer deposition, ALD

Atomic layer deposition, ALD (or atomic layer epitaxy, ALE) is used to prepare nanometre thick conformal films (i.e. films that follow the shape and contours of the substrate), particularly for microelectronics applications. Two self-limiting and complementary reactions are used sequentially to build up films slowly one monolayer at a time. This is illustrated in Fig. 4.18 for the deposition of films of TiO_2. The two chemical reagents are $TiCl_4$ and H_2O. The substrate is first exposed to $TiCl_4$, which forms a monolayer on the substrate surface; excess $TiCl_4$ is purged from the reaction chamber, which is then exposed to H_2O vapour. Ti–Cl bonds at the substrate surface are replaced by Ti–OH bonds. Excess H_2O is then purged and the cycle repeated as many times as required to achieve the desired film thickness.

The ALD process is therefore simple and, perhaps surprisingly, works well even though the detailed mechanistic understanding of the chemisorption and reaction processes are often not clear. Temperature control is important to achieve single-layer deposition in one cycle (hence self-limiting). Since film thickness depends only on the number of reaction cycles, it can be controlled accurately in a way that is not possible with other CVD processes.

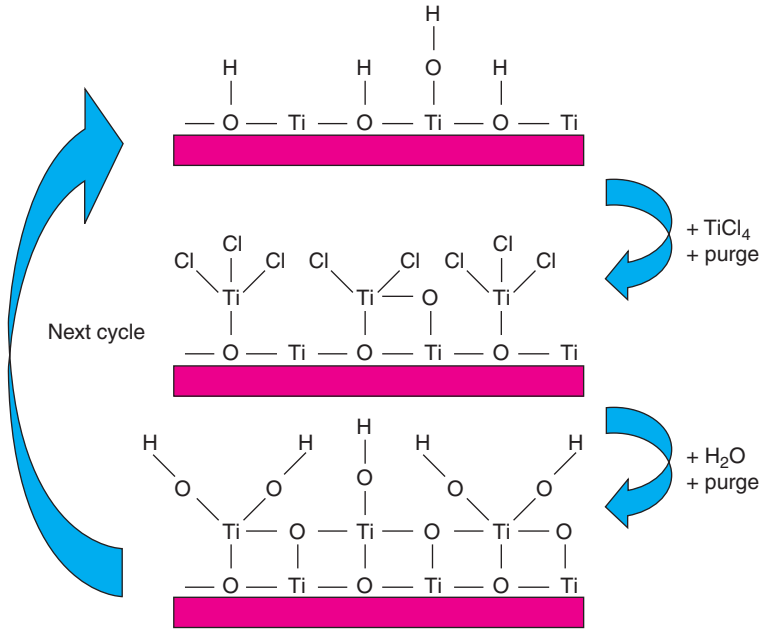


Figure 4.18 ***Deposition of TiO_2 films by ALD. Adapted with permission from M. Leskelä and M. Ritala,*** **Angew. Chem., Int. Ed.,** *42, 5548,* ***© 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.***

The main commercial application of ALD has been in the fabrication of thin-film electroluminescent flat-panel displays, TFELs. A white colour is achieved with alternating layers of yellow Mn^{2+}-doped ZnS and blue–green Ce^{3+}-doped SrS. It is also used to form coatings on high surface area porous supports of, e.g., γ-Al_2O_3 and SiO_2, for catalytic applications.

4.4.5 Aerosol synthesis and spray pyrolysis

An important method for the preparation of fine powders and thin-film deposition in industry is spray pyrolysis. The apparatus is simple and easily scaled up to mass production and the method can be a continuous rather than a batch process. The method generates aerosols, which are solid or liquid droplets suspended in a gas, in a nebuliser and are carried in a gas flow to a furnace where they are pyrolysed. One variation on this method is ultrasonic spray pyrolysis, USP, in which high-frequency ultrasound produces aerosols of micron-sized liquid droplets from a liquid precursor solution. The aerosol droplets effectively become chemical microreactors as they are carried into the furnace, Fig. 4.19. Evaporation of the solvent is followed by chemical reactions of the remaining solid or direct precipitation from the evaporating solution.

A wide range of nanomaterials have been prepared by USP, including nanoporous oxides, sulfides, high surface area carbons, semiconductor quantum dots, QDs, and conductive metal inks. QDs are nanoparticles whose properties may depend on their physical dimensions. For example, the band gap of CdSe QDs depends on the particle size, which is tuneable by varying the furnace temperature during USP. The band gap controls the optical absorption spectrum in the visible region of the electromagnetic spectrum and hence the colour, as shown in Fig. 4.20. QDs find applications as fluorescent labelling agents in living cells and may show a blue shift resulting from photo-oxidation and bleaching.

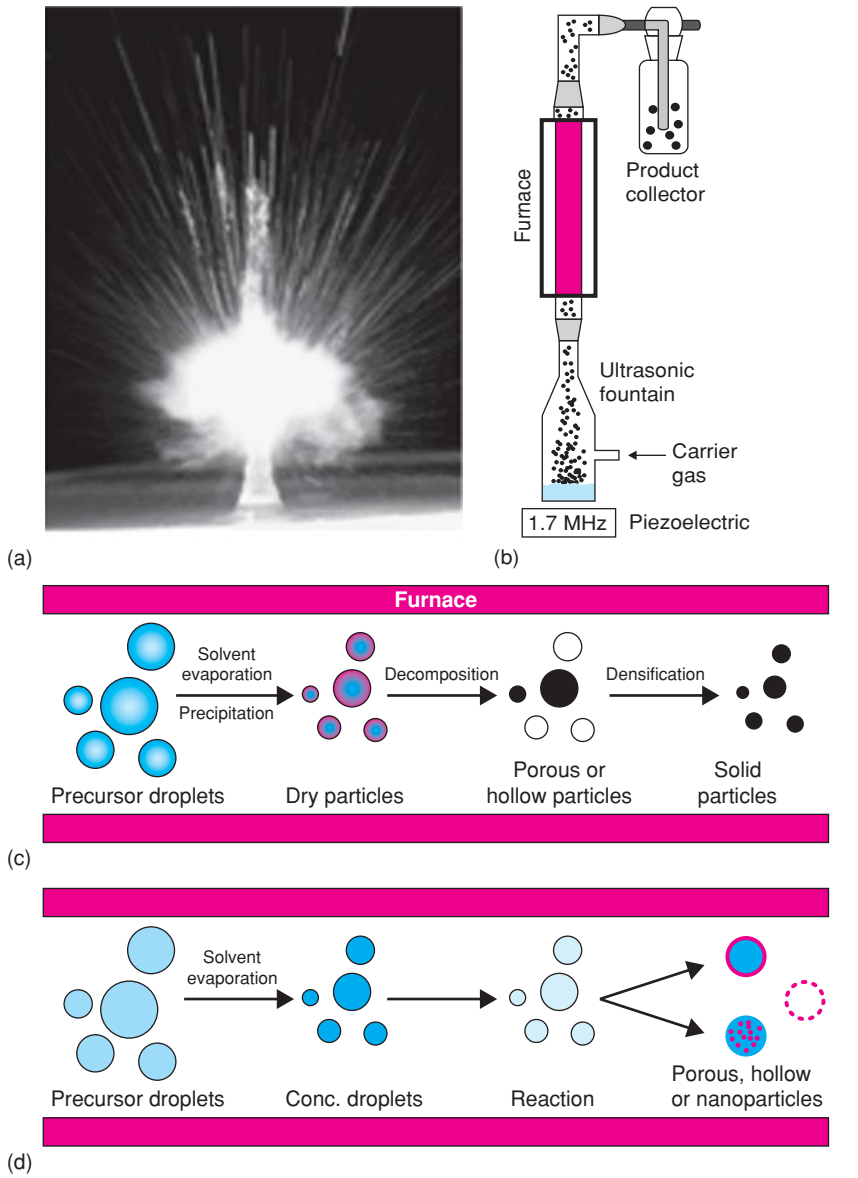


Figure 4.19 ***Ultrasonic spray pyrolysis: (a) photograph of an aerosol fountain produced using 1.7 MHz ultrasound; (b) a typical laboratory-scale USP set-up; (c) and (d) typical spray pyrolysis processes. Photograph courtesy of K. S. Suslick. Bang, J. H.; Didenko, Y. T.; Helmich, R. J.; Suslick, K. S. 'Nanostructured Materials through Ultrasonic Spray Pyrolysis'* Aldrich Materials Matter*, 2012, 7(2), 15–18.***

Spray pyrolysis is used in a number of industrial applications, especially for depositing films of active catalysts, sensors and phosphors on a range of substrates. This is illustrated in Fig. 4.21 for the fabrication of SnO_2 sensors, which find a variety of applications for detecting gases such as CO. Catalyst supports and photocatalysts are made from TiO_2 by spray pyrolysis for applications such as self-cleaning glass.

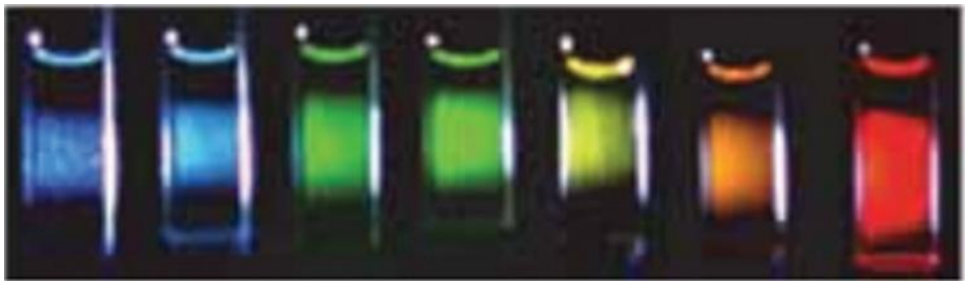


Figure 4.20 ***Fluorescence of CdSe quantum dots controlled by particle size which is tuneable from 2 to 4 nm by varying the furnace temperature during USP. Photograph courtesy of K. S. Suslick. Bang, J. H.; Didenko, Y. T.; Helmich, R. J.; Suslick, K. S. 'Nanostructured Materials through Ultrasonic Spray Pyrolysis'* Aldrich Materials Matter*, 2012, 7(2), 15–18.***

4.5 High-Pressure Methods

Techniques are now available for generating static pressures of several hundred kilobars at both ambient and high temperatures; using shock wave methods, the attainable P,T range may be extended still further. Experimental techniques vary from a simple 'opposed anvil' arrangement, whereby the sample is effectively squeezed between two pistons or rams, one of which is fixed and the other connected to a hydraulic jack, to more complex designs involving three or four anvils or rams, as in the *tetrahedral anvil*.

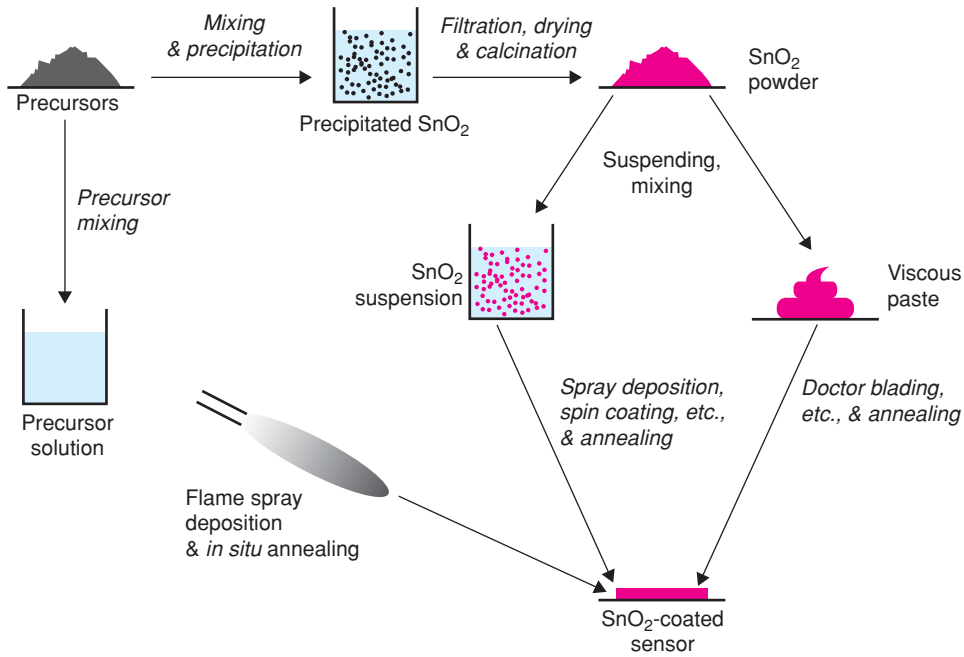


Figure 4.21 ***Process diagram for preparation of SnO_2 gas sensors comparing wet-chemical routes and direct flame deposition. Adapted with permission from R. Strobel and S. E. Pratsinis,*** J. Mater. Chem.***, © 2007 Royal Society of Chemistry.***

Table 4.5 ***High-pressure polymorphism of some simple solids***

Solid	Structure and coordination number	Typical transformation conditions		High-pressure structure and coordination number
		P/kbar	*T*/°C	
C	Graphite, 3	130	3000	Diamond, 4
CdS	Wurtzite, 4:4	30	20	Rock salt, 6:6
KC1	Rock salt, 6:6	20	20	CsCl, 8:8
SiO_2	Quartz, 4:2	120	1200	Rutile, 6:3
Li_2MoO_4	Phenacite, 4:4:3	10	400	Spinel, 6:4:4
$NaAlO_2$	Ordered Wurtzite, 4:4:4	40	400	Ordered Rock salt, 6:6:6

Phases synthesised at high pressures tend to have higher densities than phases synthesised at atmospheric pressure, and this sometimes gives rise to unusually high coordination numbers. For example, the coordination of Si in SiO_2 and the silicates is, with very few exceptions, tetrahedral. The high-pressure polymorph of SiO_2, stishovite, which is formed at pressures above 100–120 kbar, has the rutile structure and therefore contains octahedrally coordinated Si. Some other examples of increased coordination number in high-pressure polymorphs are given in Table 4.5.

By use of high pressures, it is also possible to stabilise ions in unusual oxidation states, such as Cr^{4+}, Cr^{5+}, Cu^{3+}, Ni^{3+} and Fe^{4+}. Cr normally occurs as Cr^{3+} and Cr^{6+} with octahedral and tetrahedral coordination, respectively. However, various perovskites containing octahedrally coordinated Cr^{4+} have been prepared at high pressures, such as $MCrO_3$: M = Ca, Sr, Ba, Pb. An industrially important high-pressure method is the growth of synthetic diamonds from graphite.

4.6 Crystal Growth

Crystals may be grown from vapour, liquid or solid phases although, usually, only the first two give crystals of sufficient size for applications or property measurements. Use of hydrothermal and vapour-phase transport methods of crystal growth were discussed in Sections 4.3.5 and 4.4.1. Other methods are as follows.

4.6.1 Czochralski method

A crystal is grown from a melt of the same composition by starting with a seed crystal in contact with (on the surface of) the melt, whose temperature is maintained slightly above its melting point. As the seed is gradually pulled out of the melt, Fig. 4.22, the melt solidifies on the surface of the seed to give a rod-shaped crystal in the same crystallographic orientation as the original seed. The melt and growing crystal are usually rotated counter-clockwise during pulling. The method is widely used for semiconducting materials, Si, Ge, GaAs, etc.; usually an inert gas atmosphere at high pressure is used in order to prevent loss of As, P, etc. It has also been used to produce laser generator materials such as Nd-doped $Ca(NbO_3)_2$.

4.6.2 Bridgman and Stockbarger methods

These methods are also based on the solidification of a stoichiometric melt but crystallisation is controlled by passing the melt through a temperature gradient such that crystallisation occurs at the cooler end. This

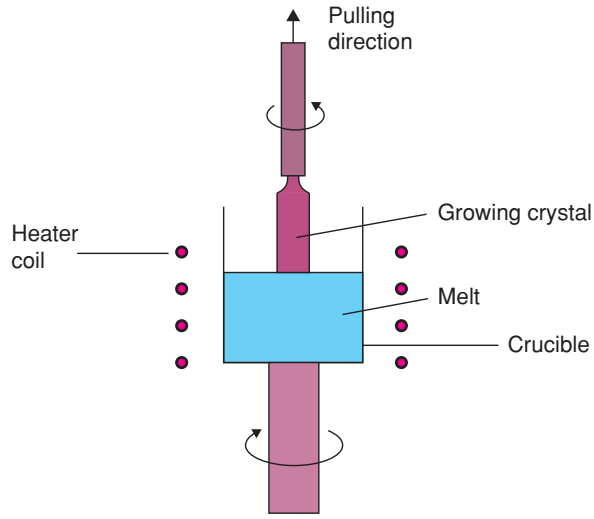


Figure 4.22 ***Czochralski method for crystal growth.***

is achieved in the Stockbarger method by arranging displacement of the melt within a temperature gradient, Fig. 4.23(a). In the Bridgman method, the melt is inside a temperature gradient furnace and the furnace is gradually cooled so that solidification begins at the cooler end, (b). In both methods, it is again advantageous to use a seed crystal and atmospheric control may be necessary.

4.6.3 Zone melting

This is related to the Stockbarger method but the thermal profile through the furnace is such that only a small part of the charge is molten at any one time, Fig. 4.23(c). Initially that part of the charge in contact with the seed crystal is melted. As the boat is pulled through the furnace, oriented solidification onto the seed occurs and, at the same time, more of the charge melts. This is a well-known method for purification of solids, the *zone-refining technique*, Fig. 7.27. It makes use of the principle that impurities usually concentrate in the liquid rather than in the solid phase. Impurities are therefore 'swept out' of the crystal by the moving molten zone. The method has been used for the purification and crystal growth of high melting metals such as tungsten.

4.6.4 Precipitation from solution or melt: flux method

In contrast to the above methods in which crystals have the same composition as the melt, precipitation methods involve the growth of crystals from a solvent of different composition. The solvent may be one of the constituents of the desired crystal, e.g. crystallisation of salt hydrate crystals from water, or the solvent may be an entirely separate liquid in which the crystals of interest are partially soluble, e.g. various high-melting silicates may be precipitated from low-melting borate or halide melts. In these cases, the solvent melts are sometimes referred to as *fluxes* since they effectively reduce the melting point of the crystals by a considerable amount.

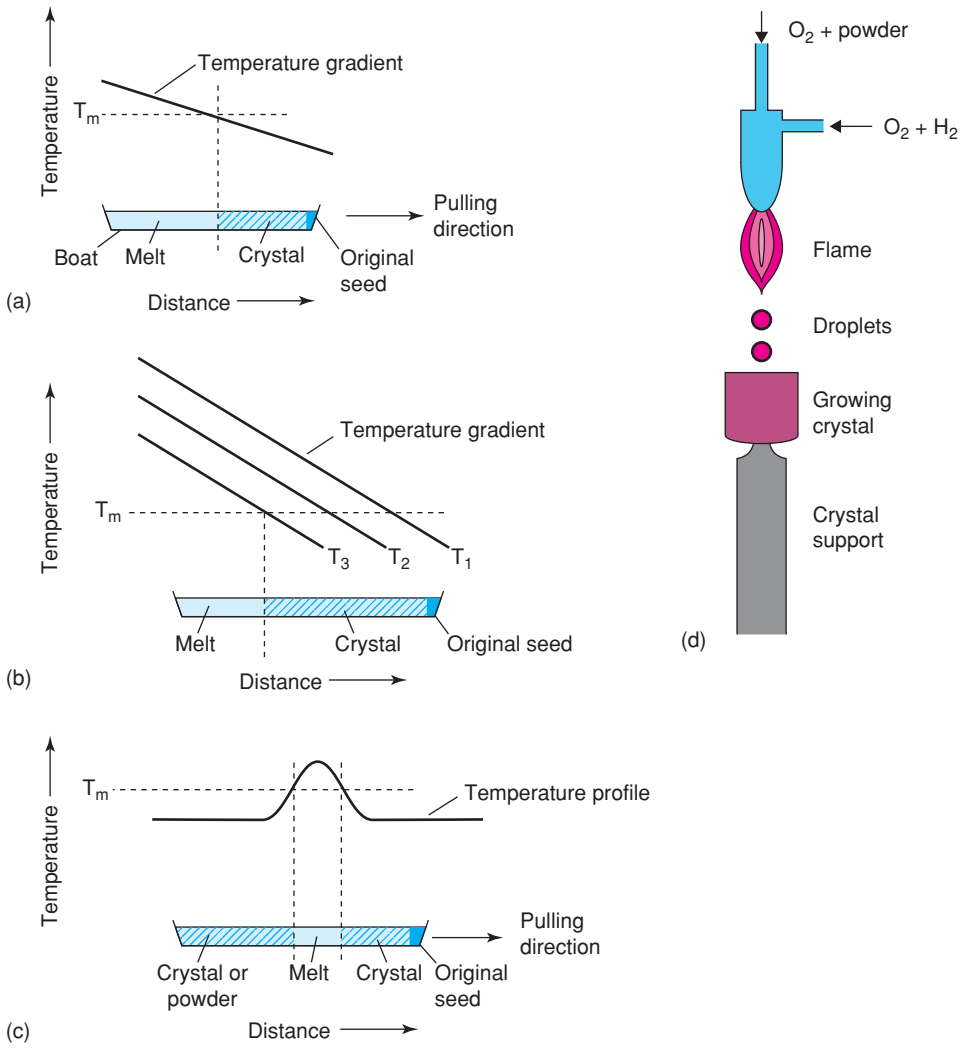


Figure 4.23 ***(a) Stockbarger method. T_m = crystal melting point. (b) Bridgman method. Times t_1, t_2 and t_3 are shown. (c) Zone melting method. (d) Verneuil method.***

4.6.5 Verneuil flame fusion method

This method was first used in 1904 to grow crystals of high-melting oxides, including artificial gemstones such as ruby and sapphire. The starting material, in the form of a fine powder, is passed through an oxyhydrogen flame or some other high-temperature torch or furnace, Fig. 4.23(d). After melting, the droplets fall on to the surface of the growing crystal or seed, where they solidify. The method has been used to prepare single crystals of CaO with a melting point of ~2600 °C by using a plasma torch to melt the CaO powder.

5

Crystallography and Diffraction Techniques

5.1 General Comments: Molecular and Non-Molecular Solids

5.1.1 Identification of crystalline solids

The simplest and most obvious first question to ask about an inorganic substance is: 'What is it?' The methods that are used to answer this come into two main categories, depending on whether the substance is molecular or non-molecular. If the substance is molecular, whether it be solid, liquid or gaseous, identification is usually carried out by some combination of spectroscopic methods and chemical analysis. Thus, mass spectrometry usually gives information on the formula weight of the molecules and that of resulting molecular fragments when the molecules disintegrate inside the spectrometer; IR spectroscopy gives information on the functional groups present in the molecule: carboxylic acids, ketones, alcohols, etc. From NMR spectroscopy, it is usually possible to deduce how the molecules are pieced together.

If the substance is non-molecular and crystalline, identification is usually carried out by powder X-ray diffraction, XRD, supplemented, where necessary, by chemical analysis. Each crystalline solid has its own characteristic XRD pattern which may be used as a 'fingerprint' for its identification. The powder patterns of most known inorganic solids are included in the *Powder Diffraction File* (Section 5.3.6); by using an appropriate search procedure, unknowns can usually be identified rapidly and unambiguously.

Once the substance has been identified, the next stage may be to determine its structure, if this is not known already. For molecular materials, details of the molecular geometry may be obtained from spectroscopic measurements. Alternatively, if the substance is crystalline, X-ray crystallography may be used, in which case information is also obtained on the way in which molecules pack together in the crystalline state. For molecular substances, this usually completes the story as far as identification and structure determination are concerned; attention may then be focused on other matters such as properties or chemical reactivity.

5.1.2 Structure of non-molecular crystalline solids

For non-molecular substances, however, the word 'structure' takes on a whole new meaning, as illustrated in Fig. 5.1. Obviously, we need to know the *crystal structure*, as given by the unit cell and its contents. However, defects and impurities are also often extremely important and sometimes control properties. Thus, the colour and lasing action of ruby, Cr-doped Al_2O_3, depend exclusively on the presence of Cr^{3+} impurities in the

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

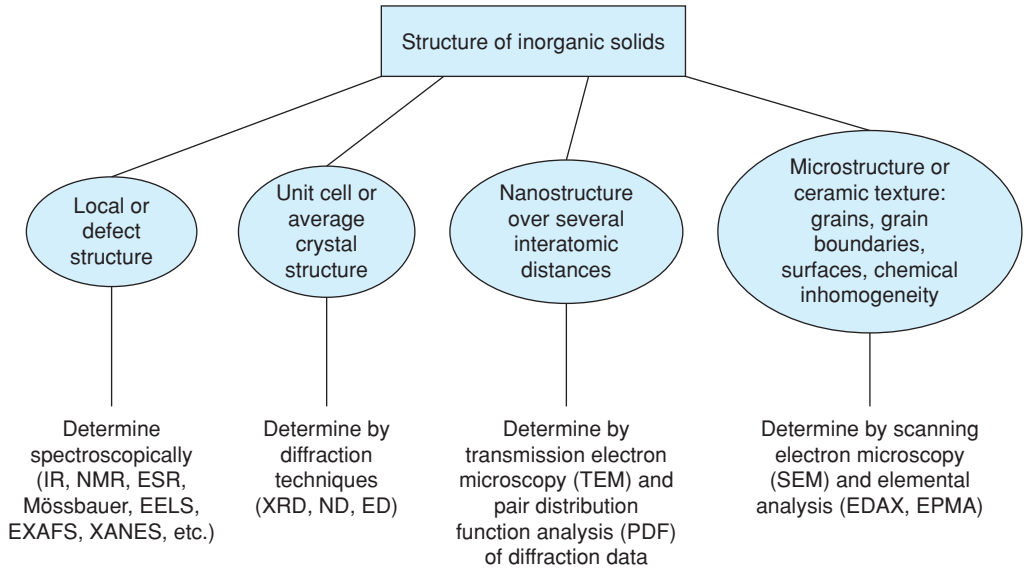


Figure 5.1 ***Structural features of inorganic solids across the length scales and some of the techniques used to study them.***

corundum crystal structure of Al_2O_3. In such cases, the crystal structure or *average structure* of the host is important, but the *local structure* centred on the impurities or defects controls the properties.

On a somewhat larger length scale, the optical properties of colloids (or nanoparticles, to give them a name that is currently more fashionable) depend on crystallite size. For instance, the colour, band gap and photoconductivity of CdS particles depend on their grain size and, hence, on their *nanostructure*.

On a still larger scale, the mechanical and electrical properties of ceramics are often determined by their *microstructure*, which covers the size, shape and distribution of crystalline grains, the bonding between grains and the segregation of any impurities to the surfaces or intergranular regions. To give one example, ZnO ceramic varistors are rather special materials whose electrical properties do not obey Ohm's law, and this is associated with compositional inhomogeneities in the ceramic, in particular the segregation of dopants such as Bi and Co to the intergranular regions.

When we refer to the structure of non-molecular inorganic materials, we may be interested in some or all of the above aspects of structure, ranging from local structure, over distances of a few ångströms (1 Å = 0.1 nm = 10^{-4} μm = 10^{-7} mm = 10^{-10} m), to microstructure at the micron level. This contrasts hugely with molecular substances for which 'structure' refers mainly to the atomic-level arrangement of atoms and, occasionally, especially with pharmaceuticals, to the packing arrangement of molecules in crystals. Since in non-molecular materials we are interested in the structure at several levels, a wide range of techniques is needed to *characterise* the solids.

5.1.3 Defects, impurities and stoichiometry of crystalline solids

The prime reason for many of the great differences between molecular and non-molecular molecular materials lies in the very different status of defects and impurities in the two categories of material. In molecular substances, defects are not allowed! If a certain molecule has atoms missing, or extra atoms present, then the resulting molecule is quite different to the parent molecule and can be separated by standard purification

Table 5.1 ***Comparison of a molecular substance, toluene, $C_6H_5CH_3$, and a non-molecular substance, alumina, Al_2O_3***

Feature	Toluene	Alumina
Stoichiometry	Fixed, $C_6H_5CH_3$	Fixed, Al_2O_3
Impurities?	Not allowed	Readily doped
Nature of the material	Volatile liquid	Can be powder, fibre, ceramic, single crystal, film
Properties		Depend on the nature of the material and dopants/impurities
Applications	Solvent	Abrasive (powder) Thermal insulation (Saffil fibres) Electrical insulation (thin film or ceramic substrate) Ruby gem and laser (Cr doped) Solid electrolyte (Na β-alumina)

methods. Further, the presence of any such 'defective' molecules is most unlikely to modify the properties of the parent, non-defective molecules. Thus, molecules have accurately fixed formulae or stoichiometries and are defect free.

In non-molecular materials, by contrast, defects and impurities are almost unavoidable. If they form an intrinsic part of the crystal structure, they cannot be readily eliminated and are always present from thermodynamic considerations. Such impurities give rise to *non-stoichiometry*, i.e. variable composition, and may induce dramatic changes in the properties of the parent structure.

An illustration of the vastly different chemistries of molecular and non-molecular substances is presented in Table 5.1 for two apparently simple substances, one in each category. Molecular chemists will tell you that toluene is an extremely well understood molecule and has few surprises left. Aluminium oxide, by contrast, shows a rich diversity of structures, properties and applications and is still being actively researched.

5.2 Characterisation of Solids

In order for a solid to be well characterised, Fig. 5.1, one needs to know about:

(a) the crystal structure, as given by the unit cell, its dimensions and the fractional coordinates of the atoms present in the cell;
(b) the crystal defects that are present, their nature, number and distribution;
(c) the impurities that are present and whether they are distributed at random or are concentrated into small regions;
(d) for nanostructured materials, the atomic arrangements over intermediate length scales, especially for amorphous solids or for crystalline solids whose structural details are modified at the nanoscale;
(e) for polycrystalline solids – powders or ceramics – the number, size, shape and distribution of the crystalline particles;
(f) the surface structure, including any compositional inhomogeneities, absorbed surface layers or structural differences between surface and interior.

No single technique is capable of providing complete characterisation of a solid. Rather, a variety of techniques are used in combination. The three main categories of physical techniques are diffraction, microscopy

and spectroscopy. Other techniques such as thermal analysis and physical property measurements give valuable information in certain cases. This chapter deals with diffraction techniques; other techniques are considered in Chapter 6.

X-ray diffraction has been used for over a century in two main areas: for the fingerprint identification of crystalline materials and for the determination of their structure. It is the principal technique of solid state chemistry and, accordingly, is given most space in this chapter. A brief description is also given of electron and neutron diffraction, two specialised techniques which have important applications.

5.3 X-Ray Diffraction

5.3.1 Generation of X-rays

5.3.1.1 Laboratory sources utilising inner shell electronic transitions

X-rays are electromagnetic radiation of wavelength ~1 Å (10^{-10} m). They occur in that part of the electromagnetic spectrum, Fig. 5.2, between γ-rays and the ultraviolet (UV) region. X-rays are produced when high-energy charged particles, e.g. electrons accelerated through 30 000 V, collide with matter. The resulting X-ray spectra usually have two components, a broad spectrum of wavelengths known as *white radiation* and a number of fixed, or monochromatic, wavelengths. White radiation arises when the electrons are slowed or stopped by the collision and some of their lost energy is converted into electromagnetic radiation. White radiation has wavelengths ranging upwards from a certain lower limiting value. This lower wavelength limit corresponds to the X-rays of highest energy and occurs when all the kinetic energy, $\frac{1}{2}mv^2$, of the incident electrons of mass m and velocity v is converted into X-rays of energy eV, where V is the accelerating voltage and e is the electronic charge. From the relation $E = hc/\lambda$, where E is energy, h is Planck's constant and c is the velocity of light *in vacuo*, and substituting for the values of the constants, the minimum wavelength λ_{min} may be calculated from λ_{min} (Å) = 12 398/V.

The X-rays which are used in almost all laboratory-based diffraction experiments are produced by a different process that leads to *monochromatic X-rays*. A beam of electrons, again accelerated through, say, 30 kV, is allowed to strike a metal *target*, often Cu. The incident electrons have sufficient energy to ionise some of the Cu 1s (K shell) electrons, Fig. 5.3(a). An electron in an outer orbital (2p or 3p) immediately drops down to occupy the vacant 1s level and the energy released in the transition appears as X-radiation. The transition energies have fixed values and so a spectrum of characteristic X-rays results, Fig. 5.3(b). For Cu, the $2p \rightarrow 1s$ transition, called $K\alpha$, has a wavelength of 1.5418 Å and the $3p \rightarrow 1s$ transition, $K\beta$, 1.3922 Å. The electronic transitions involved in both the ionisation of inner shell electrons and emission of X-rays are shown in Fig. 6.28(b).

The $K\alpha$ transition occurs much more frequently than the $K\beta$ transition and the more intense $K\alpha$ radiation that is produced is used in diffraction experiments. In fact, the $K\alpha$ radiation is a doublet, $K\alpha_1 = 1.54051$ Å and $K\alpha_2 = 1.54433$ Å, because the $2p \rightarrow 1s$ transition has a slightly different energy for the two possible

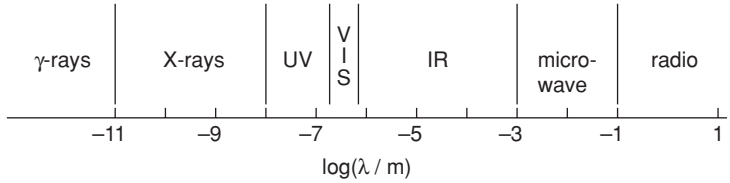


Figure 5.2 ***The electromagnetic spectrum.***

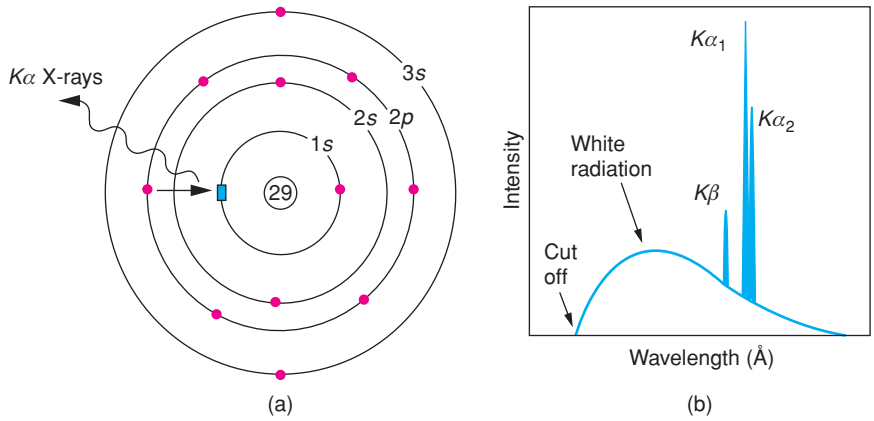


Figure 5.3 ***(a) Generation of Cu Kα X-rays. A 1s electron is ionised; a 2p electron falls into the empty 1s level (▮) and the excess energy is released as X-rays. (b) X-ray emission spectrum of Cu.***

spin states of the $2p$ electron which makes the transition, relative to the spin of the vacant $1s$ orbital. More specifically, the $2p$ electron has total angular momentum given by the quantum number $j = l \pm s$, where l is the orbital quantum number and s is the spin quantum number; for a $2p$ electron $l = 1$ and therefore $j = 3/2$ or 1/2 and the two possible transitions are

$$2p_{\frac{1}{2}} \xrightarrow{K\alpha_1} 1s_{\frac{1}{2}} \text{ and } 2p_{\frac{3}{2}} \xrightarrow{K\alpha_2} 1s_{\frac{1}{2}}$$

In some X-ray experiments, diffraction by the $K\alpha_1$ and $K\alpha_2$ *r*adiations is not resolved and a single line or spot is observed instead of a doublet (e.g. in powder diffractometry at low angle). In other experiments, separate diffraction peaks may be observed; if desired, this can be overcome by removing the weaker $K\alpha_2$ beam from the incident radiation.

The wavelengths of the $K\alpha$ lines of the target metals commonly used for X-ray generation are given in Table 5.2. They are related to the atomic number, Z, of the metal by *Moseley's law*:

$$\lambda^{-\frac{1}{2}} = C(Z - \sigma) \tag{5.1}$$

where C and σ are constants. Hence the wavelength of the $K\alpha$ line decreases with increasing atomic number.

Table 5.2 ***X-ray wavelengths (Å) of commonly used target materials***

Target	$K\alpha_1$	$K\alpha_2$	$K\bar{\alpha}$[a]	Filter
Cr	2.2896	2.2935	2.2909	V
Fe	1.9360	1.9399	1.9373	Mn
Cu	1.5405	1.5443	1.5418	Ni
Mo	0.7093	0.7135	0.7107	Nb
Ag	0.5594	0.5638	0.5608	Pd

[a] $\bar{\alpha}$ is the intensity-weighted average of α_1 and α_2.

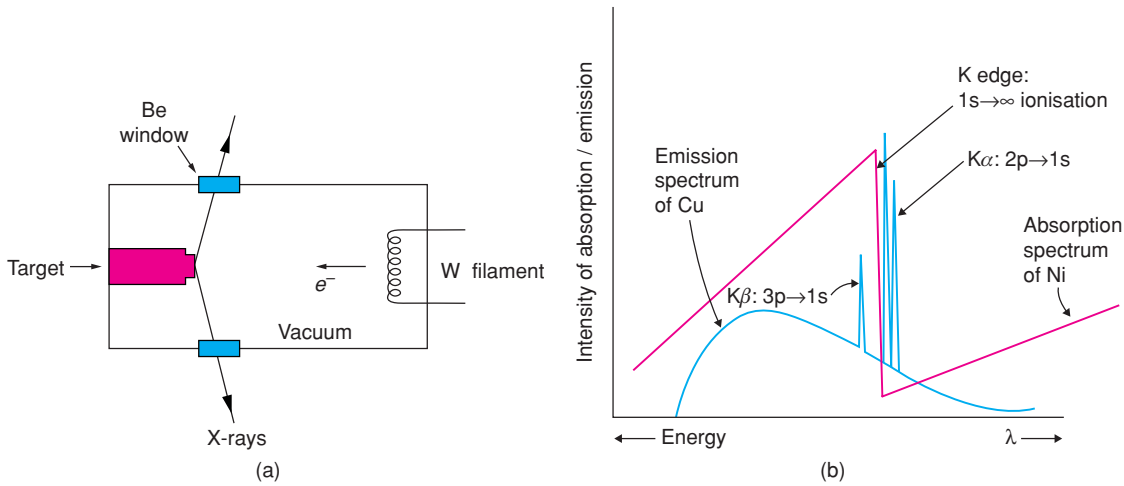


Figure 5.4 ***(a) Schematic design of a filament X-ray tube. (b) Use of Ni to filter out Cu Kβ and white radiation.***

The X-ray emission spectrum of an element such as Cu, Fig. 5.3(b), has two main features. The intense, monochromatic peaks, caused by electronic transitions within the atoms, have wavelengths that are characteristic of the element, i.e. Cu. These peaks are superposed on a background of white radiation produced by the general interaction of high-velocity electrons with matter. In order to generate the characteristic monochromatic radiation, the voltage used to accelerate the electrons needs to be sufficiently high ($\geq$10 kV) that ionisation of the Cu 1*s* electrons occurs, Fig. 6.28(b).

In the generation of X-rays, Fig. 5.4(a), the electron beam, provided by a heated tungsten filament, is accelerated towards an anode by a potential difference of $\sim$30 kV. The electrons strike the target, a piece of Cu fixed to the anode, and a spectrum of X-rays, such as shown in Fig. 5.3(b), is emitted. The chamber, known as the *X-ray tube*, is evacuated to prevent oxidation of the W filament. The X-rays leave the tube through 'windows' made of Be. The absorption of X-rays on passing through materials depends on the atomic weight of the elements present. Be with an atomic number of 4 is, therefore, one of the most suitable window materials. For the same reason, Pb is a very effective material for shielding X-ray equipment and absorbing stray radiation. While an X-ray tube is in operation, continuous cooling of the anode is necessary. Only a small fraction of the energy of the incident electron beam is converted into X-rays. Most of the energy is converted into heat and the anode would soon melt if it were not cooled. In *rotating anode* generators, the anode spins to allow distribution of a higher heat load over the target, allowing the generation of a higher intensity X-ray beam.

For most diffraction experiments, a monochromatic beam of X-rays is desired and not a continuous spectrum. In the spectrum of X-rays emitted by Cu (or any metal), the $K\alpha$ line(s) is the most intense and it is desired to filter out all other wavelengths, leaving the $K\alpha$ line for diffraction experiments. For Cu radiation, a sheet of Ni foil is a very effective filter. The energy required to ionise 1*s* electrons of Ni corresponds to a wavelength of 1.488 Å, which lies between the values for the $K\alpha$ and $K\beta$ lines of the Cu emission spectrum, as shown in Fig. 5.4 (b). Cu $K\beta$ radiation, therefore, has sufficient energy to ionise 1*s* electrons of Ni whereas Cu $K\alpha$ radiation does not. Ni foil is effective in absorbing the Cu $K\beta$ radiation and most of the white radiation, leaving a monochromatic, reasonably clean beam of $K\alpha$ radiation. A lighter element, such as Fe, would absorb Cu $K\alpha$ radiation in addition to $K\beta$, because its *absorption edge* is displaced to higher wavelengths. On the other hand, a heavier element, such as Zn, would transmit both $K\alpha$ and $K\beta$ radiation while still absorbing much of the higher energy white radiation. The atomic number of the element in the filter generally is one

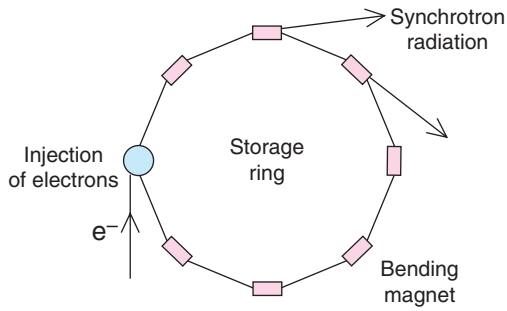


Figure 5.5 ***Schematic diagram of a synchrotron storage ring.***

or two less than that of the target material, Table 5.2. An alternative method of obtaining monochromatic X-rays, using a single-crystal monochromator, is discussed later.

5.3.1.2 Synchrotron X-ray sources

Synchrotron radiation is emitted when charged particles such as electrons, travelling at relativistic speeds, are forced to change direction on passing through a magnetic field. To generate synchrotron radiation, electrons or positrons are accelerated to speeds close to that of light and circulate in ultra-high vacuum tubes, or *storage rings*, guided by arrays of magnets. Synchrotron sources are large-scale national facilities such as the European Synchrotron Radiation Facility, ESRF, at Grenoble, France, the National Synchrotron Light Source, NSLS, at Brookhaven, Upton, NY, USA, and the Photon Factory, at Tsukuba, Japan.

A simplified sketch of a storage ring, which is typically hundreds of metres in diameter, is shown in Fig. 5.5 and contains both linear and curved segments. Electrons steered through the curved segments by a bending magnet emit streams of synchrotron radiation tangentially to the electron path and form the beam lines for various instruments. In addition, the linear sections have insertion devices known as wigglers and undulators, which also have their magnetic fields perpendicular to the storage ring, cause deflection of the electrons in the horizontal plane and emit radiation which is linearly polarised with the electric component lying in the plane of the ring.

Many different diffraction and spectroscopic instruments are arranged around the circumference of a storage ring. The spectrum of synchrotron radiation in each beam line can be varied depending on the energy of the electrons in the storage ring, the curvature of their path and, for an undulator, interference effects arising from radiation associated with different oscillations. The beams can be continuous (bending magnets and wigglers) or pulsed (undulators) with a series of peaks at integral multiples of a fundamental wavelength, which depend on the strength of the magnetic field. Synchrotron radiation has several advantages compared with conventional X-ray sources. Synchrotron radiation is of very high intensity, and the beams are accurately parallel, giving high resolution in powder diffraction data, are continuous over the complete spectral range from radio frequencies to γ-rays and are tuneable, allowing energy-dispersive experiments (not discussed here). For diffraction experiments, a monochromatic beam of X-rays is usually required and this is separated from the continuous spectrum using crystal monochromators (see later).

5.3.2 Interaction of X-rays with matter

X-rays can interact with matter in two general ways: by *scattering* and by *absorption*. If there is no loss of energy on scattering, the scattered X-rays are *coherent* with the incident X-rays and these are used for X-ray

diffraction experiments, which is the main subject of this chapter. Scattered X-rays may also lose some of their energy, in which case they are *incoherent* and of longer wavelength than the incident X-rays. Incoherent scattering, also known as Compton scattering, is generally a nuisance and contributes to the background scattering in diffraction patterns.

X-rays may also be absorbed by matter, leading either to ionisation of electrons or to promotion of electrons to higher, unoccupied energy levels. The latter process generates atoms in excited states, which can then return to their ground state accompanied by emission of radiation of appropriate energy. This energy may also be in the X-ray region when the electrons are ionised from inner shell levels; the process is known generally as *fluorescence*. Characteristic emission spectra, as shown in Fig. 5.3(b), are also examples of fluorescence produced by the same process as used in the generation of X-rays. Fluorescence resulting from interaction of matter with X-rays is therefore a secondary X-ray generation process and can be a nuisance, leading to additional background scattering in diffraction patterns.

Under other circumstances, however, fluorescence can be of great value since the resulting X-rays have wavelengths characteristic of the atom from which they are emitted and form the basis of several analytical techniques, discussed further in Chapter 6. Thus, *X-ray fluorescence* is a major analytical technique to determine quantitatively the concentrations of atoms that are present in a solid. Similar fluorescence processes lie behind the electron microscope-related techniques of *electron probe microanalysis*, EPMA, and *energy-dispersive analysis of X-rays*, EDAX or EDX.

In addition to using fluorescent X-rays for diffraction and analytical purposes, the absorption process also forms the basis of important analytical and characterisation techniques (see Chapter 6). In the region of X-ray absorption edges, Fig. 5.4(b), the precise shape and wavelength of the absorption edge are very sensitive to the local environment of the absorbing atoms and can be used, for instance, to determine oxidation states of elements in *X-ray absorption near-edge structure*, XANES, or to determine local structure such as coordination numbers and bond lengths in *extended X-ray absorption fine structure*, EXAFS. The various scattering, absorption and emission processes occur simultaneously on interaction of X-rays with matter; in practice, the experimental procedures are optimised to focus on the technique that is of particular interest.

5.3.3 Optical grating and diffraction of light

As an aid to understanding the diffraction of X-rays by crystals, let us consider the diffraction of light by an optical grating. This gives a one-dimensional (1D) analogue of the 3D process that occurs in crystals. An optical grating is a piece of glass on which have been ruled a large number of accurately parallel and closely spaced lines. The separation of the lines is slightly larger than the wavelength of light, say 10 000 Å. The grating is shown in projection as a row of points in Fig. 5.6(a). Consider what happens to a beam of light which hits the grating perpendicular to the plane of the grating. A piece of glass without the lines would simply transmit the light, but in the grating the lines act as secondary point (or, rather, line) sources of light and re-radiate light in all directions. Interference then occurs between the waves originating from each line source.

In certain directions, adjacent beams are in-phase with each other and *constructive interference* occurs to give a resultant diffracted beam in that direction. Two such directions are shown in Fig. 5.6(b). In direction 1, parallel to the incident beam, the diffracted beams are obviously in-phase. In direction 2, the beams are also in-phase although beam B is now exactly one wavelength behind beam A. For directions between 1 and 2, B lags A by a fraction of one wavelength, and *destructive interference* occurs. For a certain direction, 3, B is exactly half a wavelength behind A and complete destructive interference or *cancellation* occurs. For other directions between 1 and 2, the destructive interference is only partial. Thus, directions 1 and 2 have maximum light intensity and this falls off gradually to zero as the angle changes to direction 3. In the optical grating, however, there are not just two parallel diffracted beams A and B but several hundred, one for each line on the grating. This causes the resultant diffracted beams to sharpen enormously after interference so

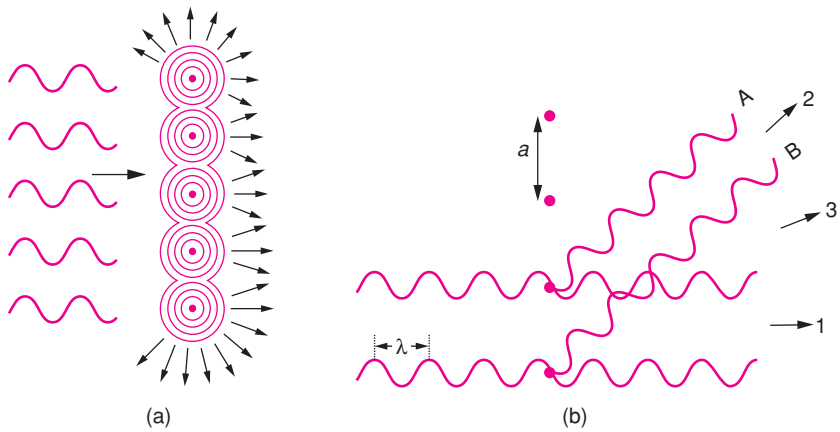


Figure 5.6 ***(a) Lines on an optical grating act as secondary sources of light. (b) Constructive interference in directions 1 and 2.***

that intense beams occur in directions 1 and 2 with virtually no intensity over the whole range of directions between 1 and 2.

The directions, such as 2, in which constructive interference occurs are governed by the wavelength of the light, λ, and the separation, a, of the lines on the grating. Consider the diffracted beams 1 and 2, Fig. 5.7, which are at an angle ϕ to the direction of the incident beam. If 1 and 2 are in-phase, the distance AB must equal a whole number of wavelengths, i.e.

$$\mathrm{AB} = \lambda,\ 2\lambda,\ \ldots,\ n\lambda$$

However,

$$\mathrm{AB} = a \sin \phi$$

and therefore

$$a \sin \phi = n\lambda \tag{5.2}$$

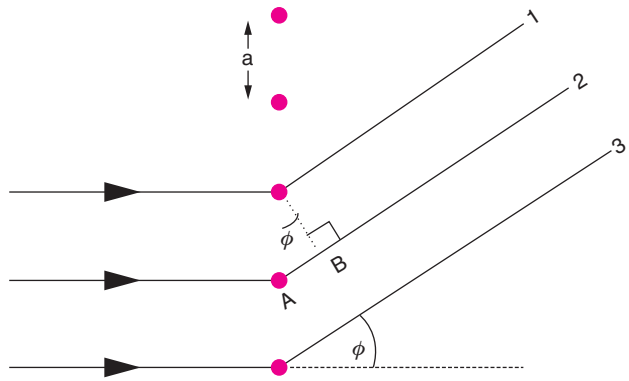


Figure 5.7 ***Diffraction of light by an optical grating.***

This equation gives the conditions under which constructive interference occurs and relates the spacing of the grating to the light wavelength and the diffraction order, n. Hence, depending on the value of $a\sin\phi$, one or more *diffraction orders*, corresponding to $n = 1, 2$, etc., may be observed.

We can now understand why the separation of lines on the grating must be of the same order of magnitude as, but somewhat larger than, the wavelength of light. The condition for the first-order (i.e. $n = 1$) diffracted beam to occur is $a\sin\phi = \lambda$. The maximum value of sin φ is 1, corresponding to $\varphi = 90^\circ$ but realistically, in order to observe first-order diffraction, sin $\varphi < 1$ and, therefore, $a > \lambda$. If $a < \lambda$, only the zero order direct beam is observable.

If, on the other hand, $a \gg \lambda$, individual diffraction orders ($n = 1, 2, 3, \ldots$, etc.) are so close together as to be unresolved and, effectively, a diffraction continuum results. This is because, for large values of a, $\sin\phi$ and, hence, ϕ must be very small. Therefore, $\phi_{n=1} \approx 0$ and the first-order beam is not distinguishable from the primary beam. Visible light has wavelengths in the range 4000–7000 Å and so, in order to observe well-separated spectra, grating spacings are usually 10 000–20 000 Å.

The other condition to be observed in the construction of an optical grating is that the lines should be accurately parallel. If this were not so, ϕ would vary over the grating and the diffraction spectra would be blurred or irregular and of poor quality generally.

5.3.4 Crystals and diffraction of X-rays

By analogy with the diffraction of light by an optical grating, crystals, with their regularly repeating structures, should be capable of diffracting radiation that has a wavelength similar to interatomic separations, ~2–3 Å. Three types of radiation are used for crystal diffraction studies: X-rays, electrons and neutrons. Of these, X-rays are the most useful, but electron and neutron diffraction also have important specific applications and are discussed later. The X-ray wavelength commonly employed is the characteristic $K\alpha$ radiation, $\lambda = 1.5418$ Å, emitted by Cu. When crystals diffract X-rays, the atoms or ions act as secondary point sources and scatter the X-rays; in the optical grating, the lines scratched on the glass surface cause scattering.

Historically, two approaches have been used to treat diffraction by crystals: the Laue equations and Bragg's law.

5.3.4.1 The Laue equations

Diffraction from a hypothetical 1D crystal, constituting a row of atoms, may be treated in the same way as diffraction of light by an optical grating because, in projection, the grating is a row of points. Equation (5.2) is obtained which relates the separation, a, of the atoms in the row, the X-ray wavelength, λ, and the diffraction angle, ϕ. A real crystal is a 3D arrangement of atoms for which three *Laue equations* may be written as follows:

$$a_1 \sin\phi_1 = n\lambda$$

$$a_2 \sin\phi_2 = n\lambda$$

$$a_3 \sin\phi_3 = n\lambda$$

These equations correspond to each of the three crystallographic axes needed to represent the atomic arrangement in the crystal. For a diffracted beam to occur, these three equations must be satisfied simultaneously.

The Laue equations provide a rigorous and mathematically correct way to describe diffraction by crystals. The drawback is that they are cumbersome to use. The alternative theory of diffraction, based on Bragg's law,

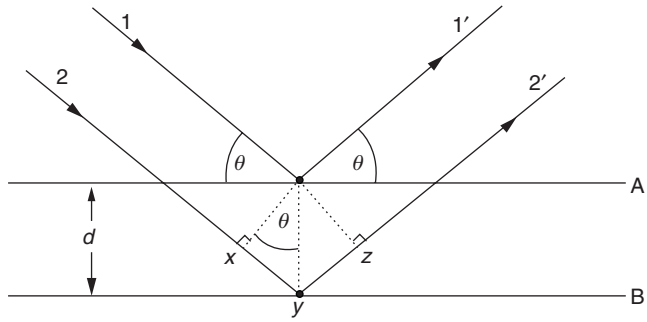


Figure 5.8 ***Derivation of Bragg's law.***

is much simpler and is used almost universally in solid state chemistry. The Laue equations are not considered further here.

5.3.4.2 Bragg's law

The Bragg approach to diffraction is to regard crystals as built up in layers or planes such that each acts as a semi-transparent mirror. Some of the X-rays are reflected off a plane with an angle of reflection equal to the angle of incidence, but the rest are transmitted to be subsequently reflected by succeeding planes.

The derivation of Bragg's law is shown in Fig. 5.8. Two X-ray beams, 1 and 2, are reflected from adjacent planes, A and B, within the crystal and we wish to know under what conditions the reflected beams $1'$ and $2'$ are in-phase. Beam $22'$ has to travel an extra distance xyz compared with beam $11'$; for $1'$ and $2'$ to be in-phase, the distance xyz must equal a whole number of wavelengths. The perpendicular distance between pairs of adjacent planes, the *d-spacing*, d, and the angle of incidence, or *Bragg angle*, θ, are related to the distance xy by

$$xy = yz = d \sin\theta$$

Hence

$$xyz = 2d \sin\theta$$

However,

$$xyz = n\lambda$$

and therefore

$$2d\sin\theta = n/\lambda\text{: Bragg}'\text{s law} \tag{5.3}$$

When Bragg's law is satisfied, the reflected beams are in-phase and interfere constructively. At angles of incidence other than the Bragg angle, reflected beams are out-of-phase and destructive interference or cancellation occurs. In real crystals, which contain thousands of planes and not just the two shown in Fig. 5.8, Bragg's law imposes a stringent condition on the angles at which reflection may occur. If the incident angle is incorrect by more than a few tenths of a degree, cancellation of the reflected beams is usually complete.

For a given set of planes, several solutions of Bragg's law are usually possible, for $n = 1, 2, 3$, etc. It is customary, however, to set n equal to 1 and for situations where, say, $n = 2$, the d-spacing is instead halved by doubling the number of planes in the set; hence n is kept equal to 1. [Note that $2\lambda = 2d\sin\theta$ is equivalent to $\lambda = 2(d/2)\sin\theta$.]

It is difficult to give an explanation of the nature of the semi-transparent layers or planes that is immediately convincing, as they are a concept rather than a physical reality. Crystal structures, with their regularly repeating patterns, may be referred to a 3D grid and the repeating unit of the grid, the *unit cell*, can be found. The grid may be divided up into sets of planes in various orientations and it is these planes which are considered in the derivation of Bragg's law. In some cases, with simple crystal structures, the planes also correspond to layers of atoms, but this is not generally the case. See Section 1.5 for further information.

Some of the assumptions upon which Bragg's law is based may seem to be rather dubious. For instance, it is known that diffraction occurs as a result of interaction between X-rays and atoms. Further, the atoms do not *reflect* X-rays but scatter or diffract them in all directions. Nevertheless, the highly simplified treatment that is used in deriving Bragg's law gives exactly the same answers as are obtained by a rigorous mathematical treatment. We therefore happily use terms such as reflexion (often deliberately with this alternative, but incorrect, spelling!) and bear in mind that we are fortunate to have such a simple and picturesque, albeit inaccurate, way to describe what in reality is a very complicated process.

5.3.5 X-ray diffraction methods

When reduced to basic essentials, the X-ray diffraction experiment, Fig. 5.9, requires an X-ray source, the sample under investigation and a detector to pick up the diffracted X-rays. Within this broad framework, three variables govern the different X-ray techniques:

(a) radiation – monochromatic or of variable λ;
(b) sample – single crystal, powder or a solid piece;
(c) detector – radiation counter or photographic film.

These are summarised for the most important techniques in Fig. 5.10. With the exception of the Laue method, which is used almost exclusively by metallurgists to study metal textures and is not discussed here, monochromatic radiation is nearly always used.

5.3.6 The powder method – principles and uses

The principles of the powder method are shown in Fig. 5.11. A monochromatic beam of X-rays strikes a finely powdered sample that, ideally, has crystals randomly arranged in every possible orientation. In such

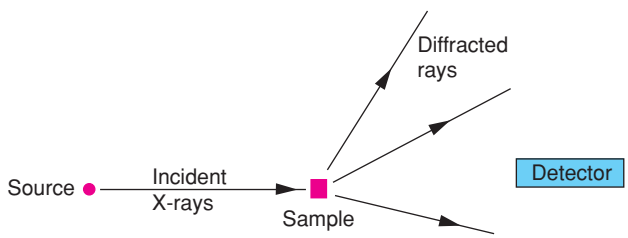


Figure 5.9 ***The X-ray diffraction experiment.***

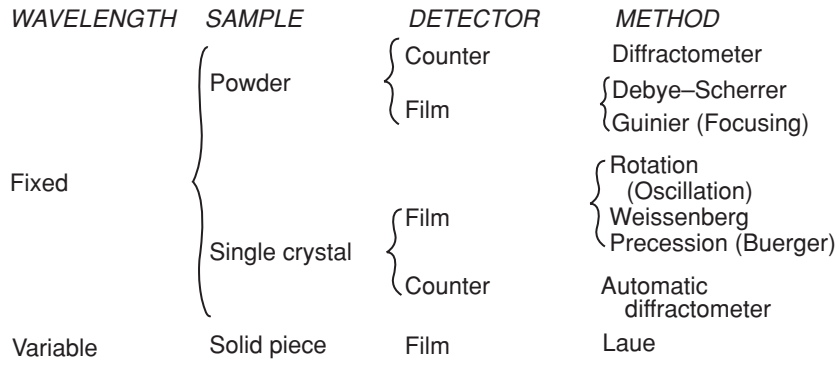


Figure 5.10 ***The different X-ray diffraction techniques.***

a powder sample, the various lattice planes are also present in every possible orientation. For each set of planes, therefore, at least some crystals must be oriented at the Bragg angle, θ, to the incident beam and thus diffraction occurs for these crystals and planes. The diffracted beams may be detected either by surrounding the sample with a strip of photographic film (Debye–Scherrer and Guinier focusing methods) or by using a movable detector, such as a Geiger counter, scintillation counter or image plate, connected to a chart recorder or computer (diffractometer).

The original powder method, the *Debye–Scherrer method*, is little used nowadays, but more modern methods operate on the same principle, it is easy to understand and it provides a very good introduction to the principles of X-ray diffraction. For any set of lattice planes, the diffracted radiation forms the surface of a cone, as shown in Fig. 5.12. The only requirement for diffraction is that the planes be at the Bragg angle θ to the incident beam; no restriction is placed on the angular orientation of the planes about the axis of the incident beam. In a finely powdered sample, crystals are present at every possible angular position about the incident beam and the diffracted beams that result appear to be emitted from the sample as cones of radiation (each cone is in fact a large number of closely spaced diffracted beams). If the Bragg angle is θ, then the angle between diffracted and undiffracted beams is 2θ and the angle of the cone is 4θ. Each set of planes gives its own cone of radiation. The cones are detected by a thin strip of film wrapped around the sample, Fig. 5.11; each cone intersects the film as two short arcs, Fig. 5.13, which are symmetrical about the two holes in the film (these allow entry and exit of incident and undiffracted beams). In a well-powdered sample, each arc appears as a continuous line, but in coarser samples the arcs may be spotty due to the relatively small number of crystals present.

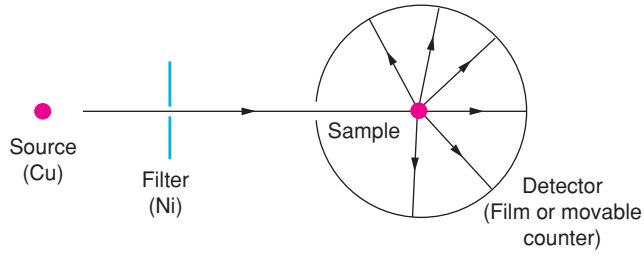


Figure 5.11 ***The powder method.***

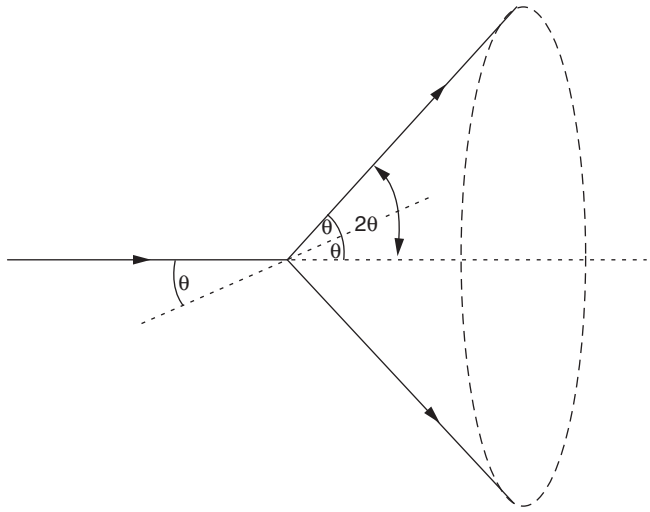


Figure 5.12 ***Formation of a cone of diffracted radiation.***

To obtain *d*-spacings from the Debye–Scherrer film, the separation, *S*, between pairs of corresponding arcs is measured. If the camera (film) radius, *R*, is known, then

$$\frac{S}{2\pi R} = \frac{4\theta}{360} \tag{5.4}$$

from which 2θ and therefore *d* may be obtained for each pair of arcs. The disadvantages of this method are that exposure times are long (6–24 h) and closely spaced arcs are not well resolved. This is because, although the incident beam enters the camera through a pinhole slit and collimator tube, the beam is somewhat divergent and the spread increases in the diffracted beams. If, in an effort to increase the resolution, a finer collimator is used, the resulting diffracted beams have much less intensity and longer exposure times are needed. Apart from considerations of the extra time involved, the amount of background radiation detected by the film (as fogging) increases with exposure time and, consequently, weak lines may be lost altogether in the background.

In modern film methods (*Guinier focusing methods*), a *convergent*, intense incident beam is used, with the result that excellent resolution of lines is obtained and exposure times are much reduced (10 min–1 h). Methods for obtaining a convergent beam of X-rays are discussed in the next section.

The other modern and more widely-used powder technique is *diffractometry*, which gives a series of peaks on a strip of chart paper or on a computer screen. A convergent incident beam is again used to give good resolution of peaks. Both peak positions and intensities (peak heights or areas) are readily obtained, which makes this a very useful and rapid method of phase analysis.

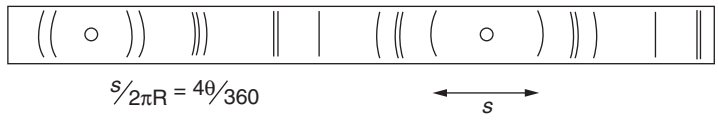


Figure 5.13 ***Schematic Debye–Scherrer photograph.***

The most important use of the powder method is in the qualitative identification of crystalline phases or compounds. Whereas most chemical methods of analysis give information about the *elements* present in a sample, powder XRD is very different, and perhaps unique, in that it tells which *crystalline compounds* or *phases* are present but gives no direct information about their chemical constitution.

Each crystalline phase has a characteristic powder XRD pattern which can be used as a fingerprint for identification purposes. The two variables in a powder pattern are peak position, i.e. d-spacing, which can be measured very accurately if necessary, and intensity, which can be measured either qualitatively or quantitatively. It is rare but not unknown that different crystalline phases have identical XRD patterns. More often, different materials may have one or two lines with common d-spacings, but on comparing the whole patterns, which may contain between $\sim$5 and $>$100 observed lines, the two are found to be different. The normal practice in using XRD patterns for identification purposes is to pay most attention to the d-spacings but, at the same time, check that the intensities are roughly correct.

For the identification of unknown crystalline materials, an invaluable reference source is the *Powder Diffraction File* (International Centre for Diffraction Data, USA), previously known as the ASTM or JCPDS file, which contains the XRD patterns of $>$300 000 materials; new entries are added at a current rate of $\sim$2000 p.a. In the search indices, materials are classified according to either their most intense peaks or the first eight lines in the XRD pattern in order of decreasing d-spacing. Identification of an unknown is usually possible within a few minutes of obtaining its XRD pattern.

Although the powder XRD method is invaluable for the identification of crystalline materials, it is important to recognise that it does have limitations, as follows:

1. If the material is not included in the powder diffraction file (obviously!).
2. If the material is not pure but contains lines from more than one phase; with care, however, mixtures of crystalline materials can be identified.
3. If the material belongs to a family of phases with very similar XRD patterns, it may not be possible to identify unambiguously the phase concerned; for instance, many niobates and tantalates have very similar XRD patterns, e.g. $LiNbO_3$ and $LiTaO_3$, as do many compounds of adjacent rare earth elements.
4. If a crystalline phase has variable composition or is part of a solid solution series, then identification of the composition from a single XRD pattern alone is generally not possible. Using the above example, $LiNbO_3$ and $LiTaO_3$ form a complete range of homogeneous solid solutions but it would not be possible, using the Powder Diffraction File, to identify a particular composition in this series.

5.3.6.1 Focusing of X-rays: theorem of a circle

A major disadvantage of Debye–Scherrer cameras is that incident and diffracted beams are, inevitably, divergent and of low intensity. In diffractometers and focusing cameras, a convergent X-ray beam is used; this gives a dramatic improvement in resolution and, because much more intense beams may be used, exposure times are greatly reduced. It is not possible to focus or converge X-rays using the X-ray equivalent of an optical lens; instead, use is made of certain geometric properties of the circle in order to obtain a convergent X-ray beam. These properties are illustrated in Fig. 5.14(a). The arc XY forms part of a circle and all angles subtended on the circumference of this circle by the arc XY are equal, i.e. $XCY = XC'Y = XC'Y = \alpha$. Suppose that X is a source of X-rays and XC, XC′ represent the extremities of a divergent X-ray beam emitted from X. If the beam is diffracted by a sample which covers the arc between C and C′ such that the diffracting planes are tangential to the circle, then the diffracted beam, represented by CY and C′Y, will focus to a point at Y. The principle of the focusing method is therefore to arrange that the source of X-rays, the sample and the detector all lie on the circumference of an appropriate circle.

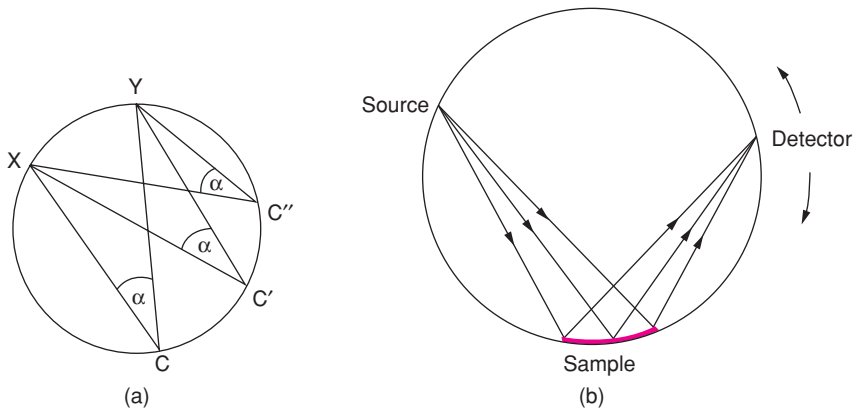


Figure 5.14 ***(a) Theorem of a circle used to focus X-rays. (b) Arrangement of sample, source and detector on the circumference of a circle.***

5.3.6.2 Crystal monochromators

The same focusing principle is used in the construction of diffractometers and focusing cameras, although several different arrangements are found in commercial instruments. An additional feature is often the inclusion of a *crystal monochromator*, which serves two functions: to give highly monochromatic radiation and to produce an intense, convergent X-ray beam. There are several sources of background scattering in diffraction experiments, one of which is radiation of wavelength different from that of the $K\alpha$ radiation. $K\alpha$ radiation may be separated from the rest by the use of filters or, better, by a crystal monochromator.

A crystal monochromator consists of a large single crystal of, for example, quartz, oriented such that one set of planes which diffracts strongly ($10\bar{1}1$ for quartz) is at the Bragg angle to the incident beam. This Bragg angle is calculated for $\lambda_{K\alpha_1}$ and so only the $K\alpha_1$ rays are diffracted, giving monochromatic radiation. If a flat crystal monochromator is used, much of the $K\alpha_1$ radiation would be lost since the X-ray beam emitted from a source is naturally divergent; only a small amount of the $K\alpha_1$ component would therefore be at the correct Bragg angle to the monochromator. To improve the efficiency, the crystal monochromator is bent, in which case a divergent X-ray beam may be used which is diffracted by the crystal monochromator to give a beam that is intense, monochromatic and convergent.

5.3.6.3 Powder diffractometers

The powder diffractometer has a proportional, scintillation or Geiger counter or an image plate detector which scans a range of 2θ values at constant angular velocity (it is common practice to refer to the angle 2θ between diffracted and undiffracted beams, Fig. 5.12, rather than to the Bragg angle, θ). Usually, the range 10–80° 2θ is sufficient to cover the most useful part of the powder pattern. A typical diffractometer trace is shown in Fig. 5.15(a) for the cristobalite polymorph of SiO_2. The scale is linear in 2θ and d-spacings of the peaks may be calculated from Bragg's law or obtained from standard tables of d versus 2θ. The scanning speed of the counter is usually 2° 2θ min^{-1} and, therefore, about 30 min are needed to obtain a trace. Intensities are taken as either peak heights or, for accurate work, peak areas; the most intense peak is given an intensity of 100 and the rest are scaled accordingly. For accurate d-spacings, an internal standard (a pure material, such as KCl, whose d-spacings are known accurately) may be mixed with the sample. A correction factor, which may vary

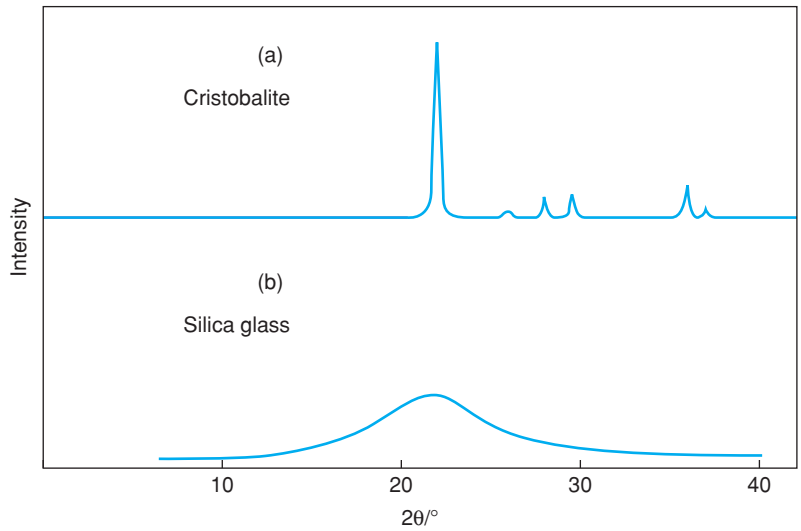


Figure 5.15 ***X-ray powder diffraction patterns of (a) cristobalite and (b) glassy SiO_2; Cu Kα radiation.***

with 2θ, is obtained from the discrepancy between observed and true d-spacings of the standard and is then applied to the pattern that is being measured.

Samples for diffractometry take various forms: they include thin layers of fine powder sprinkled onto a glass slide smeared with vaseline and thin flakes pressed onto a glass slide. The objective is always to obtain a sample which contains a random arrangement of crystal orientations. If the crystal arrangement is not random, then *preferred orientation* exists and can introduce errors, sometimes very large, into the measured intensities. Preferred orientation is a serious problem for materials that crystallise in a characteristic, very non-spherical shape, e.g. clay minerals, which usually occur as thin plates, or some cubic materials, which crystallise as cubes and, on crushing, break up into smaller cubes. In a powder aggregate of such materials, the crystals tend to sit on their faces, resulting in a non-random average orientation.

5.3.6.4 Guinier focusing cameras

The arrangement of a focusing or Guinier camera, which uses a crystal monochromator M and also makes use of the theorem of the circle described above, is shown in Fig. 5.16(a). A convergent beam of monochromatic radiation passes through the sample at X. Radiation that is not diffracted comes to a focus at A, where a beam stop is placed in front of the film to prevent its blackening. Various beams diffracted by the sample focus at B, C, etc. We know from the theorem of the circle that A, B, C and X must lie on the circumference of a circle. The film is placed in a cassette which is in the form of a short cylinder and lies on the circle ABC. The scale of the film is linear in 2θ. A schematic film is as shown in Fig. 5.16(b) except that instead of peaks of different height, lines of different intensity or different degrees of blackness are seen. Film dimensions are $\sim 1 \times 15$ cm, which makes them very convenient to handle. The line at 0° 2θ or ∞ d-spacing corresponds to the undiffracted beam at A in (a) and is the reference position on the film.

The Guinier method is capable of giving accurate d-spacings with results comparable to those obtained by diffractometry. Intensities are either estimated visually or measured using microdensitometry. Sample sizes are small, ≤ 1 mg, and exposure times vary between 5 min and 1 h, depending on factors such as the crystallinity of the sample and the presence or absence of heavy elements which absorb X-rays.

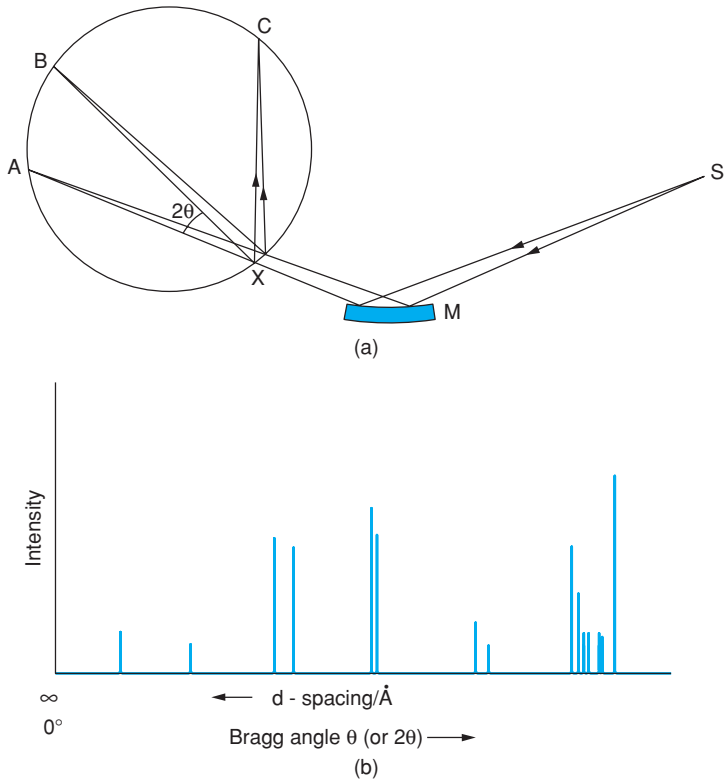


Figure 5.16 ***(a) Crystal monochromator M, source S and sample X, in a focusing camera. (b) Schematic Guinier powder XRD pattern.***

5.3.6.5 A powder pattern of a crystalline phase is its 'fingerprint'

There are two main factors which determine XRD patterns: (i) the size and shape of the unit cell and (ii) the atomic number and position of atoms in the cell. Thus, two materials may have the same crystal structure but almost certainly will have distinct powder patterns. For example, KF, KCl and KI all have the rock salt structure and should show the same set of lines in their powder patterns, but, as can be seen from Table 5.3, both the positions and intensities of the lines are different in each. The positions or *d*-spacings vary because

Table 5.3 ***X-ray powder diffraction patterns for potassium halides***

	KF, $a = 5.347$ Å		KCl, $a = 6.2931$ Å		KI, $a = 7.0655$ Å	
(*hkl*)	*d*/Å	*I*	*d*/Å	*I*	*d*/Å	*I*
111	3.087	29		–	4.08	42
200	2.671	100	3.146	100	3.53	100
220	1.890	63	2.224	59	2.498	70
311	1.612	10	–	–	2.131	29
222	1.542	17	1.816	23	2.039	27
400	1.337	8	1.573	8	1.767	15

the unit cells are of different size and, therefore, the a parameter in the d-spacing equation varies. Intensities differ because different anions with different atomic numbers and therefore different scattering powers are present in the three materials, even though the atomic coordinates are the same for each (i.e. cations at corner and face centre positions, etc.). KCl is a rather extreme example because the intensities of 111 and 311 reflections are too small to measure, but it serves to illustrate the importance of scattering power of the atoms present. Intensities are discussed in more detail in the next section.

A powder XRD pattern has two characteristic features, therefore: the d-spacings of the lines and their intensity. Of the two, the d-spacing is more useful and capable of precise measurement. The d-spacings should be reproducible from sample to sample unless impurities are present to form a solid solution. Intensities are more difficult to measure quantitatively and often vary from sample to sample, especially if a preferred orientation is present. Thus, the differences in tabulated intensities for, say, the (220) reflection of the three materials in Table 5.3 are probably not absolute, quantitatively.

The likelihood of two materials having the same cell parameters and d-spacings decreases considerably with decreasing crystal symmetry. Thus, cubic materials have only one variable, a, and there is a fair chance of finding two materials with the same a value. On the other hand, triclinic powder patterns have six variables, a, b, c, α, β and γ, and so accidental coincidences are far less likely. Problems of identification, if they occur, are most likely to be experienced with high symmetry, especially cubic, materials or in cases where similar-sized ions may replace each other in a particular structure as indicated in Section 5.3.6.

5.3.6.6 Powder patterns and crystal structures

Although powder patterns are invaluable and unique for fingerprint identification of crystalline materials, a great deal of additional crystallographic information can be obtained from powder patterns without the necessity for full structure determination (which is also possible, as discussed later). As a first step, it is necessary to be able to index the powder patterns, i.e. to assign h, k, l indices to all lines in the powder diffraction pattern, and at the same time obtain the unit cell symmetry and dimensions. A general observation is that crystals with higher symmetry have simpler powder patterns, i.e. with fewer lines, than samples with lower symmetry recorded under the same experimental conditions. Thus, a cubic material may have 10 lines in a powder pattern recorded by a conventional diffractometer whereas a similar-sized triclinic crystal may have ~100 lines, many of which are of low intensity. This is because in a cubic powder pattern, each line represents the overlap of a number (up to 48) of lines of the same d-spacing, whereas in a material of lower symmetry, these lines do not overlap if they have different d-spacings.

The number of lines that overlap as a consequence of crystal symmetry is the *multiplicity*. For example, in a cubic material, a line with indices $h00$ has multiplicity of 6 ($h00$, $\bar{h}00$, $0k0$, $0\bar{k}0$, $00l$, $00\bar{l}$). In a tetragonal crystal for which $a = b$, the multiplicity of $h00$ is 4 (i.e. $h00$, $\bar{h}00$, $0k0$, $0\bar{k}0$), whereas the multiplicity of $00l$ is only 2 ($00l$, $00\bar{l}$). In an orthorhombic crystal, for which $a \neq b \neq c$ the reflection $h00$ has a multiplicity of 2 ($h00$, $\bar{h}00$). This means that a single line of multiplicity 6 would appear in a powder pattern of a cubic substance whereas two lines of multiplicities 4 and 2 would appear in a tetragonal powder pattern and three lines of multiplicities 2, 2 and 2 in an orthorhombic powder pattern.

A good example of this effect is seen in the powder patterns of different polymorphs of $BaTiO_3$, shown in Fig. 5.17. $BaTiO_3$ shows the following sequence of polymorphs and symmetries with increasing temperature:

$$\text{rhombohedral} \xrightarrow{-110^\circ\text{C}} \text{orthorhombic} \xrightarrow{28^\circ\text{C}} \text{tetragonal} \xrightarrow{125^\circ\text{C}} \text{cubic} \xrightarrow{1470^\circ\text{C}} \text{hexagonal}$$

The cubic polymorph has a single line for the $h00$ reflection whereas the tetragonal polymorph shows a doublet with intensity ratio 2:1, reflecting the fact that the multiplicity ratio is 2:1; the orthorhombic

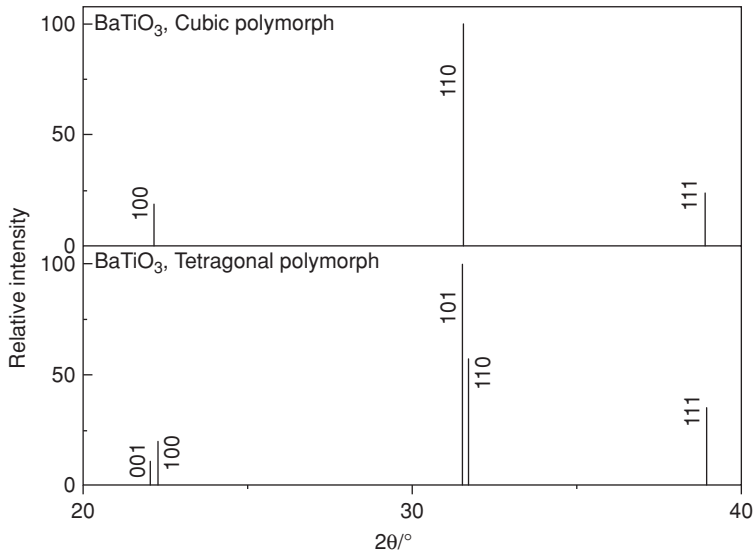


Figure 5.17 ***Section of the powder XRD patterns of $BaTiO_3$ showing the cubic polymorph and line splittings associated with the tetragonal polymorph. Note the difference in intensities of 001 and 100 lines and of the 101 and 110 lines of tetragonal $BaTiO_3$, which are a direct consequence of their different multiplicities. By contrast, the 111 reflection is a singlet in both cubic and tetragonal polymorphs.***

polymorph (not shown) gives a triplet. The powder patterns enable us not only to identify the material as $BaTiO_3$, but also to specify which polymorph, and symmetry, is being measured. This is important if one is interested in the material properties since, for instance, tetragonal $BaTiO_3$ is ferroelectric whereas cubic $BaTiO_3$ is not.

Order–disorder phenomena are very important in many inorganic materials and can be seen clearly from changes in the XRD powder patterns. In Fig. 5.18 are shown powder patterns of several phases that have the basic rock salt structure but with various cation arrangements. In CoO (a), the structure is *fcc* rock salt and the first two lines in the cubic XRD pattern have indices (111) and (200). In $LiCoO_2$, (b), Li and Co cations occupy layers of alternate octahedral sites within the *ccp* oxide array; this cation ordering gives a larger unit cell, which is rhombohedral. Extra lines appear in the XRD pattern; the lines that correspond to the rock salt subcell are still present, but appear at higher 2θ (lower d-spacing) since there is a contraction in interatomic distances. In addition, the *hkl* indices are different since the unit cell is different.

In Li_2MnO_3, Fig. 5.18(c), the cation arrangement is also layered but fully occupied Li layers alternate with layers containing a mixture of Li and Mn arranged in ordered fashion; this structure is monoclinic and again, extra reflexions appear in comparison with CoO (a). All three powder patterns therefore contain the same basic sub-cell lines that are observed in a cubic rock salt structure, such as that of CoO, (a), but in addition, the patterns of $LiCoO_2$ and Li_2MnO_3 contain extra sets of reflexions associated with the cation order, which is different in these two cases.

5.3.7 Intensities

Intensities of X-ray reflections are important for two main reasons. First, quantitative measurements of intensity are necessary in order to solve crystal structures. Second, qualitative or semi-quantitative intensity

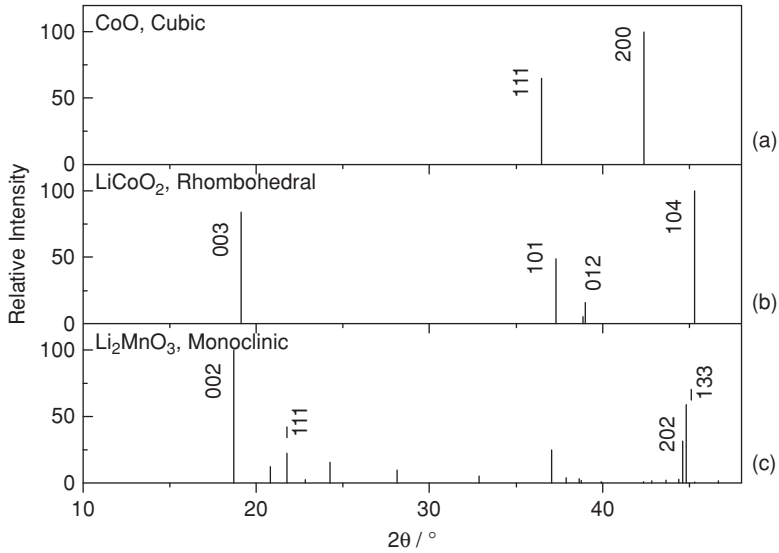


Figure 5.18 ***XRD patterns of some rock salt-related phases.***

data are needed in using the powder fingerprint method to characterise materials and especially in using the *Powder Diffraction File* to identify unknowns. Although this book is not concerned with the details of the methods of structure determination, it is important that the factors which control the intensity of X-ray reflections be understood. The topic falls into two parts: the intensity scattered by individual atoms and the resultant intensity scattered from the large number of atoms in a crystal.

5.3.7.1 Scattering of X-rays by an atom: atomic scattering factors or form factors

Atoms diffract or scatter X-rays because an incident X-ray beam, which can be described as an electromagnetic wave with an oscillating electric field, sets each electron of an atom into vibration. A vibrating charge such as an electron emits radiation which is in phase or *coherent* with the incident X-ray beam. The electrons therefore act as secondary point sources of X-rays. Coherent scattering may be likened to an elastic collision between the wave and the electron: the wave is deflected by the electron without loss of energy and, therefore, without change of wavelength. The intensity of the radiation scattered coherently by 'point source' electrons is given by the *Thomson equation*:

$$I_P \propto \frac{1}{2}(1 + \cos^2 2\theta) \tag{5.5}$$

where I_P is the scattered intensity at any point, P, and 2θ is the angle between the directions of the incident beam and the diffracted beam that passes through P. From this equation, the scattered beams are most intense when parallel or antiparallel to the incident beam and are weakest when at 90° to the incident beam. The Thomson equation is also known as the *polarisation factor* and is one of the standard angular correction factors that must be applied when processing intensity data for use in structure determination.

At this point, it is worth mentioning that X-rays can interact with electrons in a different way to give *Compton scattering*. Compton scattering is an inelastic collision since the X-rays lose some of their energy

on impact and the scattered X-rays are therefore of longer wavelength. They are also no longer in-phase with the incident X-rays; nor are they in-phase with each other. Compton scattering is caused by interaction between X-rays and more loosely held valence electrons; it is an important effect with light elements and can have a deleterious effect on the powder patterns of organic materials such as polymers. A close similarity exists between Compton scattering and the generation of white radiation in an X-ray tube; both are examples of incoherent scattering that are sources of background radiation in XRD experiments.

The X-rays scattered by an atom are the result of the waves scattered by each electron in the atom. For this purpose, electrons may be regarded as particles that momentarily occupy different positions in an atom and interference occurs between their scattered waves. For scattering in the direction of the incident beam, Fig. 5.19(a), beams 1′ and 2′, all electrons scatter in-phase irrespective of their position and the scattered intensity is the sum of the individual intensities. The *scattering factor*, or *form factor*, *f*, of an atom is proportional to its atomic number, *Z*, or, more strictly, to the number of electrons possessed by that atom.

For scattering at some angle 2θ to the direction of the incident beam, a phase difference, corresponding to the distance XY, exists between beams 1′ and 2′. This phase difference is usually less than one wavelength (i.e. XY $<$1.5418 Å for Cu $K\alpha$ X-rays) because distances between electrons within an atom are short and only partial destructive interference occurs between 1′ and 2′. The net effect of such interference is to cause a gradual decrease in scattered intensity with increasing angle, 2θ. For example, *f* of Cu is proportional to 29 (i.e. *Z*) at $2\theta = 0^\circ$, to 14 at 90° and to 11.5 at 120°. It should also be apparent that for a given angle, 2θ, the net intensity decreases with decreasing λ because the phase difference XY gives a greater degree of cancellation for smaller λ. The form factors of atoms are given in *International Tables for X-Ray Crystallography*, Vol. 3. They are tabulated against sin θ/λ to include the effect of both angle and X-ray wavelength; examples are shown in Fig. 5.19(b).

Two consequences of the dependence of *f* on sin θ/λ and *Z* are as follows. First, XRD patterns at high angles (above ~60–70° 2θ) are usually weak. Second, in X-ray crystal structure determination it is difficult to locate light atoms because their diffracted radiation is so weak. Thus H cannot usually be located unless all the other elements present are also extremely light. Atoms that have as many electrons as oxygen can usually be located easily unless heavy atoms such as uranium are present. Structures that may be difficult to solve are those in which a considerable number or all of the atoms have similar atomic number, e.g. large organic molecules with C, N and O. In such cases, a common ploy is to make a derivative of the compound of interest which contains a heavy metal atom. The heavy atoms are detected readily because they determine the phase of diffracted beams, and this helps towards locating the remaining atoms. Because of their similar atomic numbers, Al and Si are difficult to distinguish, which may cause problems in the determination of aluminosilicate structures. One advantage of using neutrons instead of (or in addition to) X-rays for crystallographic work is that neutron form factors are not a simple function of atomic number. Light atoms, e.g. H and Li, are often strong neutron scatterers.

5.3.7.2 Scattering of X-rays by a crystal – systematic absences

Earlier, we treated diffraction in terms of Bragg's law, with crystals divided up into sets of imaginary lattice planes, defined with reference to the shape and dimensions of the unit cell. Each set of planes, has a characteristic *d*-spacing and the Bragg angle can be calculated for a given wavelength. The number of possible sets of planes is limited since, using equation (1.1) for orthogonal crystals, i.e. with $\alpha = \beta = \gamma = 90^\circ$, *h*, *k* and *l* must be integers. It is possible to calculate all possible *d*-spacing values using equation (1.1) or the appropriate equations for other unit cell shapes, although the calculation is usually terminated when either a minimum *d*-spacing or maximum set of indices is reached. This has been done for a hypothetical orthorhombic crystal, with all possible *h*, *k*, *l* combinations of 0 and 1, in Table 5.4. Obviously, the list could be extended for higher indices.

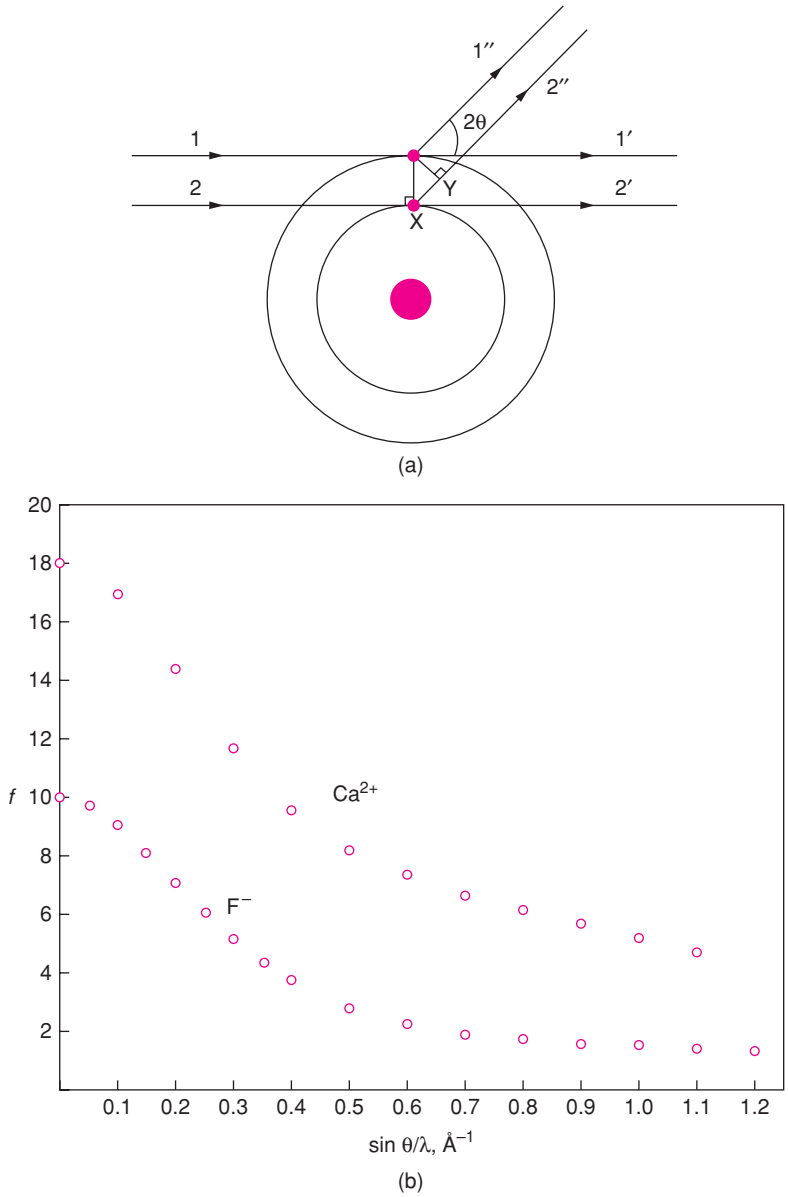


Figure 5.19 ***(a) Scattering of X-rays by electrons in an atom. (b) Form factors of Ca^{2+} and F^{-}.***

In principle, each set of lattice planes can give rise to a diffracted beam. In practice, the intensity of the beams diffracted by certain sets of lattice planes may be zero. These are known as *systematic absences*. Systematic absences arise if the lattice type is non-primitive (I, F, etc.) or if elements of space symmetry (screw axes, glide planes) are present.

As an example of absences due to lattice type, consider α-Fe, Fig. 5.20(a), which is *bcc*. Reflection from the (100) planes (b) has zero intensity and is systematically absent. This is because, at the Bragg angle for

Table 5.4 ***Calculated* d-*spacings for an orthorhombic cell, for* a = *3.0,* b = *4.0,* c = *5.0 Å***

hkl	*d*/Å	*hkl*	*d*/Å
001	5.00	101	2.57
010	4.00	110	2.40
011	3.12	111	2.16
100	3.00		

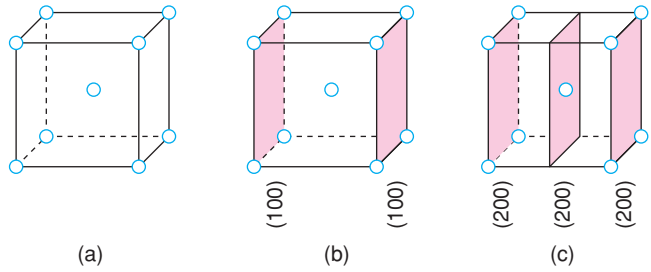


Figure 5.20 ***(a)* bcc *α-Fe; (b) (100) planes; (c) (200) planes.***

these planes, the body centre atoms which lie midway between adjacent (100) planes diffract X-rays exactly 180° out-of-phase relative to the corner atoms which lie on the (100) planes. Averaged over the whole crystal, there are equal numbers of corner and body centre atoms and the beams diffracted by each cancel completely. In contrast, a strong 200 reflection is observed because all the atoms lie on (200) planes (c) and there are no atoms lying between (200) planes to cause destructive interference. It is easy to show, by similar arguments, that the (110) reflection is observed whereas (111) is systematically absent in α-Fe. For each non-primitive lattice type there is a simple characteristic equation for systematic absences, Table 5.5. For a body centred cell, reflections for which $(h + k + l)$ is odd are absent, e.g. reflections such as (100), (111), (320), etc., are systematically absent (see the next section for proof of this condition).

Systematic absences are an extreme case of destructive interference between X-ray beams diffracted by individual atoms. They arise when one set of atoms diffracts X-rays that are exactly out-of-phase with those diffracted by a second set of atoms of the same type. Two conditions must be met for systematic absences:

Equal numbers of diffracted beams must be (i) out-of-phase (by λ/2 or π) and (ii) of the same amplitude (determined by form factors, f).

Table 5.5 ***Systematic absences due to lattice type***

Lattice type	Rule for reflection to be observed[a]
Primitive, P	None
Body centred, I	hkl; $h + k + l = 2n$
Face centred, F	hkl; h, k, l either all odd or all even
Side centred, e.g. C	hkl; $h + k = 2n$
Rhombohedral, R	hkl; $-h + k + l = 3n$ or $(h - k + l = 3n)$

[a]If space symmetry elements are present, additional rules may apply. These are not dealt with here.

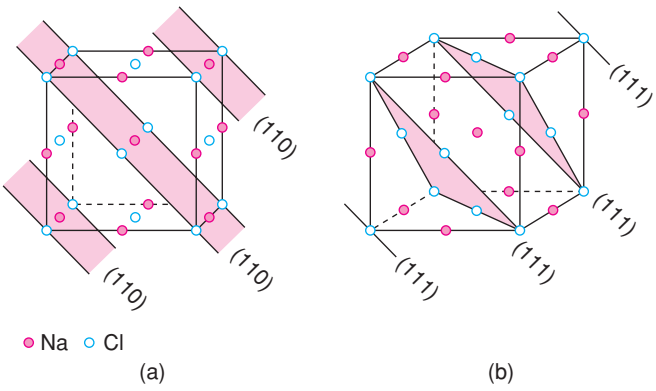


Figure 5.21 ***(a) (110) and (b)(111) planes in NaCl.***

In cases where destructive interference is not complete and reduced intensities are observed, then one or both of these conditions is not met.

Let us consider now an example of partially destructive interference in the rock salt structure. Rock salt is *fcc* and, therefore, only those reflections for which h, k, l are either all odd or all even may be observed (Table 5.5). From this rule, for instance, (110) is systematically absent. In Fig. 5.21(a), (110) planes have Na^+ and Cl^- ions on the planes but equal numbers of the same ions are midway between the planes. Both conditions specified above are met and complete cancellation occurs. For the (111) planes, however, Cl^- ions lie on the planes and Na^+ ions lie exactly midway between them. Hence the Na^+ and Cl^- ions scatter exactly 180° out-of-phase with each other for these planes, but since they have different form factors, the destructive interference that occurs is only partial. The intensity of the (111) reflection in materials that have the rock salt structure is, therefore, related to the difference in atomic number of anion and cation. For the potassium halides, (111) has essentially zero intensity for KCl, since K^+ and Cl^- are isoelectronic, and the intensity of (111) should increase in the order

$$KCl < KF < KBr < KI$$

Some data which illustrate this are given in Table 5.3.

Similar effects occur in other simple structures. In primitive cubic CsCl, if the difference between caesium and chlorine is ignored, the atomic positions are the same as in body centred α-Fe, Fig. 5.20. The (100) reflection is systematically absent in α-Fe but is observed with CsCl because the form factors of Cs^+ and Cl^- are different, i.e. $f_{Cs^+} \neq f_{Cl^-}$.

5.3.7.3 General equation for phase difference, δ

Each atom in a crystal scatters X-rays by an amount related to the scattering power, f, of that atom. In summing the individual waves to give the resultant diffracted beam, both the *amplitude* and *phase* of each wave are important. If we know the atomic positions in the structure, the amplitude and phase of the wave scattered by each atom in the unit cell in the direction that satisfies the Bragg condition for a particular *hkl* reflection may be calculated. To determine the overall intensity of that reflection, waves scattered by each atom of the unit cell may be summed mathematically, therefore simulating what happens during diffraction. Let us consider first the *relative phases* of different atoms in the unit cell.

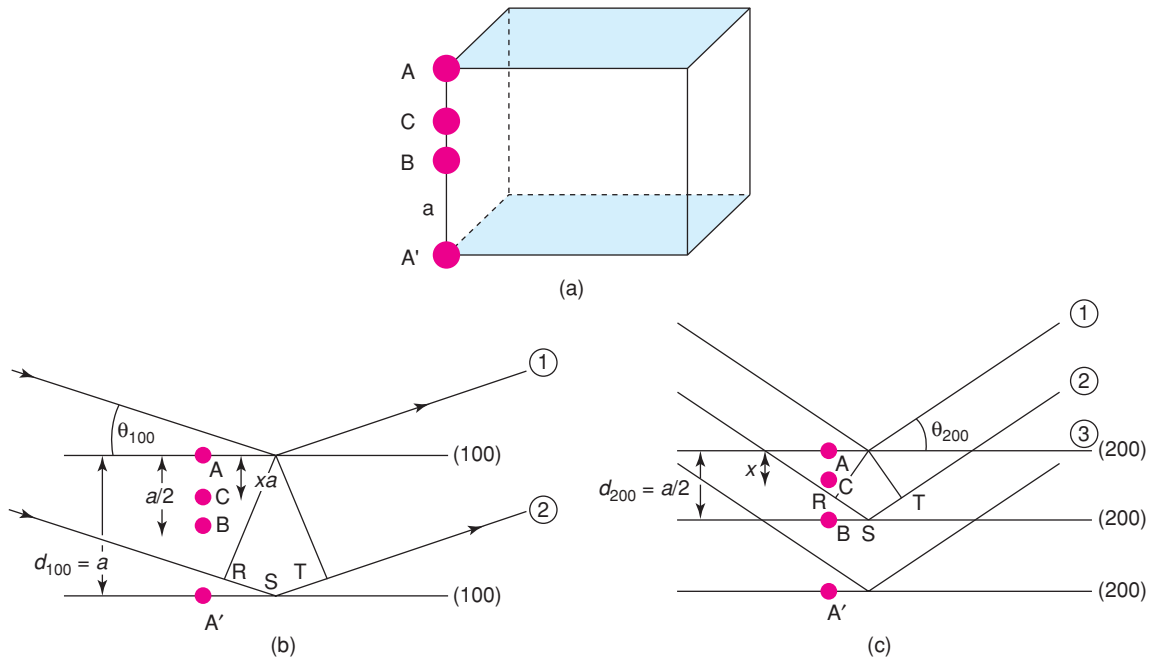


Figure 5.22 ***(a, b) (100) planes for an orthogonal unit cell (α = β = γ = 90°). Atoms A, B, C, A′ lie on the a cell edge. (c) (200) planes for the same unit cell as in (a, b).***

In Fig. 5.22(a) are drawn two (100) planes of a crystal that has an orthogonal unit cell (note that the term 'orthogonal' refers to a unit cell whose angles are 90°; this is different from 'orthorhombic', which describes a unit cell symmetry). Atoms A, B, C, A′ lie on the a axis [perpendicular to (100) planes] with A and A′ at the origin of adjacent unit cells. For the 100 reflection, from Bragg's law, A and A′ scatter in-phase because their phase difference is exactly one wavelength, 2π rad. Atom B, situated half way between adjacent (100) planes, has fractional x coordinate (relative to A) of $\frac{1}{2}$. The phase difference between (waves diffracted from) A and B is $\frac{1}{2} \times 2\pi = \pi$, i.e. atoms A and B are exactly out-of-phase. Atom C has a general fractional coordinate x (at distance xa from A) and, therefore, a phase relative to A of $2\pi x$.

Consider, now, the (200) reflection for the same unit cell, Fig. 5.22(c). Since $d_{200} = \frac{1}{2}d_{100}$, then from Bragg's law, $\sin\theta_{200} = 2\sin\theta_{100}$ and therefore $\theta_{200} \gg \theta_{100}$. Atoms A and B have a phase difference of 2π for the (200) reflection and scatter in-phase, whereas their phase difference is π for the (100) reflection. Thus, the effect of halving d and increasing the Bragg angle is to double the relative phase difference between pairs of atoms such as A and B; therefore, A and C have a phase difference of $(2x \times 2\pi)$ for the (200) reflection.

For the general case of an $h00$ reflection, the d-spacing between adjacent $(h00)$ planes is $(1/h)a$ and the *phase difference*, δ, between A and C is given by

$$\delta = 2\pi hx \tag{5.6}$$

The phase difference between atoms depends, therefore, on two factors: the indices of the reflection that is being considered and the fractional coordinates of the atoms in the unit cell. This reasoning may be extended

readily to diffraction in 3D. For reflection from the set of planes with indices (*hkl*), the phase difference, δ, between atoms at the origin and a position with fractional coordinates x, y, z is given by

$$\delta = 2\pi(hx + ky + lz) \tag{5.7}$$

This important equation is applicable to all unit cell shapes. Let us use it on a simple structure, γ-Fe, which is *fcc* with atoms at the corner and face centre positions, i.e. with fractional coordinates

$$0, 0, 0; \tfrac{1}{2}, \tfrac{1}{2}, 0; \tfrac{1}{2}, 0, \tfrac{1}{2}; 0, \tfrac{1}{2}, \tfrac{1}{2}$$

These coordinates may be substituted into the above equation for δ to give four phases:

Atomic position	Phase δ relative to origin
0, 0, 0	0
$\frac{1}{2}, \frac{1}{2}, 0$	$\pi(h + k)$
$\frac{1}{2}, 0, \frac{1}{2}$	$\pi(h + l)$
$0, \frac{1}{2}, \frac{1}{2}$	$\pi(k + l)$

How do these vary with h, k and l? If h, k and l are either all even or all odd, the phases are in multiples of 2π and, therefore, are in-phase with each other. If, however, one, say h, is odd and the other two, k and l, are even, the four phases reduce to

$$0, \quad (2n + 1)\pi, \quad (2n + 1)\pi, \quad 2n\pi$$

The first and last are π out-of-phase with the middle two and complete cancellation occurs. The γ-Fe structure is a simple example of an *fcc* lattice in which the Fe atoms correspond to lattice points; we have in fact just proved the condition for systematic absences due to face centring! (Table 5.5). The reader may like to prove the condition for systematic absences in a *bcc* structure by working out the phases of the atoms for the structure of α-Fe.

5.3.7.4 Intensities and structure factors

Let us now generalise the treatment to consider any atom, j, in the unit cell. The diffracted wave of amplitude f_j and phase δ_j is a sine wave of the form

$$F_j = f_j \sin(\omega t - \delta_j) \tag{5.8}$$

The waves diffracted from each atom in the cell have the same angular frequency, ω, but may differ in f and δ. The resultant intensity is obtained from the summation of the individual sine waves. Mathematically, addition of waves may be carried out by various methods, including vector addition and the use of complex numbers. In complex notation, wave j may be written as

$$F_j = f_j(\cos\delta_j + i \sin\delta_j) \tag{5.9}$$

or as

$$F_j = f_j \exp(i\delta_j) \tag{5.10}$$

where $i = \sqrt{-1}$

The intensity of a wave is proportional to the square of its amplitude, i.e.

$$I \propto f_j^2 \tag{5.11}$$

and is obtained by multiplying the equation for the wave by its complex conjugate; i.e.

$$I \propto f_j \exp(i\delta_j) \times f_j \exp(-i\delta_j)$$

and therefore

$$I \propto f_j^2$$

Alternatively,

$$[f_j(\cos\delta_j + i\sin\delta_j)][f_j(\cos\delta_j - i\sin\delta_j)] = f_j^2(\cos^2\delta_j + \sin^2\delta_j) = f_j^2$$

Substituting the expression for δ, equation (5.7), into equation (5.10), the equation of a diffracted wave becomes

$$\begin{aligned} F_j &= f_j \exp 2\pi i(hx_j + ky_j + lz_j) \\ &= f_j[\cos 2\pi(hx_j + ky_j + lz_j) + i\sin 2\pi(hx_j + ky_j + lz_j)] \end{aligned} \tag{5.12}$$

When written in this form, the summation over the j atoms in the unit cell may be carried out readily, to give the *structure factor* or *structure amplitude*, F_{hkl}, for the hkl reflection, i.e.

$$F_{hkl} = \sum_{j=1\to n} f_j \exp(i\delta_j)$$

or

$$F_{hkl} = \sum_j f_j(\cos\delta_j + i\sin\delta_j) \tag{5.13}$$

The intensity of the diffracted beam I_{hkl} is proportional to $|F_{hkl}|^2$ and is obtained from

$$\begin{aligned} I_{hkl} \propto |F_{hkl}|^2 &= \left[\sum_j f_j(\cos\delta_j + i\sin\delta_j\right]\left[\sum_j f_j(\cos\delta_j + i\sin\delta_j)\right] \\ &= \sum_j (f_j\cos\delta_j)^2 + \sum_j (f_j\sin\delta_j)^2 \end{aligned} \tag{5.14}$$

This is a very important equation in crystallography because, using it, the intensity of any *hkl* reflection may be calculated from a knowledge of the atomic coordinates in the unit cell. It is instructive to see one example of its use, although in practice, the different stages and calculations form part of automated software requiring little input. Calcium fluoride, CaF_2, has the fluorite structure with atomic coordinates x, y, z in the *fcc* unit cell, listed in column 2:

Atom	*x y z*	δ	δ for (202)	δ for (212)
Ca	0 0 0	$2\pi(0)$	0	0
	$\frac{1}{2}\ \frac{1}{2}\ 0$	$\pi(h+k)$	2π	3π
	$\frac{1}{2}\ 0\ \frac{1}{2}$	$\pi(h+l)$	4π	4π
	$0\ \frac{1}{2}\ \frac{1}{2}$	$\pi(k+l)$	2π	3π
F	$\frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}$	$\pi/2(h+k+l)$	2π	$5\pi/2$
	$\frac{1}{4}\ \frac{1}{4}\ \frac{3}{4}$	$\pi/2(h+k+3l)$	4π	$9\pi/2$
	$\frac{1}{4}\ \frac{3}{4}\ \frac{1}{4}$	$\pi/2(h+3k+l)$	2π	$7\pi/2$
	$\frac{3}{4}\ \frac{1}{4}\ \frac{1}{4}$	$\pi/2(3h+k+l)$	4π	$9\pi/2$
	$\frac{3}{4}\ \frac{3}{4}\ \frac{1}{4}$	$\pi/2(3h+3k+l)$	4π	$11\pi/2$
	$\frac{3}{4}\ \frac{1}{4}\ \frac{3}{4}$	$\pi/2(3h+k+3l)$	6π	$13\pi/2$
	$\frac{1}{4}\ \frac{3}{4}\ \frac{3}{4}$	$\pi/2(h+3k+3l)$	4π	$11\pi/2$
	$\frac{3}{4}\ \frac{3}{4}\ \frac{3}{4}$	$\pi/2(3h+3k+3l)$	6π	$15\pi/2$

Substitution of the atomic coordinates into equation (5.7) gives the δ values listed in column 3. Since the fluorite structure is *fcc*, h, k and l must be all odd or all even for an observed reflexion. This is shown in column 4 for the (202) reflexion. For any other combination, $F = 0$, as shown for the (212) reflexion, column 5. Therefore, making the summations for all 12 atoms in the unit cell we obtain

	Reflexion (202)	Reflexion (212)
$\Sigma f_{Ca}\cos\delta$	$4f_{Ca}$	0
$\Sigma f_{Ca}\sin\delta$	0	0
$\Sigma f_{F}\cos\delta$	$8f_{F}$	0
$\Sigma f_{F}\sin\delta$	0	0
F_{202}	$4f_{Ca} + 8f_{F}$	
F_{212}		0

The (202) reflexion in CaF_2 has a d-spacing of 1.929 Å ($a = 5.464$ Å). Therefore,

$$\theta_{202} = 23.6^\circ \text{ and } \sin\theta/\lambda = 0.259 \text{ for } \lambda = 1.5418\ \mathring{A}\ (\text{Cu } K\alpha)$$

Form factors for Ca and F are given in Fig. 5.19(b); for $\sin\theta/\lambda = 0.259$, by interpolation,

$$f_{Ca} = 12.65 \text{ and } f_{F} = 5.8$$

Table 5.6 ***Structure–factor calculations for CaF_2: powder XRD***

d/Å	hkl	I	Multiplicity[a]	I/(multiplicity × L_p)[b]	F^{obs}	F^{calc}	F^{obs} scaled	$\|\|F^{obs}\| - \|F^{calc}\|\|$
3.143	111	100	8	0.409	0.640	67	90	23
1.929	202	57	12	0.476	0.690	97	97	0
1.647	311	16	24	0.098	0.313	47	44	3
1.366	400	5	6	0.193	0.439	75	62	13
1.254	331	4	24	0.047	0.217	39	31	8

$\sum F^{obs}\text{scaled} = 324.$

$\sum \left| |F^{obs}| - |F^{calc}| \right| = 47.$

$$R = \frac{\sum |\Delta F|}{\sum F^{obs}} = \frac{47}{324} = 0.15.$$

[a]The multiplicity of an X-ray powder line is the number of equivalent sets of planes that diffract at the same Bragg angle. Thus, overlapping with 111 are $\bar{1}11$, $1\bar{1}1$, $11\bar{1}$, $\bar{1}\bar{1}1$, $\bar{1}1\bar{1}$, $1\bar{1}\bar{1}$ and 111, giving a multiplicity value of 8, where a negative index indicates that a negative crystallographic direction was used in defining it.
[b]The Lorentz polarisation factor, L_p, is an angular correction factor that includes the effect of equation (5.5) and certain instrumental factors. Available from standard tables.

Therefore,

$$F_{202} = 97$$

This calculation may be made for a series of *hkl* reflections and the results, after scaling, compared with the observed values, Table 5.6. In solving unknown structures, the objective is always to obtain a model structure for which the calculated structure factors, F_{hkl}^{calc}, are in the best possible agreement with experimental values, F_{hkl}^{obs}.

An important feature which simplifies the above calculations is that all the sine terms are zero. This is because the origin of the unit cell coincides with a centre of symmetry. Therefore, for each atom at position x, y, z there is a centrosymmetrically related atom at $-x$, $-y$, $-z$ (e.g. F at $^1/_4$, $^1/_4$, $^1/_4$ and $-^1/_4$, $-^1/_4$, $-^1/_4$, i.e. $1 - ^1/_4$, $1 - ^1/_4$, $1 - ^1/_4$ or $^3/_4$, $^3/_4$, $^3/_4$) and since $\sin(-\delta) = -\sin\delta$, the summation of the sine terms over the unit cell contents is zero. If, on the other hand, one of the F atoms was taken as the origin of the cell, the sine terms would be non-zero because F, with its immediate coordination environment of 4Ca arranged tetrahedrally, does not lie on a centre of symmetry. Many structures, of course, belong to non-centric space groups, in which case the complete calculation of F using both cosine and sine terms cannot be avoided.

5.3.7.5 Temperature factors

In a crystal at all real temperatures, atoms are vibrating at infrared frequencies. Since the vibrational frequency is several orders of magnitude less than that of the X-rays used for diffraction experiments, the resulting diffracted beams arise from the time average of the positions of the vibrating atoms. Effectively, therefore, the electron density of each atom is smeared out over a volume whose size depends on the amplitude of thermal vibration. The net effect is to increase the destructive interference between X-rays scattered from different parts of the same vibrating atom and cause a slightly greater decrease in form factors as a function of $\sin\theta/\lambda$ than that shown in Fig. 5.19(b). The form factor is given by

$$f = f_0 \exp\left(-B\frac{\sin^2\theta}{\lambda^2}\right) \tag{5.15}$$

where f is the form factor at the temperature of measurement, f_0 is the form factor calculated for the atom at rest and B (or sometimes U) is the temperature factor. B is related to the thermal vibration of atoms by

$$B = 8\pi^2\bar{u}^2 \tag{5.16}$$

where $\bar{u}^2$is the mean square displacement of the atom from its rest position. For most purposes, it is assumed that atoms vibrate isotropically and effectively are spherical, but for precise work, especially in the final details of structure refinement, anisotropic temperature factors, B_{aniso}, may be used.

5.3.7.6 R-factors and structure determination

It was shown above how the structure factor, F_{hkl}^{calc}, may be calculated for any *hkl* reflection from a knowledge of the coordinates of the atoms in the unit cell. The values of F_{hkl}^{calc} for the first five lines in the powder pattern of CaF_2 are given in Table 5.6. The experimental intensities, the intensities after correction for the L_p factor and *multiplicities* and the observed structure factor, F_{hkl}^{obs}, given by $F_{hkl}^{\text{obs}} = \sqrt{I_{\text{corr}}}$, are also listed. In order to compare the values of F_{hkl}^{obs} and F_{hkl}^{calc}, they are scaled such that $\sum F_{hkl}^{\text{obs}} = \sum F_{hkl}^{\text{calc}}$. Multiplication of each F_{hkl}^{obs} value by 141 gives the scaled values listed. The measure of agreement between the individual, scaled F_{hkl}^{obs} and F_{hkl}^{calc} values is given by the *residual factor* or *R-factor*, defined by

$$R = \frac{\sum \left| \left| F^{\text{obs}} \right| - \left| F^{\text{calc}} \right| \right|}{\sum \left| F^{\text{obs}} \right|} \tag{5.17}$$

An R-factor of 0.15 (or 15% after multiplying by 100) is obtained.

In solving unknown structures, one is guided by, among other things, the value of R; the lower it is, the more likely is the structure to be correct. The calculation given for CaF_2 is rather artificial since only five reflections were used (normally, hundreds or thousands are used), but it serves as an illustration. It is not possible to give hard and fast rules linking the magnitude of R and the likely correctness of a structure but, usually, when R is less than 0.1–0.2, the proposed structure is essentially correct. A structure which has been solved fully using good-quality intensity data has R typically in the range 0.02–0.06.

5.3.7.7 Structure refinement from powder data: Rietveld refinement

In recent years, the Rietveld refinement method (named after the originator of the method, Dutch physicist Hugo Rietveld) has become extremely valuable as a method to confirm structural details of powdered samples. It is not an *ab initio* structure determination method for which single-crystal data are much preferred, but is extremely valuable for materials in which the approximate structure type may be known but for which details need to be determined, or confirmed. Rietveld refinement is a whole pattern refinement in which the experimental powder XRD profile is compared with a calculated profile, whose parameters are adjusted by refinement. The method was devised by Rietveld initially for processing powder neutron diffraction data but has since been extended to XRD data, including those collected from laboratory instruments. Today, widely available software packages make this a semi-routine method of structure analysis (although there are pitfalls; see later).

Rietveld refinement requires a starting model for a crystal structure from which a calculated powder XRD profile can be obtained. For a known structure type, this means that reasonable guesses for the values of a number of refinable parameters are required. If a trial structure model is not available, structure solution from powder XRD data may be attempted using the various methods outlined in Section 5.3.8; it is necessary to deconvolute all the contributing peaks in the powder XRD pattern so that their intensities can be determined and converted into $|F_{hkl}^{\text{obs}}|$ values for inputting into Fourier calculations. The limitation of powder XRD (and

powder neutron) patterns is that they are one-dimensional summations (with a single variable 2θ) of what in reality is a 3D diffraction process. With lower symmetry phases in particular, this means that the powder patterns may contain many partially overlapping reflections, raising the question as to how their individual intensities can be determined, especially if there is uncertainty over individual peak shapes. Since the Rietveld method is a whole pattern refinement, this includes the background scatter, which must be known as accurately as possible in order to model it in the calculated diffraction profile.

Finally, a word of caution. If the parameters of the trial structure differ somewhat from those of the true structure, it is very easy for the refinement to produce a false minimum in which some of the structural parameters may be incorrect without there being any easy way of knowing which, if any, of the parameters are incorrect. This may happen if too many variables are allowed to refine at any time. The strategy should therefore be to proceed with caution: refine each variable one-by-one and then either fix them at their refined values or allow the number of refined parameters to increase gradually.

The R-factor used to assess the quality of a single crystal structure refinement, equation (5.17), may also be used for assessing the quality of a Rietveld refinement, provided that individual peak intensities have been extracted; numerous other indicators of refinement quality may also be determined, such as the R_p-factor, given by

$$R_p \frac{\sum \left(I^{obs} - I^{calc}\right)}{\sum I^{obs}} \times 100\% \tag{5.18}$$

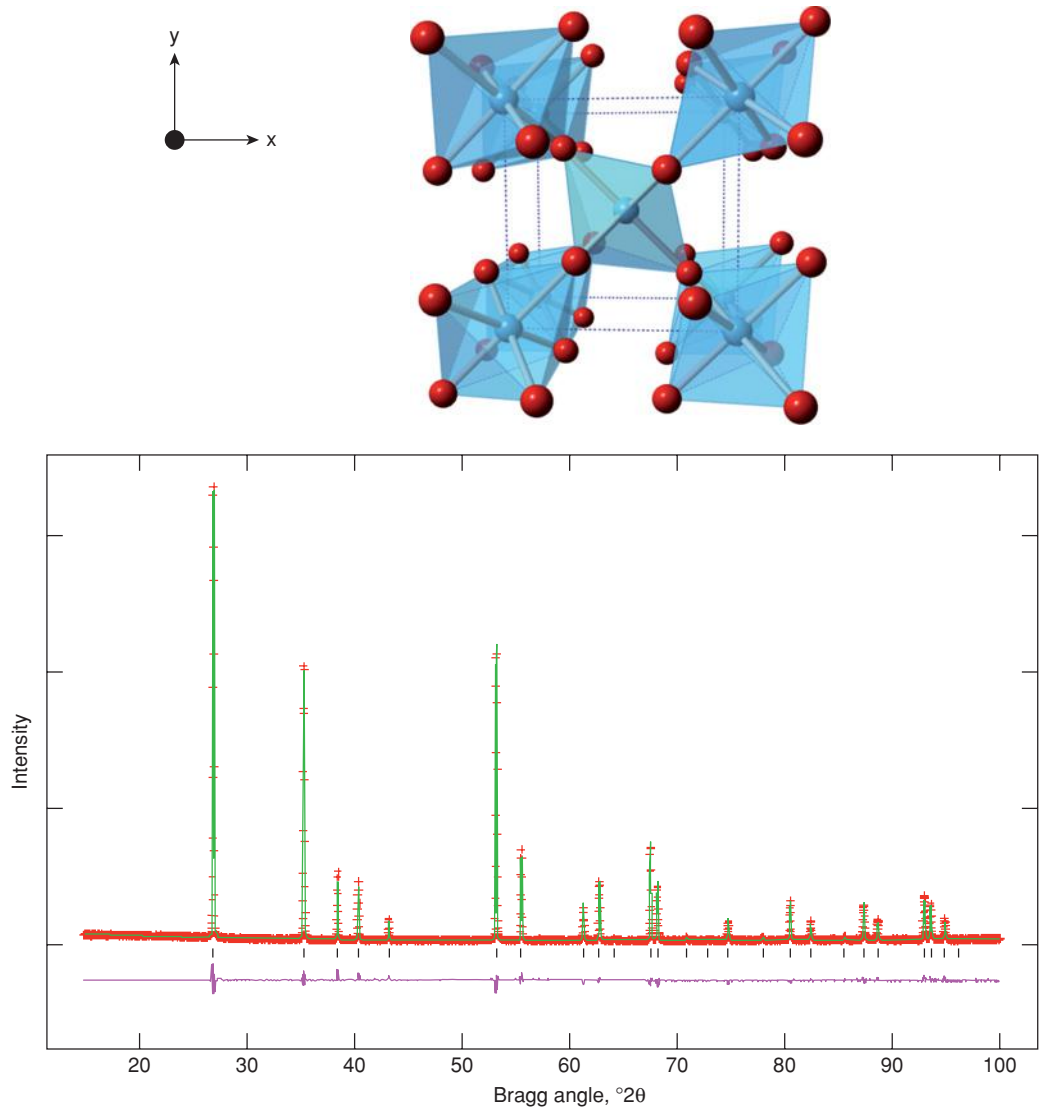
For this, intensity values are measured and calculated at regular intervals of, e.g., 0.02° 2θ. The objective in refinement is therefore to minimise the value of R_p. It is usual practice to show graphically the quality of a refinement by presenting experimental and calculated profiles together with a difference profile and markers showing the positions of the expected Bragg peaks. A typical example is shown in Fig. 5.23; the experimental and calculated diffraction profiles match well, as indicated by the difference profile shown, which has only small residual peaks.

5.3.8 X-ray crystallography and structure determination – what is involved?

For determining the structures of crystalline materials, X-ray crystallography reigns supreme. For molecular materials, it complements the use of spectroscopic techniques – NMR, mass spectrometry, etc. – and often one may use either crystallography or spectroscopy to determine molecular structures. For non-molecular materials, or for molecular materials whose arrangement in the crystalline state is important, or whose bond lengths and angles must be determined, then X-ray crystallography is by far the most important technique for structure determination.

Nowadays, solving structures is mathematically complex but usually highly automated, with computer-controlled collection and processing of diffraction data. It requires large, expensive instruments to do the diffraction experiments: given a reasonable-sized single crystal, $^1/_2$–2 days are often required to collect the data and a few hours of computation to solve the structure. It is not yet the kind of experiment that is carried out in an afternoon in a teaching laboratory, but nevertheless, it is important to have at least a passing familiarity with the processes and problems involved. We have already covered much of the groundwork earlier in this chapter.

Solving an unknown structure is rather like solving a set of simultaneous equations. The unknowns are the atomic coordinates and the equations are the experimental intensity data. Obviously, we must have at least as many equations (i.e. intensities) as variables, but in practice, we need many more intensities than variables to obtain good-quality structure determinations. This is partly because the intensity data may have errors, partly because the computational methods involve statistical analyses of data which function properly only with large data sets, and partly because the Fourier methods used in structure determination are effective only with a large



a = 4.67791(3), c = 3.02234(2) Å
Space group $P4_2/mnm$ (No 136)
Cations, site 2a: 000;½½½; U_{iso} = 0.0080 (3) Å^2
Oxygens, site 4f: x,x,0; $\bar{x}$, $\bar{x}$,0; $\bar{x}$+½,x+½,½; x+½,$\bar{x}$+½,½ with x = 0.03035(3); U_{iso} = 0.0119(7) Å^2
χ^2: 6.575, R_{wp}: 7.27%, R_p: 5.58%,
Cation-oxygen bond lengths : 1.994 (1) x 4 and 2.008(2) x 2 Å

Figure 5.23 ***Calculated (red crosses), observed (green curves) and difference (purple) profiles, crystal structure and crystallographic data for rutile-structured $Ti_{1-3x}Cu_xNb_{2x}O_2$: x = 0.02 refined from powder XRD data. Adapted from N. Reeves-McLaren* et al.*, J. Solid State Chem.,* 184*, 1813, © 2011 Elsevier. Reproduced under the terms of the STM Agreement.***

number of coefficients (intensities). Let us consider some examples with structures of varying complexity, because the number of variables to be determined dictates the type of diffraction experiment that is required.

First, suppose that we have material, MX_2, suspected to have the rutile structure; this was the case for the material that forms the basis of Fig. 5.23; how many variables must be refined in order to confirm the structure? From the rutile structure, Fig. 1.37, atoms M are in fixed positions, at the corner and body centre, and hence have no positional variables. The atoms X have a single positional variable, x, since the coordinates of the four atoms in the unit cell are given by

$$x, x, 0; 1-x, 1-x, 0; \tfrac{1}{2}-x, \tfrac{1}{2}+x, \tfrac{1}{2}; \tfrac{1}{2}+x, \tfrac{1}{2}-x, \tfrac{1}{2}$$

Hence a *trial structure* may be proposed in which x is given a typical value for rutile of 0.3.

The main structural variables to be refined in order to compare the calculated structure factors, F_{hkl}^{calc}, with the observed values, F_{hkl}^{obs}, are the x coordinate for oxygen and the B (or U_{iso}) values for Ti and O. In order to determine these accurately, intensities of at least 10–20 reflections would normally be required; since a powder X-ray pattern may contain at least this number of lines, the structure could be determined satisfactorily from powder data. In this case, the final refined values of the lattice parameters, x value, U_{iso} and statistical parameters are as shown in Fig. 5.23 for a phase which has the rutile structure but with partial substitution of Ti by a mixture of Cu and Nb.

Second, a moderately complex structure is that of $YBa_2Cu_3O_x$, the so-called YBCO or Y123 ceramic superconductor. It is orthorhombic with a structure related to but more complex than that of perovskite. The unit cell contains one formula unit with $x = 7.0$. There are five positional variables to be determined to locate all the atoms and eight B_{iso} values (four for oxygen, since there are four crystallographically distinct oxygens in the structure, two for Cu and one each for Ba and Y). One of the oxygens is present in variable amounts (x is variable in the formula) and the fractional occupancy of its site is a variable. This gives a total of 14 variables and for a good structure determination, 200–300 intensities would be required. Routine powder methods do not give this number of well-resolved reflections; either special powder techniques, such as high-resolution neutron diffraction or synchrotron X-ray diffraction, or single-crystal methods must be used. In these, the data set is increased both by collecting data to lower d-spacings using shorter wavelength radiation and, with single crystals, by recording data that are too weak to appear in powder patterns.

Third, for yet more complex structures, such as many silicates and complex organic molecules, especially if they are of low symmetry, the number of variables may be 50–100 and single-crystal data, giving perhaps 2000–3000 intensities, are usually essential.

Thus far, we have not considered the *ab initio* determination of structures for which we have no trial structure in mind, but have simply considered the number of variables that need to be determined in order to specify the crystal structure. Let us now consider the steps involved when we do not have a trial structure available. Once the intensity data have been collected, corrected for factors such as Lorentz polarisation and converted to observed structure factors, the process of solving the structure can begin. The problem is essentially to determine the values of the atomic coordinates x, y, z that, when substituted into equation (5.12), yield calculated F values that match the observed ones.

In cases where both the signs (+ or –) and magnitudes of the F^{obs} values are known, there are standard mathematical procedures, based on Fourier series, for attacking this problem. The observed diffracted X-ray beams may be combined, as a Fourier transform, to give the crystal structure in the form of an electron density map (as in Fig. 5.24). The relevant equation is

$$\rho(u, v, w) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} \cos 2\pi(hu + kv + lw) \tag{5.19}$$

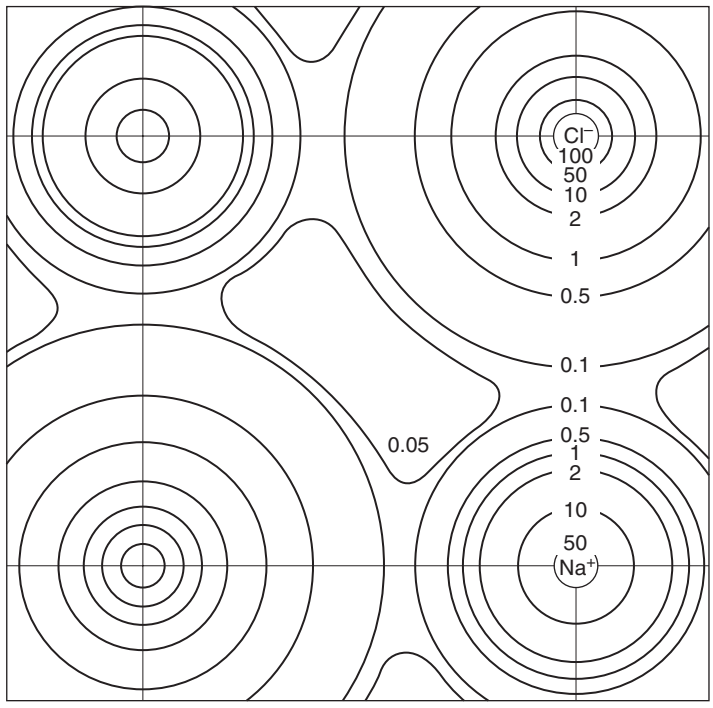


Figure 5.24 ***Electron density map for NaCl.***

from which, for any point u, v, w in the unit cell, the electron density, ρ, is calculated by substituting into the summation all the observed F values for each h, k, l. In order to obtain a good-quality electron density map, there must be a large number of terms in the summation and this is the main reason why the number of intensity data required for a structure determination is much greater than the number of variables.

There is a clear analogy between this mathematical addition and the production of images in optical and electron microscopes. In microscopes, the diffraction pattern obtained by shining a beam of either light or electrons on the sample is combined to give an image using either an optical lens or an electromagnetic lens. There is, unfortunately, no material that can act as an 'X- ray lens' and we must resort to mathematical methods to perform this combination. This brings us to the key problem in crystallography: while the amplitudes of the F^{obs} values (written as $|F^{obs}|$ and spoken as mod F^{obs}) are calculated directly and unequivocally from the intensities, their phases are not. Thus, intensity, I, values must be positive, but F values, given by $\sqrt{I}$, may be positive or negative. Many means have been devised to attack this '*phase problem*'. The most successful ones are outlined next.

5.3.8.1 The Patterson method

This method uses a Fourier summation similar to equation (5.19) but in which intensity (or F^2_{hkl}) data form the *coefficients*, i.e.

$$\rho(u, v, w) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}|^2 \cos 2\pi(hu + kv + lw) \tag{5.20}$$

The resulting Patterson map looks similar to a Fourier map, but the regions of high electron density correspond to vectors between pairs of atoms. The peak heights are proportional to the products of the atomic numbers, hence vectors formed from pairs of the heaviest atoms in the unit cell give rise to the largest peaks. The peak positions give the separation, vectorially, of these two atoms in the structure. Thus, two atoms at x_1, y_1, z_1 and x_2, y_2, z_2 will give a Patterson peak, relative to the origin of the Patterson map, at $x_1 - x_2$, $y_1 - y_2$, $z_1 - z_2$. This does not give the positions of heavier atoms directly, but it does give their relative positions in the unit cell and is often of great assistance in getting started with structure determination. Thus, if one atom can be located with confidence, the Patterson map may suggest the location of others. Once such a start has been made, other Fourier methods may be used to complete the structure determination.

5.3.8.2 Fourier methods

These may be used even though the complete list of F^{obs} values, with their signs, is not available. Use is made of composite F values, whose magnitudes are taken from the F^{obs} listing and whose signs are taken from a partial structure factor calculation based on the heavy atom positions only. Since the heavy atoms scatter X-rays most strongly, they are likely to dominate the intensities and, in particular, control the phases (i.e. whether + or –). Hence most of the signs can probably be determined correctly at this stage, especially for the larger F values. A Fourier map, made from equation (5.19) and these composite F values, should then reveal considerably more electron density peaks and enable lighter atoms to be located. This process may be repeated: a more accurate structure factor calculation is performed using the more extensive list of atomic coordinates obtained from the Fourier map; consequently, the signs of more F values are determined correctly and an improved Fourier map can be calculated.

Once most of the atoms have been located approximately, *least-squares refinement* procedures may be used to improve the agreement between F^{obs} and F^{calc} and thereby minimise the R-factor, equation (5.17). In these procedures, atom positions are permitted to vary somewhat, the effect on R is noted and the optimum positions (for minimum R) are found by trial and error. It is also useful to construct a *difference Fourier map* using $F^{obs} - F^{calc}$ values as the coefficients:

$$\Delta(u, v, w) = \frac{1}{V} \sum_h \sum_k \sum_l (F^{obs} - F^{calc}) \cos 2\pi (hu + kv + lw) \tag{5.21}$$

A difference map may show regions of low, positive electron density associated with previously undetected light atoms such as H; or it may show regions of negative electron density, indicating that an atom has been wrongly placed at that position; or it may be completely featureless, indicating that the structure refinement is correct and complete.

5.3.8.3 Direct methods

These are very useful for structure determination and work best when all the component atoms have similar atomic number, thus complementing the Patterson method, which works best when a small number of heavy atoms are present. Direct methods are used to determine phases. They are based on the statistical probabilities of phases being either + or –. For example, for centrosymmetric structures, the *Sayre probability relationship* indicates that for three reflections h, k, l, h', k', l' and h'', k'', l'' that are related by $h'' = h - h'$, $k'' = k - k'$ and $l'' = l - l'$, then the sign of one phase is likely to be the same as the product of the signs of the other two. Thus, if 312 and 111 are both –, then 201 is probably +. By choosing three reflections, whose phases are not known, there are eight possible combinations of + and –. For each combination, the phases of other

reflections may be predicted. A range of methods are available for predicting phases and optimising the value of the predictions. The end result is an *E-map*, similar to an electron density map and from which atomic positions may be determined.

5.3.8.4 Electron density maps

An electron density map shows the variation of electron density throughout the unit cell. When solving an unknown structure, it is often useful to construct electron density maps (Fourier maps) in order to try to locate atoms. As the structure refinement proceeds, the quality of the electron density map usually improves: the background electron density decreases and more peaks due to individual atoms become resolved. The electron density map obtained at the end of a structure determination is an important piece of information.

Electron density maps take the form of sections through the structure at regular intervals; by superposing these, a 3D picture of the electron density distribution may be obtained. In Fig. 5.24 is shown the electron density distribution for a section through a very simple structure, NaCl. This is essentially a contour map showing lines of equal electron density by interpolation from the calculated electron density values at each point on a grid superposed on the unit cell. The section shown is parallel to one unit cell face and passes through the centres of the Na^+, Cl^- ions. It has the following features.

An electron density map resembles a geographical contour map. The contours are lines of constant electron density throughout the structure. Electron density maxima correspond to atoms; the coordinates of the atoms in the unit cell are given by the coordinates of the peak maxima. The peak height is proportional to the number of electrons possessed by that atom, which apart from very light atoms equals the atomic number of that atom. In Fig. 5.24, peaks of relative height 100 and 50 are seen and are assigned to Cl and Na, respectively. (The atomic numbers of Cl and Na are 17 and 11; for ions, the number of electrons are 18 and 10. The experimental maxima are therefore in fair agreement with the expected values.)

Electron density maps also show that our mental picture of atoms as spheres is essentially correct, at least on a time average. The electron density drops to almost zero at some point along the line connecting pairs of adjacent atoms in Fig. 5.24 and this supports the model of ionic bonding in NaCl. In structures which have covalent bonding, there is residual electron density between atoms on the electron density map. However, in other than very simple structures, such as the alkali halides, it is difficult to use electron density maps to determine quantitatively the distribution of valence electrons.

In most structure refinements, both the position and temperature factors (or thermal vibration factors) of atoms are varied in order to achieve the best agreement between measured and calculated structure factors. In refining simple structures, the atomic coordinates are usually known accurately, but in more complex structures, various kinds of local structure distortions are possible which cannot be determined individually but are averaged in the final refinement data. The temperature factors allow for both local structure distortions, involving small atomic displacements, and also temperature-induced atomic vibrations.

5.4 Electron Diffraction

Electrons have wave characteristics which allow them to be used for diffraction experiments. Their wavelength is related to velocity, which is governed by the voltage through which they are accelerated in an ‘electron gun’ and is usually about 0.04 Å in conventional electron microscopes. The two techniques of *electron diffraction* and *electron microscopy* are closely related. The latter is used more widely and is discussed in Chapter 6; here we shall see briefly the general characteristics of electron diffraction and how it compares with X-ray diffraction.

Electrons interact strongly with matter and intense diffraction patterns are obtained from very small samples; indeed, for transmission studies, the samples should not be thicker than about 100 nm (1 nm = 10 Å), otherwise the electrons may be unable to pass through the sample. This contrasts markedly with X-ray diffraction: the efficiency with which X-rays are diffracted by matter is low and, for single-crystal studies, relatively large crystals, 0.05 mm or greater in dimensions, are usually required.

One disadvantage of electron diffraction is that *secondary diffraction* commonly occurs. Because the scattering efficiency of electrons is high, the diffracted beams are strong. Secondary diffraction occurs when these diffracted beams effectively become the incident beam and are diffracted by another set of lattice planes. There are two undesirable consequences of secondary diffraction. First, under certain circumstances, extra reflections may appear in the diffraction pattern; care is therefore needed in their interpretation. Second, the intensities of diffracted beams are unreliable and cannot be used quantitatively for structure determination.

In spite of these disadvantages, electron diffraction is very useful and complements the various X-ray techniques. Thus, with X-rays, the scattering efficiency is small, secondary diffraction is rarely a problem and intensities are reliable, but (relatively) large samples are needed. With electrons, the scattering efficiency is high and intensities are unreliable, but very small samples can be studied. The technique is very useful for obtaining unit cell and space group information for crystals smaller than 0.01–0.02 mm in their lateral dimensions; it is, in fact, the only reliable method for obtaining such information since it gives a direct picture in 3D (in reality, the pictures are 2D sections through 3D space) of the directions of diffracted beams. From such information, the diffracted beams can be assigned h, k, l indices and the unit cell determined. From the pattern of systematic absences, further information on the lattice type and space symmetry, leading to determination of the space group, can be obtained.

In the absence of direct unit cell and space group information, powder XRD patterns may be indexed using various trial-and-error procedures but there can be uncertainty in the results obtained, especially in cases where a superstructure may be present that shows up only weakly in the powder XRD pattern. In spite of its utility, electron diffraction is underused in solid state chemistry for unit cell and space group determination, probably because it is time consuming and requires a high level of expertise for data interpretation. It is, however, of limited use for full structure determination because of the unreliability of diffraction intensities.

Electron diffraction is unsuitable as a routine method of phase identification in relatively large (e.g. 10 mg or more) samples because of sampling issues. The technique focuses on individual grains, one at a time, rather than giving an overall average analysis of the sample. It is useful, however, (a) when very small quantities are available, (b) for thin-film samples and (c) for detecting small amounts of impurity phases. In all these cases, there would be insufficient material to show up in XRD analysis.

5.5 Neutron Diffraction

Neutron diffraction is a very expensive technique. In order to get a sufficiently intense source of neutrons, a nuclear reactor is needed. Few laboratories have their own neutron facility and, instead, experiments are carried out at central laboratories which provide a user service (e.g. at the ILL, Grenoble, France, the Rutherford-Appleton Laboratory, UK, and the Argonne Laboratory, USA). In spite of its high cost, neutron diffraction is a valuable technique and can provide information, especially on magnetic materials, that is not attainable with other techniques. Clearly, it is never used when alternative techniques, such as XRD, can solve a particular problem.

Neutron beams are usually of low intensity and therefore, for diffraction experiments, the sample size required is relatively large, at least 1 mm^3. Since crystals of this size are often not available, crystallographic studies are usually carried out on polycrystalline samples. A powder neutron diffraction pattern looks very much like an X-ray pattern.

There are several characteristic differences between neutron and X-ray diffraction. First, the neutrons obtained from a nuclear reactor give a continuous spectrum of radiation, without the intense characteristic peaks that are present in X-ray spectra (as in Fig. 5.3b). The neutrons that are used for diffraction (so-called thermal neutrons whose velocity and therefore wavelength are controlled by the temperature inside the nuclear reactor) have wavelengths of the order 0.5–3 Å. In order to have monochromatic neutrons for conventional diffraction experiments, it is necessary to select a particular wavelength and filter out the remainder using a crystal monochromator. Most of the available neutron energy is wasted, therefore, and the beam that is used is weak and not particularly monochromatic.

A recent exciting advance uses pulsed neutron sources coupled with *time-of-flight* analysis. The neutrons are obtained using particle accelerators to bombard a heavy metal target with high-energy particles such as protons. The efficiency of the *spallation* process, yielding about 30 neutrons per proton, gives a high neutron flux suitable for diffraction experiments. In the time-of-flight method, the entire neutron spectrum (variable wavelength) is used with a fixed diffraction angle, θ. The neutron wavelength depends on velocity, given by the de Broglie relation, $\lambda = h/mv$, where m is the mass of the neutron, 1.675×10^{-27} kg. Hence the diffracted radiation arriving at the detector is separated according to its time of flight and wavelength. The fundamental law of diffraction is again Bragg's law, $n\lambda = 2d\sin\theta$. In the time-of-flight method, λ and d (the d-spacing) are the variables at fixed θ. This compares with conventional diffraction techniques in which d and θ are the variables at fixed λ. The pulse method gives rapid data collection. It may therefore also be used for studies of short time phenomena, especially in experiments where samples are subjected to pulsed magnetic fields.

A second difference between neutron and X-ray diffraction is that the scattering powers of atoms towards neutrons are completely different from those towards X-rays. In the latter, scattering power is a simple function of atomic number and light atoms such as hydrogen diffract X-rays only weakly. With neutrons, the atomic nuclei, rather than the extranuclear electrons, are responsible for the scattering, and in fact hydrogen is a strong scatterer of neutrons. In addition, since atomic nuclei are effectively point scatterers of neutrons there is no dependence of neutron scattering power (the so-called neutron scattering length) on $\sin\theta/\lambda$, unlike the form factors for X-ray scattering. This means that the intensity of powder neutron diffraction patterns does not show the general decrease with increasing Bragg angle that is seen with powder XRD patterns.

There is no simple dependence of neutron scattering power on atomic number. Also, there is not such a wide range of scattering powers of neutrons across the Periodic Table as there is for X-rays. This means that heavy atoms in crystal structures dominate XRD patterns and therefore form the basis of methods such as the Patterson method for structure determination, but can create difficulties in structure refinement using XRD data if light atoms are present or atoms of a similar atomic number that cannot be readily distinguished by XRD. Neutron diffraction, by contrast, is less useful for structure determination because of the absence of particularly strong neutron scatterers but is good for structure refinement using the Rietveld method because almost all atoms contribute significantly to neutron scattering. Additionally, some atoms cause a change of phase of π (or $\lambda/2$) in the diffracted neutron beam and others, notably vanadium, do not diffract neutrons at all. For this reason, vanadium cans are used as sample holders in diffraction experiments because they are transparent to neutrons.

5.5.1 Crystal structure determination

Neutron diffraction may be used for crystallographic work in cases where XRD is inadequate. It has been much used to locate light atoms, especially hydrogen in hydrides, hydrates and organic structures. Usually, the main part of the structure is solved by X-ray methods and neutron diffraction is used to locate the light atoms. Neutron diffraction is also used to distinguish atoms that have similar X-ray scattering powers, such as Mn, Fe, Co and Ni. The neutron scattering powers of these atoms are different and, for instance, superlattice phenomena, associated with Mn/Fe ordering in alloys, are readily observed by neutron diffraction.

5.5.2 Magnetic structure analysis

Magnetic properties depend on the presence of unpaired electrons, especially in *d* or *f* orbitals. Since neutrons possess a magnetic dipole moment, they interact with unpaired electrons and, consequently, are diffracted by both atomic nuclei and unpaired electrons. This forms the basis of a powerful technique for studying the magnetic structure of materials. A simple example of magnetic structure and order is shown by NiO. By XRD, NiO has the *fcc* rock salt structure. When examined by neutron diffraction, however, extra peaks are observed which indicate the presence of a superstructure. This arises because the unpaired *d* electrons (in the e_g orbitals) are arranged so as to be antiparallel in alternate layers of nickel atoms, Fig. 5.25. Neutrons detect this ordering of spins whereas X-rays do not. The unit cell of 'antiferromagnetic' NiO, which is stable below 250 °C, has eight times the volume of paramagnetic NiO, stable above 250 °C. The structure also shows a slight rhombohedral distortion, involving a contraction in the [111] direction perpendicular to the planes of ordered Ni^{2+} ions. This distortion is very small and is detected by the splitting of reflections such as [111] and [$\bar{1}$11] in high-resolution neutron powder diffraction patterns. For the present purposes, we may ignore this distortion and treat the structure, geometrically, as cubic.

Neutron powder diffraction patterns for MnO (which behaves very similarly to NiO) below and above the Néel temperature, T_N, together with a schematic powder XRD pattern at room temperature, are shown in Fig. 5.26. Comparison of the two patterns above T_N (b and c) shows that lines appear in the same positions but are of very different intensities. In the rock salt structure, the condition for reflection is that *h*, *k*, *l* should be either all odd or all even. Hence the first four lines to be expected in the powder pattern are 111, 200, 220 and 311. All four lines appear in both patterns but 200 and 220 are weak in the neutron pattern (b). The small intensity of 200 and 220 in (b) is largely because the neutron scattering powers of Mn^{2+} and O^{2-} are opposite in sign although slightly different in magnitude. Partial cancellation therefore occurs for the 200 and 220 reflections since Mn^{2+} and O^{2-} ions on the same planes scatter out-of-phase with each other. This is precisely the opposite of the case with scattering of X-rays, for which the scattering factors of all elements have the same sign and therefore for the 200 and 220 reflections, Mn^{2+} and O^{2-} scatter in-phase with each other.

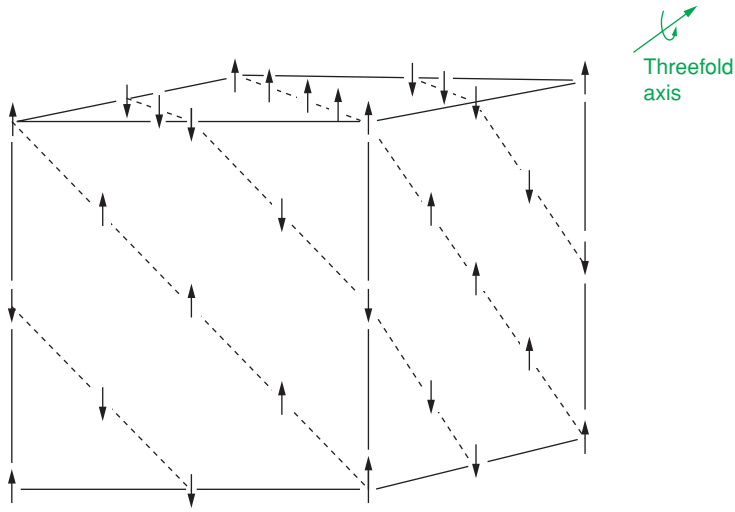


Figure 5.25 ***Antiferromagnetic superstructure in MnO, FeO and NiO, showing pseudocubic unit cell for which a (supercell) = 2a (subcell). Oxygen positions are not shown.***

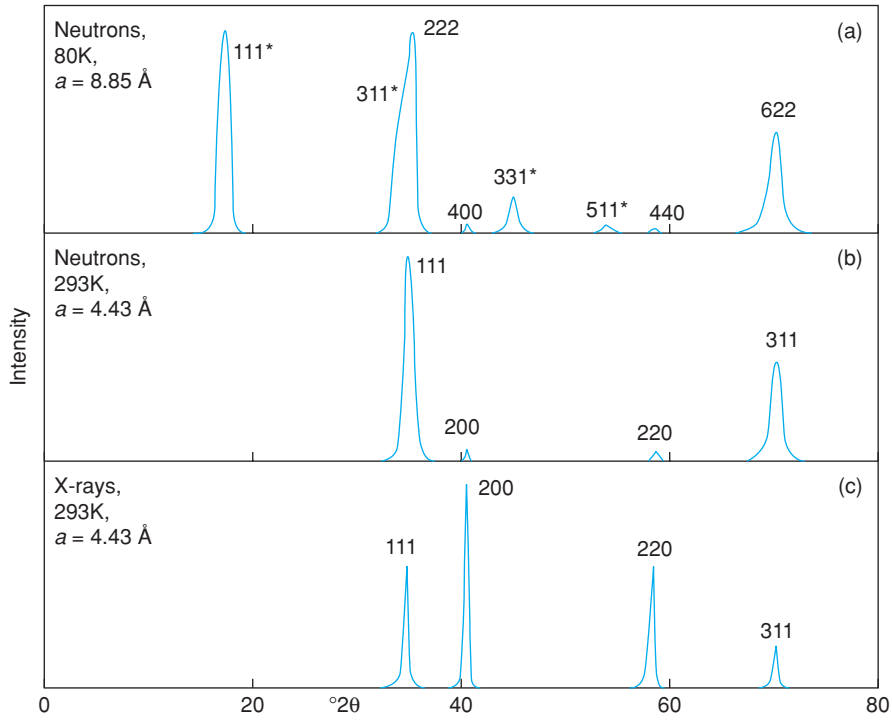


Figure 5.26 ***Schematic neutron and powder XRD patterns for MnO for λ = 1.542 Å. Peaks are assigned Miller indices for the cubic unit cells given. Neutron data adapted from Shull, Strauser and Wollan,*** **Phys. Rev.,** 83, ***333, © 1951 American Physical Society.***

Comparison of Fig. 5.26(a) and (b) shows that below T_N, extra lines (asterisked) appear in the neutron diffraction pattern. These extra lines are associated with the antiferromagnetic superstructure. Although, as mentioned above, the true symmetry of the antiferromagnetic structure is rhombohedral, to a first approximation it can be treated as cubic with cell dimensions that are twice the value for the high-temperature paramagnetic structure, i.e. $a = 8.85$ Å at 80 K ($<T_N$), whereas at 293 K ($>T_N$), $a = 4.43$ Å. The volumes of the unit cells are therefore in the ratio 8:1. The extra lines in the powder pattern of the antiferromagnetic structure may be indexed as shown; observed reflexions are those for which h, k and l are odd. Note that, because the unit cell for (a) is twice that for (b), the reflexions that are common to both have doubled indices in (a) compared to those in (b). For example, reflexion (200) in (b) has indices (400) in (a).

5.5.3 Inelastic scattering, soft modes and phase transitions

'Slow' neutrons possess kinetic energy that is comparable to the thermal energy levels in a solid. Such neutrons are inelastically scattered by phonons (i.e. vibrational modes) in the solid. From an analysis of the energy of the scattered neutrons, information on phonons and interatomic forces is obtained. For magnetic materials, further information on their electron exchange energy is obtained. These are spectroscopic techniques using incoherent neutrons rather than diffraction techniques which use coherently scattered neutrons, but are mentioned here, briefly, for convenience.

Displacive phase transitions are believed to be associated with the instability of a lattice vibration. A certain type of vibrational mode, known as a *soft mode*, in the low-temperature structure effectively collapses at the critical temperature. Soft modes may be studied by IR and Raman spectroscopy, provided that the vibrations involved are spectroscopically active, and also by neutron scattering. The latter technique is useful since it is not limited by the spectroscopic selection rules; by measuring the inelastic scattering about a number of 'Bragg reflections', the atomic displacements that are responsible for the soft mode and the phase transition may be determined. For instance, the mechanism of the displacive transition in quartz, SiO_2, at 573 °C has been analysed by recording the neutron spectra at several temperatures below and above 573 °C.

6

Other Techniques: Microscopy, Spectroscopy, Thermal Analysis

A wide variety of spectroscopic, microscopic and other techniques, in addition to the diffraction techniques described in Chapter 5, may be used to obtain information on the identity, structure and characterisation of solids. The main objective of this chapter is to give a brief summary of some of these techniques and especially to indicate the kind of information that they can give about solids. There is a considerable amount of specialised terminology; such words and phrases are given in italics; explanations are not usually given, but the meanings should be clear from the text.

First, each technique, or group of techniques, is considered with a brief summary of how the technique operates and the kind of information that can be obtained by using it. We then change the focus to ask the question, 'I have an unknown solid; what techniques should I use to analyse, identify and characterise it?'.

6.1 Diffraction and Microscopic Techniques: What Do They Have in Common?

These two groups of techniques operate with the same basic sequence of events:

$$\text{source} \xrightarrow{\text{radiation}} \text{sample} \xrightarrow{\text{(coherent) scattering}} \text{diffraction pattern} \xrightarrow{\text{recombination}} \text{image}$$

In principle, each source of radiation produces a diffraction pattern by a process of coherent scattering and interference; there is essentially no loss of energy and the scattered radiation has the same wavelength as the incident radiation. The rays which form the diffraction patterns may subsequently be recombined to give an image of the object. However, electrons are the only radiation which is used routinely to obtain both diffraction patterns and images. Hence X-rays and neutrons are used to obtain diffraction patterns but not images, whereas optical microscopes are used to obtain images but not diffraction patterns.

The formation of images from diffraction patterns requires a suitable lens. In the case of X-rays, there are currently no materials available from which a suitable lens can be constructed and, therefore, with X-rays we are limited to recording diffraction patterns. Similarly with neutrons, only the diffraction patterns can be obtained. It is possible, however, to obtain images mathematically from X-ray diffraction patterns

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

since, if the amplitudes and phases of diffracted beams are known, these can be combined using Fourier series to give electron density maps. By contrast, with visible light it is not feasible, for various reasons, to record diffraction patterns and, instead, optical microscopes are designed with suitable lens systems to give magnified images. Diffraction techniques using X-rays, electrons and neutrons are discussed in Chapter 5. In this chapter, Section 6.2, we look at microscopy techniques utilising both visible light and electrons.

There is also a close link between microscopic and many spectroscopic techniques involving electrons, especially when the electrons interact inelastically, or incoherently, with samples. In these, the incident electrons lose energy in the collisions, which is released as either secondary electrons or radiation. Numerous imaging and spectroscopic techniques make use of these emissions; those that are based on electron microscope instrumentation are also in considered in Section 6.2.

6.2 Optical and Electron Microscopy Techniques

6.2.1 Optical microscopy

As a first step in examining a solid, it is usually well worthwhile to have a look at it under magnification, especially if the particles are micron sized or larger. This may involve no more than a brief look with a polarising microscope. Materials that visually appear similar, such as fine powders of white sand or table salt, may look completely different under the microscope. Thus, in these two examples, crystals have different morphology (i.e. shape) and their optical properties in plane polarised light are distinct.

With optical microscopes, particles down to a few micrometres or microns (1 μm = 10^4 Å = 10^{-3} mm) in diameter may be seen under high magnification. The lower limit is reached when the particle size approaches the wavelength of visible light, 0.4–0.7 μm. For submicron-sized particles, it is essential to use electron microscopy. With this technique, features as small as a few ångströms across can be imaged readily. Thus, nanoparticles cannot be studied by optical microscopy because, at best, they form a blurred image, although it would be possible to confirm that particles are submicron in size and that larger particles are not present.

A simple magnifying optical microscope is shown in Fig. 6.1(a). The *resolving power* of optical (and electron) microscopes is the smallest separation of two objects which may be reproduced clearly in the image and is given by the equation

$$\text{resolving power} = \frac{0.61\lambda}{n \sin \alpha} \tag{6.1}$$

where n is the refractive index of the medium in which the object is immersed and α is the maximum half-angle subtended at the objective lens by the object, Fig. 6.1(a). The product $n\sin\alpha$ is the *numerical aperture* of the objective lens. The main variable in equation (6.1) is the wavelength of the radiation used, λ, which shows why electron microscopes, with electron wavelengths $\ll$1 Å, have resolving powers that are orders of magnitude greater than those of optical microscopes.

In examining a solid, there is a logical sequence of steps, with increasing magnification at each stage, to go through. First, inspection with the naked eye gives resolution at the submillimetre level. Second, examination with an optical microscope enables particles and features down to the micron level to be seen. Third, scanning electron microscopy gives resolution down to the nanometre scale. Finally, transmission electron microscopy can give resolution down to the ångström level, with the possibility of imaging individual atoms.

In examining an ‘unknown’ solid, it is important to follow this sequence in order to get a clear picture of the material across different length scales. It may not seem very exciting to use an optical microscope when electron microscopes are available, but a great deal of intermediate-scale information can be lost if one goes directly from visual examination with the naked eye to electron microscopy.

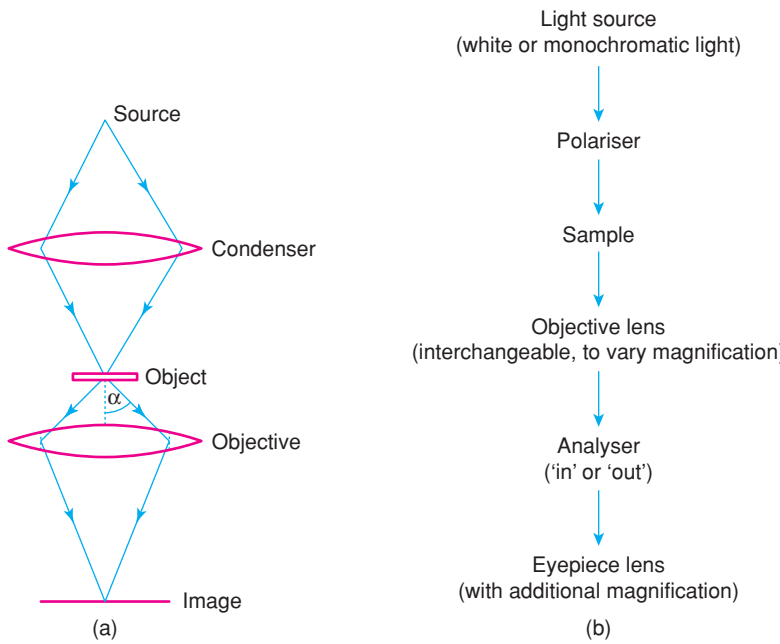


Figure 6.1 ***(a) Simplified ray diagram for an optical microscope. (b) Basic components of a polarising microscope.***

Both optical and electron microscopy use two types of instrument, where the sample can be viewed in transmission, Fig. 6.1(a) (i.e. the beam of light or electrons passes through the sample) or in reflection (the beam of light or electrons is reflected off the sample surface). The *polarising* or *petrographic microscope* is a transmission instrument, widely used by geologists and mineralogists, and can be used profitably in solid state chemistry. Samples are usually either fine powders or thin slices cut off a solid piece. On this scale, substances are often transparent whereas they would be opaque in bulk form. The sample is immersed in a liquid whose refractive index is close to that of the sample. If this is not done and the solid is examined in air, much of the light is scattered off the sample surface rather than transmitted through it. It is then difficult, if not impossible, to measure the various optical properties of the solid. The *metallurgical* or *reflected light microscope* is suitable for looking at the surfaces of materials, especially opaque samples, and is much used in metallurgy, mineralogy and ceramics.

6.2.1.1 Polarising microscope

The basic components of a polarising microscope are indicated in Fig. 6.1(b). The source may give either white or monochromatic light. This shines on the *polariser* and only that component whose vibration direction is parallel to that of the polariser is permitted to pass through. The resulting *plane polarised light* passes through lenses, apertures and accessories and is focused on to the sample, which is mounted on the *microscope stage*. Light transmitted by the sample and immersion liquid is picked up by the *objective lens*. The instrument usually has several of these, of different magnification, which are readily interchangeable. The *analyser* may be placed in or out of the path of light. It is similar to the polariser but is oriented so that its vibration direction is at 90° to that of the polariser. When the analyser is 'in', only light vibrating in the correct direction is permitted to pass through and on to the *eyepiece*. When the analyser is 'out', the microscope behaves as a simple magnifying microscope. Various other attachments may also be present between the sample and eyepiece.

In practice, samples are usually examined with the analyser alternately 'out' and 'in'. With the analyser 'out', the instrument can be focused and a first examination made, noting the size and shape of particles. With the analyser 'in', the sample is viewed between *crossed polars*, and one can immediately tell whether the sample is *isotropic* (dark), as is the case for cubic crystals or amorphous solids such as gels and glasses, or *anisotropic* (bright or coloured), in the case of non-cubic crystals. By rotating the sample and stage (in some instruments, the polars rotate and the stage is fixed), the *extinction directions* can be seen if the crystals are anisotropic; from the nature of the extinction, information on the quality of the crystals may be obtained. If it is desired to measure the *refractive index* (or indices) of the sample, the analyser is again taken out and various immersion liquids, of different refractive index, are tried until one is found in which the sample is effectively invisible. The refractive index of the sample then matches that of the liquid. It is possible to do this systematically; using the *Becke line method*, it is possible to tell whether the refractive index of the sample is higher or lower than that of the immersion liquid.

The above measurements are simple and rapid. If necessary, further measurements can be made. Thus, the variation of refractive index with crystal orientation may be studied using the concept of the *optical indicatrix*. Anisotropic crystals may be classed into *uniaxial* or *biaxial*, depending on whether they have one or two *optic axes*. When viewed down an optic axis, anisotropic crystals appear to be isotropic, i.e. they are dark between crossed polars. If a convergent beam of light is shone onto the sample, a *conoscopic examination* may be made and *interference figures* seen; from these, further information on uniaxial and biaxial crystals is obtained.

Before the advent of X-ray diffraction (∼1910), morphological data provided an important means of classifying crystals. The method uses a goniometric stage on which a crystal is mounted. The crystal can be rotated and viewed from all angles under the microscope. From the shape of the crystal, information on the internal symmetry of the crystal structure may be obtained. With well-formed crystals, the number and disposition of the crystal faces may indicate the *class* (i.e. *point group*) of the crystal. Although nowadays *goniometry* is little used for crystallographic work, it is possible to obtain useful information from a cursory examination of powdered samples using the polarising microscope. For instance:

1. Crystal fragments often have a characteristic shape, especially if the crystals cleave easily and in a particular orientation.
2. It is possible to classify substances into *isotropic* and *anisotropic*. Isotropic crystalline substances are of cubic symmetry and appear 'dark' between crossed polars since plane-polarised light passes through them without modification. Anisotropic substances cause a partial rotation of the plane of polarisation of light as it passes through them. The emergent beam has a component that is vibrating parallel to the vibration direction of the analyser and hence anisotropic substances appear 'light'. Anisotropic substances include all non-cubic crystalline solids.

   It is often possible to distinguish between uniaxial crystals (of hexagonal, trigonal or tetragonal symmetry) and biaxial crystals (orthorhombic, monoclinic and triclinic symmetry). In uniaxial crystals, the optic axis is parallel to the unique symmetry axis (six-, three- or fourfold, respectively) and when viewed down this axis such crystals appear to be isotropic. In a powder, it is fairly common to find crystals in this orientation, especially if they exhibit *preferred orientation*. For instance, crystals of hexagonal symmetry often exist as thin hexagonal plates and these naturally tend to lie on their flat faces. Although biaxial crystals have two optic axes, these are not usually parallel to any pronounced edge or feature of the crystal. The chance of finding biaxial crystals oriented with an optic axis parallel to the light beam is small.
3. When viewed between crossed polars, anisotropic substances usually appear 'light'. On rotating the samples, they become 'dark' in a certain position. This is known as *extinction* and occurs every 90°; at the 45° position, the samples exhibit *maximum brightness*. If the samples have *parallel extinction*, i.e. the extinction occurs when a pronounced feature of the crystals, such as an edge, is parallel to the direction of vibration of the polarised light, then it is likely that the crystals possess some symmetry, i.e. they have monoclinic symmetry, or higher.

By using steps (1) to (3), it is often possible to guess the symmetry and unit cell of a crystalline substance. This can be a useful prelude to making more time-consuming X-ray diffraction studies. If crystals have a recognisable morphology, use can be made of this in orienting them for X-ray work or physical property measurements as a function of crystal orientation.

Crystalline substances may be identified according to their optical properties, refractive indices, optic axes, etc. Much use is made of this in mineralogy. Standard tables are available with which optical data for an unidentified mineral may be compared. The method is little used outside mineralogy for the identification of complete unknowns. However, where only a limited number of substances are possible, optical microscopy can be a very powerful method of identification. For instance, in the synthesis of new compounds or phase diagram studies on a specific system, only a limited number of phases are likely to appear. Provided the general appearance and optical properties of most or all are known, then optical microscopy may provide an extremely rapid method of phase analysis. Thus, the purity of a sample may be checked provided that the impurities form a separate crystalline or amorphous phase. Low impurity levels are easily detected, especially if the optical properties of the impurity are markedly different from those of the major phase.

The quality of single crystals may be assessed with a polarising microscope. Good-quality crystals should show *sharp extinction*, i.e. on rotating the crystal relative to the polars, extinction should occur simultaneously throughout the entire crystal. Crystal aggregates may show *wavy* or irregular extinction. If the crystal divides into strips which extinguish alternately as the crystal rotates, then the crystal is likely to be *twinned*. A twinned crystal is not a single crystal but contains crystal domains in two, or more, symmetry-related orientations. Other optical effects associated with twinning are striation and crosshatching. Twinning may occur as a result of the crystal growth mechanism or as a consequence of a phase transition from a high symmetry to a lower symmetry phase, as in paraelectric–ferroelectric or paramagnetic–ferromagnetic transitions. Optical microscopy is a powerful method for studying phase transitions directly if a hot stage is fitted so that the temperature of the sample can be varied during microscopic examination.

The quality and homogeneity of glasses may be checked for the presence or absence of crystalline inclusions. Compositional homogeneity may be checked by measuring the refractive index of a random selection of fragments of crushed glass: the refractive index of glass is usually composition dependent. Although glass is isotropic, glasses that have not been annealed properly may show *stress birefringence* and appear light and apparently anisotropic between crossed polars.

The Becke line method used for refractive index determination is based on the same principle that is now used in fibre optic communications, optical waveguides and Christmas tree decorations (see Chapter 10). In these, light is transmitted down a cable and it does not matter whether the cable is straight or bent because the light is reflected internally off the walls of the cable and is therefore constrained to remain inside the cable. The key to achieving this is to arrange that the cable is coated ('cladded') with a substance of lower refractive index than that of the cable core. This is illustrated in Fig. 6.2; in (a) the core has higher n than the cladding and the light stays within the core; in (b) the reverse applies and the light passes out of the cable into the cladding and is, effectively, lost.

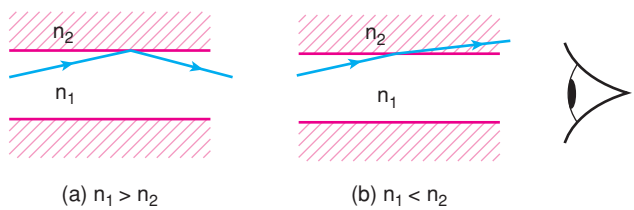


Figure 6.2 ***Illustration of the Becke line method for refractive index determination and in the transmission of light along optical fibres.***

In refractive index determination, the sample (refractive index n_1) is immersed in a liquid (refractive index n_2). When the sample is viewed slightly out of focus, the light is seen to emerge from the region of higher n. In practice, this is seen as a white border around the sample, the Becke line, which is either just inside or just outside the sample, depending on whether the sample or the immersion liquid has higher n.

6.2.1.2 Reflected light microscope

The reflected light microscope is similar to the transmission type except that the source and objective lens are on the same side of the sample. It is used to examine solid lumps of opaque material such as metals, minerals and ceramics. The amount of information that can be obtained depends on the care and skill with which the sample is prepared. It is best to have a flat polished surface which has been treated chemically to etch part of the sample preferentially. Some phases dissolve more quickly than others and this gives relief to the initially flat surface; other phases take on a coloured hue after etching. The information obtained primarily concerns the texture of the solid, i.e. the phases present and their identification and morphology.

Reflected light microscopy on polished and etched materials can provide much information on internal crystal surfaces (e.g. grain boundaries) and line defects (dislocations). Such defects are always present, even in good-quality single crystals. In the region of an emergent dislocation at the crystal surface, the crystal is in a stressed condition; if the surface is treated with a suitable chemical reagent, etching may occur preferentially at such stressed sites, resulting in etch pits. Dislocation densities may be determined from the number of etch pits per unit area. Grain boundaries also etch preferentially and low-angle grain boundaries may appear as rows of dislocation etch pits.

6.2.2 Electron microscopy

Electron microscopy is extremely versatile for providing morphological, structural and compositional information over a wide range of magnification. It is also often possible, using the same electron microscope, to record diffraction patterns in addition to magnified images, as shown schematically in Fig. 6.3, whereas with optical microscopes only the images can be recorded satisfactorily. At one extreme, *scanning electron microscopy* (SEM) complements optical microscopy for studying the texture, topography and surface features of powders or solid pieces and, because of the *depth of focus* of SEM instruments, the resulting pictures have a definite 3D quality. At the other extreme, *transmission electron microscopy* (TEM), *high-resolution electron microscopy* (HREM) and *scanning transmission electron microscopy* (STEM) are capable of giving structural and compositional information on an atomic scale. Resolution of $\leq$2 Å is readily achieved, and it is possible to 'see' individual atoms. However, lest anyone should think that these techniques are on the verge of solving all remaining problems concerning the structure of materials, it must be emphasised that there are formidable obstacles to be overcome before this goal is achieved; there is no immediate prospect of redundancy for more conventional crystallographers!

Electron microscopes operate in either transmission or reflection. For transmission, samples should usually be thinner than $\sim$2000 Å because electrons interact strongly with matter and are completely absorbed by thick particles. Sample preparation may be difficult, especially if it is not possible to prepare thin foils. Thinning techniques, such as ion bombardment, are used, but there is a danger that ion bombardment may lead to structural modification of the solid in question or that different parts of the material may etch preferentially in the ion beam. One possible solution is to use higher voltage instruments, e.g. 1 MV. Thicker samples may then be used since the beam is more penetrating; in addition, the amount of background scatter is reduced and higher resolution may be obtained. Alternatively, if the solid can be crushed into a fine powder then at least some of the resulting particles should be thin enough to be viewed in transmission.

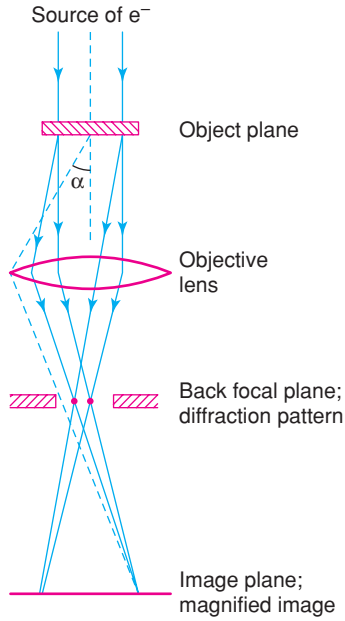


Figure 6.3 ***Simplified ray diagram for an electron microscope showing diffraction and imaging possibilities.***

In reflection mode, sample thickness is not a problem. It is usually necessary to coat the sample with a thin layer of metal, to prevent the build-up of charge on the surface. The main reflection mode, SEM, covers the magnification range between the lower resolution limit of optical microscopy ($\sim$1 μm) and the upper practical working limit of TEM ($\sim$ 0.1 μm), although SEM can be used over a much wider range, from $\sim 10^{-2}$ to $\sim 10^{2}$ μm. The approximate working ranges of different kinds of microscope are summarised in Fig. 6.4.

In considering the interactions of an electron beam with a solid target, the interactions can be either elastic or inelastic, Fig. 6.5. In elastic collisions, there is very little loss in energy of the electron during the collision and its wave characteristics remain coherent with the incident beam; coherently scattered electrons interfere

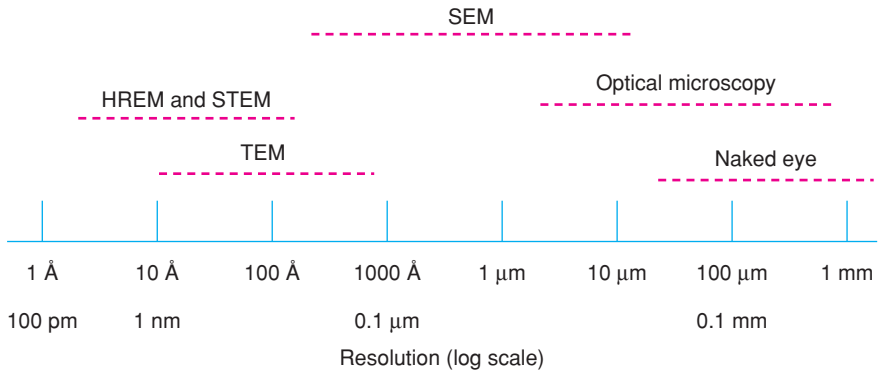


Figure 6.4 ***Working ranges of various techniques used for viewing solids. TEM = transmission electron microscopy; HREM = high-resolution electron microscopy; SEM = scanning electron microscopy.***

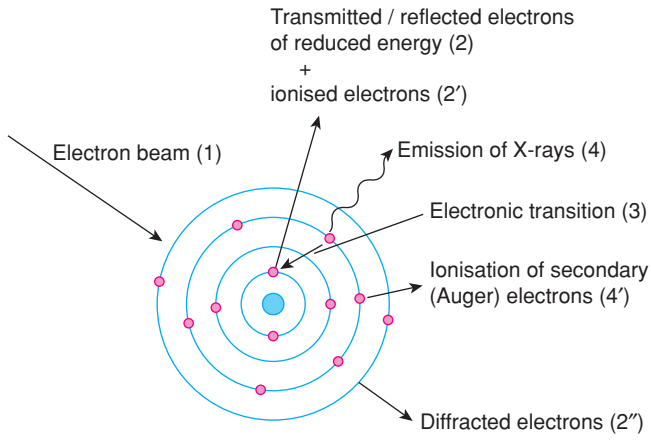


Figure 6.5 ***Some of the processes that occur on bombarding a sample with electrons and a range of ionisation and energy release mechanisms.***

giving rise to measurable diffraction patterns, as in XRD. In inelastic scattering, the incident electrons lose much of their energy by energy transfer to the solid target. The target can respond in various ways, by re-emitting secondary (or ionised) electrons, by emitting photons of energy that span the visible and X-ray part of the electromagnetic spectrum, by the excitation of lattice vibrations and associated heating of the sample or by structural changes known as radiation damage. The various emitted electrons and radiation form the basis of a number of imaging, diffraction, analytical and spectroscopic techniques.

The analytical techniques include *electron probe microanalysis* (EPMA) and *energy-dispersive analysis of X-rays* (EDS, EDX or EDAX). These utilise the characteristic X-ray emission spectra of the elements present, Fig. 5.3(b), which are generated upon electron bombardment in the electron microscope. By scanning either their wavelength (*wavelength dispersive*, WD) or their energy (*energy dispersive*, ED) and recording the emission spectra, it is possible to identify the elements present. With suitable calibration, quantitative elemental analysis may be made. Elements heavier than, and including, sodium can be determined routinely and with some instruments, all elements from Be onwards can be analysed. For lighter elements, however, the alternative techniques, *Auger spectroscopy* and *electron energy loss spectroscopy* (EELS), can be used.

Generation of low-energy Auger electrons (4′), Fig. 6.5, occurs when electrons, and not X-rays, are ejected from the sample as an energy release mechanism, in response to an atom returning from an excited state to its ground state by an electronic transition (3). The energy of Auger electrons is characteristic of the atom from which they are emitted.

EELS detects the incident electrons (2) which were responsible for the initial ionisation of electrons (2′). Since energy is required to ionise the atoms, the EELS electrons (2) are of reduced energy compared with the energy of the incident electron beam (1).

In electron microscopes, electrons emitted from a tungsten filament (*electron gun*) are accelerated through a high voltage (50–100 kV). Their wavelength, λ, may be calculated by combining the de Broglie relationship, i.e. $\lambda = h/mv$, with the equation for the kinetic energy of an electron accelerated through voltage V, i.e. $^1/_2 mv^2 = eV$, and rearranging to eliminate velocity, v, to obtain:

$$\lambda = \frac{h}{\sqrt{2meV}} = \frac{12.3}{\sqrt{V}} \text{ Å} \tag{6.2}$$

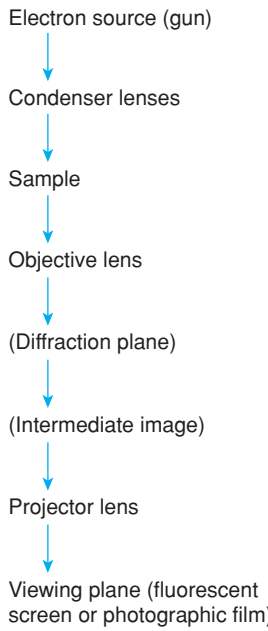


Figure 6.6 ***Basic components of an electron microscope.***

where m and e are the mass and charge of the electron, respectively. At high voltage, as the velocity of the electrons approaches the velocity of light, m is increased by relativistic effects. The electron wavelengths are smaller than X-ray wavelengths, e.g. $\lambda \approx 0.04$ Å at 90 kV accelerating voltage. Consequently, the Bragg angles for diffraction are small and the diffracted beams are concentrated into narrow cones centred on the undiffracted beam.

In order to use electrons in a microscope for image formation, it is necessary to be able to focus them. As yet, no one has found a substance that can act as a lens to focus electrons but, fortunately, electrons may be focused by either an electrostatic or a magnetic field. Electron microscopes contain several *electromagnetic lenses*, Fig. 6.6. The *condenser lenses* control the size and angular spread of the incident electron beam. Transmitted electrons pass through a sequence of lenses – *objective*, *intermediate* and *projector* – and form a magnified image of the sample on a fluorescent screen. Photographs may be taken if so desired. By changing the relative position of the viewing screen, the diffraction pattern, rather than the image, of the specimen is seen and can also be photographed. The region of the sample that is chosen for imaging is controlled by an aperture placed in the *intermediate image plane*. This is particularly important in examining polycrystalline materials that contain more than one phase. Images may be recorded in either *bright field* or *dark field*. In dark field imaging, only the beams diffracted from the particle of interest are allowed to recombine to form the image, which appears bright on a dark background. In bright field imaging, only the electrons transmitted by the sample form the image which therefore appears dark on a bright background.

There are very many techniques associated with electron microscopes; some of the most important are indicated in Fig. 6.5 and Fig. 6.7. In addition to diffraction and imaging techniques, there is a wide range of spectroscopic methods, some of which have already been referred to. Next, a brief summary of the main techniques associated with electron microscopy is given.

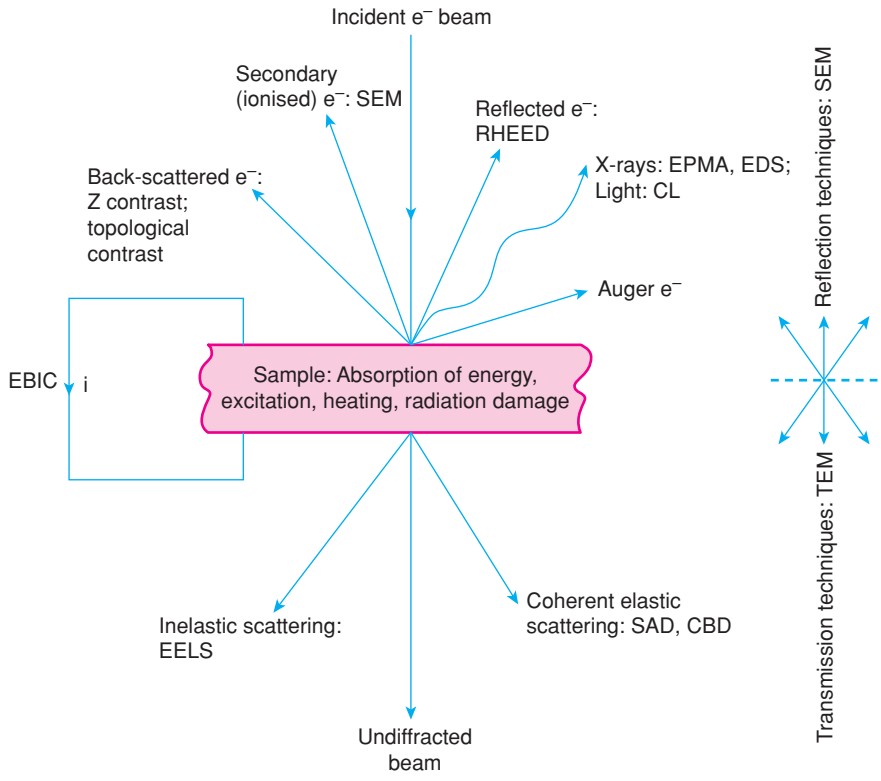


Figure 6.7 ***Reflection and transmission signals in SEM and TEM modes***

6.2.2.1 Scanning electron microscopy

In SEM, electrons from the electron gun, accelerated through 5–50 keV, are focused to a small spot, 50–500 Å in diameter, on the sample surface, Fig. 6.7 and Fig. 6.8. The electron beam scans the sample, rather like the spot on a television screen; the penetration depth of the electron beam is typically up to 1 μm. The escape depths for various electrons and radiation depend on their energies. Back-scattered electrons, BSE, are able to escape from a thicker region of the sample provided that they avoid secondary collisions after the initial back-scattering event, Fig. 6.9. The escape depth for secondary ionised electrons or emitted radiation (X-rays and visible light) is more shallow because these have much less energy than the incident beam. The incident electrons that penetrate deeper into the sample lose their energy to lattice vibrations or radiation damage and are unable to escape from the sample. In SEM, the incident electrons are unable to penetrate the sample completely unlike TEM and STEM instruments for which either thinner samples (TEM) and/or higher energy incident electrons (STEM) are used. The lower limit of resolution with SEM instruments is ~100 Å. The STEM instruments combine the scanning feature of the SEM with the intrinsically higher resolution obtainable with TEM.

On bombardment of a sample with electrons in the SEM, particles and radiation with a wide range of energies, arising from a number of processes, Fig. 6.5 and Fig. 6.7, can be detected. These include secondary electrons arising from absorption and re-emission processes; back-scattered electrons, without appreciable loss of energy; elastically scattered or diffracted electrons; Auger electrons of characteristic energies; X-rays, both characteristic and white radiation (or Bremsstrahlung); and visible light.

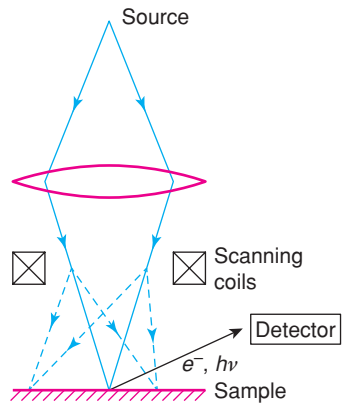


Figure 6.8 ***Principle of the scanning electron microscope.***

The main application of SEM is for surveying materials under high magnification and providing information on sizes, shapes and compositions as seen from solid surfaces. The results complement those obtained from optical microscopy by providing information on submicron-sized features, but can also extend to features up to hundreds of microns across. Images may be recorded using both secondary electrons and back-scattered electrons, with instrumental conditions optimised for good topographic contrast.

6.2.2.2 Electron probe microanalysis, EPMA, and energy-dispersive X-ray spectroscopy, EDS or EDX

These techniques operate in SEM mode and utilise the characteristic X-ray emission spectra generated when inner-shell electrons from atoms are ionised by high-energy bombardment, Fig. 6.5.

In EPMA, characteristic X-rays are detected by *crystal spectrometers* which are oriented at a suitable Bragg angle to the emitted X-ray beam. For each crystal spectrometer, a particular set of lattice planes that diffract X-rays strongly is selected. For each element, the wavelength required to satisfy Bragg's law

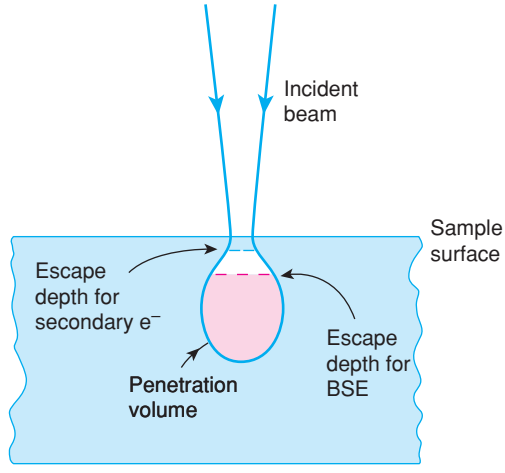


Figure 6.9 ***Penetration and escape depths in SEM.***

for a particular characteristic X-ray line is given by Moseley's law, equation (5.1), and spectrometers are therefore oriented so that the crystal is at the appropriate Bragg angle for the element in question. By using several crystal spectrometers, each with a specific *d*-spacing, elements as light as Be can be detected and analysed. In EPMA, X-rays are emitted from regions a few microns in size and this places a limit on the compositional resolution that can be achieved with this technique. The large probe size is required in order to obtain statistically relevant X-ray intensity data on a reasonable time scale. Smaller probe volumes may be used for instruments operating in EDS mode, which gives higher resolution for image contrast but with only semiquantitative accuracy.

In EDS, the method for detecting the characteristic X-rays emitted by a sample is different to that used in EPMA in which the crystal spectrometers are adjusted to detect one element at a time. In EDS, the whole X-ray emission spectrum is detected simultaneously by a process in which X-rays are absorbed by a semiconducting material, such as Si or Ge, leading to excitation of electrons and the formation of electron–hole pairs. The electron–hole pairs are swept away by an applied voltage, which is detected as a voltage pulse that is amplified and displayed on a screen. The key point is that the voltage pulse is proportional to the energy of the incoming X-rays, which is characteristic of the elements responsible for the X-ray emission. By analyzing the energies, and distribution of energies, of the excited electrons, the X-ray emission spectrum of the sample area under bombardment is obtained.

EDS is very useful for getting an overview of the elements present in a sample, although it is less sensitive for light elements. Semiquantitative analysis is possible with EDS with an approximate detection limit of 1000–3000 ppm (∼10 wt%), which compares unfavourably with EPMA for which the detection limit is an order of magnitude lower. In addition, the width of the spectral peaks, as given by the full width at half-maximum, FWHM, is an order of magnitude greater for EDS than for WDS (EPMA); this means that the resolution of EDS peaks is less, giving frequent overlaps of peaks arising from different elements. By EDS, the FWHM for a peak in the region of the Mn $K\alpha$ peak, for example, is ∼130 eV but only ∼10 eV by WDS. The background in an EDS spectrum is also subject to various interferences and stray emissions, whereas this is not an issue with EPMA for which only peaks at specific Bragg angles are detected by the spectrometer. Nevertheless, EPMA is rather a specialist technique whereas most SEM instruments are fitted with an EDS detector, which is extremely useful for routine element identification and mapping of element distributions in a sample.

6.2.2.3 Auger electron (emission) microscopy and spectroscopy, AES

We are familiar with the emission of characteristic X-rays produced when an outer shell electron drops down to fill a vacancy in an inner shell, such as $K\alpha$ and $K\beta$ transitions, Fig. 5.3. An alternative energy release mechanism is emission of outer shell electrons known as Auger electrons, Fig. 6.10. These are low-energy electrons with energy in the range 0–2.5 keV and can be detected only when they escape from near-surface atoms; Auger electrons emitted from deeper within a sample would lose energy and be absorbed by successive collisions. The escape depth for Auger electrons is typically 1 nm compared with 0.5 μm for EPMA.

Since Auger electrons arise from inter-shell electronic transitions with specific energies, their energies are characteristic of the elements responsible and can be used for identification and analysis purposes in addition to providing specific band structure information. Auger emission is independent of the incident electron beam energy provided that it is above the threshold necessary to cause the initial ionisation process, Fig. 6.10, and can also be stimulated by bombardment with photons or ions. Clearly, Auger techniques are very useful for characterising local surface structures and chemistries. However, sample surfaces are often contaminated by absorbed species which may be different from the main body of a sample. Care is therefore required in preparing sample surfaces and in interpretation of Auger data; similar comments apply to other surface-sensitive techniques such as photoelectron spectroscopy.

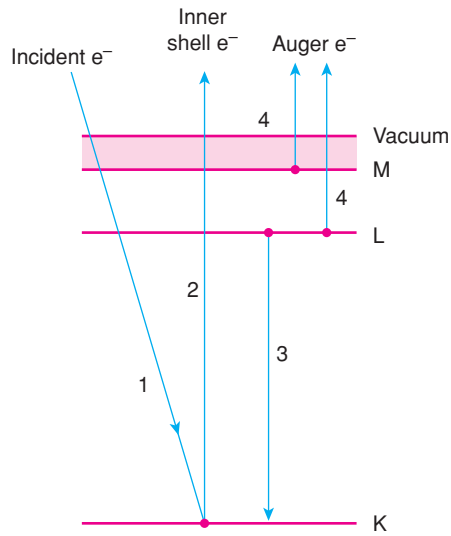


Figure 6.10 ***Processes responsible for Auger electron emission.***

Auger techniques find important applications in studying the catalytic behaviour of materials since catalytic processes usually involve surface reactions. They can be used to study compositional variations in solid solution-based ceramics since substitutional cations are often segregated either towards or away from surfaces instead of being homogeneously distributed; by Auger mapping of a cross-section through a sample, an elemental distribution map may be provided. An alternative method is to carry out Auger depth profiling; surface layers of a material are gradually removed by ion beam ablation and AES spectra recorded as a function of depth.

An example of the use of AES to study compositional variations as a function of depth for thin films of a ferroelectric material deposited on to Si substrates is shown in Fig. 6.11: film structures such as this find applications in microelectronics, especially in microelectromechanical systems, MEMS, which feature coupling between an electrical impulse and dimensional changes associated with the piezoelectric response of the film. The film structure, Fig. 6.11, consisted of an Si substrate coated with a 200 nm thick SiO_2 layer and an outer layer of $(Pb_{0.4}Sr_{0.6})TiO_3$, PST. The surface layers were gradually removed by ion beam ablation and AES spectra recorded periodically.

The AES spectra represent the elements present to a depth of ~1 nm. Initially spectral peaks assigned to Pb, Sr, Ti and O are seen. With increased ablation, the intensities of these peaks diminish and a new peak assigned to Si, at 107 eV, appears. In addition, the oxygen peak at 534 eV attributed to PST is gradually replaced by a second oxygen peak at 538 eV attributed to oxygen in a different environment, in SiO_2. With further ablation, the PST peaks disappear entirely, leaving only peaks attributable to SiO_2, Fig. 6.11(b), and these peaks subsequently disappear with further ablation to be replaced by a single peak associated with the Si substrate (c).

This example shows the compositional sensitivity of AES at the nanometre level and also shows that with further analysis, structural and bonding information can potentially be obtained: the peak positions for both Si in SiO_2 and elemental silicon and for O in SiO_2 and PST are different, reflecting different local chemistries for these elements. In some cases, interdiffusion of ions between different layers may occur and also be detected by spectral peak shifts when the ions are in different chemical and structural environments.

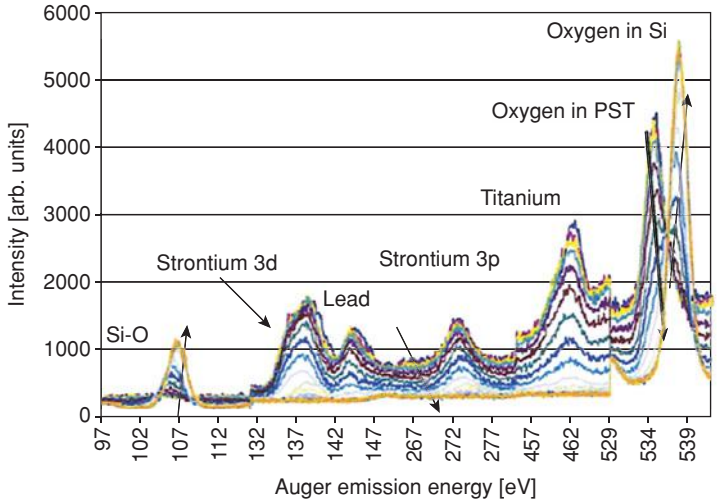


Figure 6.11 ***Auger spectra of $(Pb_{0.4}Sr_{0.6})TiO_3/SiO_2/Si$ layered structure after gradual removal of surface layers by ion sputtering. (a) Systematic decrease in Pb, Sr and Ti intensities; increase in Si and O intensities in SiO_2 and decrease in O intensities in $(Pb_{0.4}Sr_{0.6})TiO_3$; (b) SiO_2 layer exposed after removal of the outer $(Pb_{0.4}Sr_{0.6})TiO_3$ layer.***

6.2.2.4 Cathodoluminescence, CL

Cathodoluminescence refers to the emission of visible light in response to electron bombardment in the electron microscope. Electron bombardment induces various ionisation processes and electronic transitions, Fig. 6.5; included in these processes in non-metallic materials are promotion of electrons across the band gap from valence band to conduction band, Fig. 6.12. The subsequent return of the electrons to the valence band, either by a direct transition, 1, or via intermediate-level trap states, 2 or 3, is responsible for CL. When

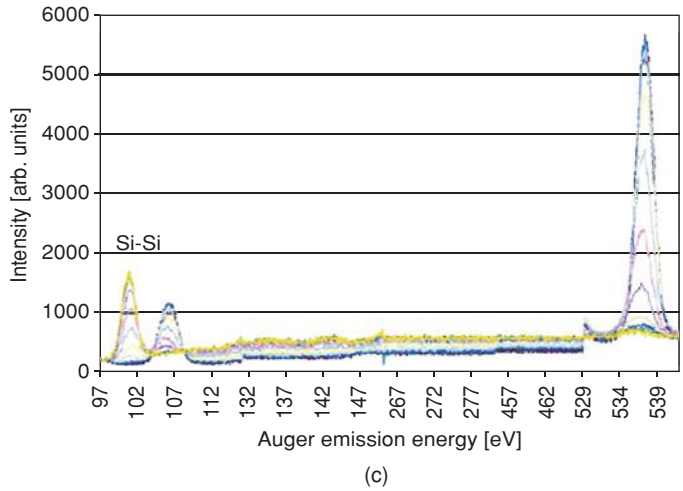


Figure 6.11 **(Continued)** ***(c) decreasing SiO_2 (blue) and increasing Si (yellow) intensities after further ablation. Reproduced with permission from A. Lüker* et al.*,* Thin Solid Films*,* 518*, 3763, © 2010 Elsevier.***

electrons are promoted into the conduction band, holes are left behind in the valence band and the reverse process may be regarded as *radiative recombination of electron–hole pairs*.

CL works best with insulating or semiconducting materials such as ceramics, organic and biomaterials, since these have a clearly defined band gap, but not with metals. It is also informative for materials containing defects or dopants with associated inter-band trap levels, Fig. 6.12, since the wavelengths of the CL emissions are characteristic of the species or defects responsible. CL can be used for both analytical and imaging purposes to (a) identify species responsible for CL emission and (b) map their location in a sample at the nanometre level. CL finds practical applications as fluorescent screens in a range of display devices; further details are given in Chapter 10.

An example of the use of CL as a powerful analytical technique, especially when combined with other techniques such as SEM and EDS, is shown in Fig. 6.13 for samples of AlN that were either undoped or doped with 0.2% Eu. The CL spectrum of both samples shows a peak at 360 nm, which is attributed to defects

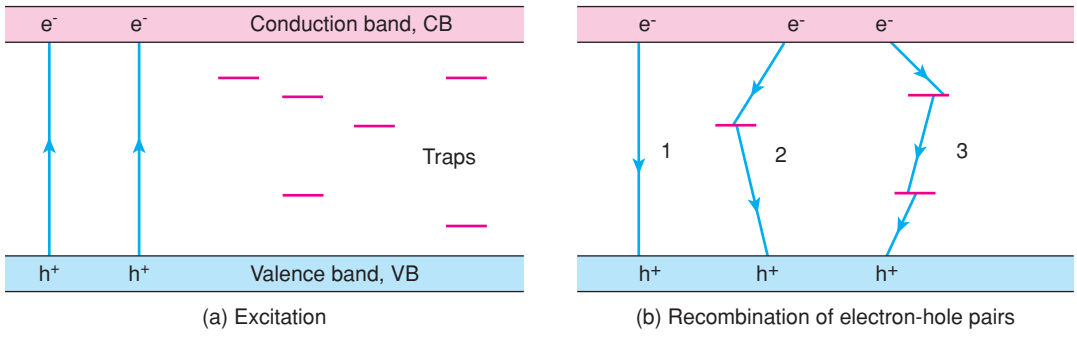


Figure 6.12 ***Cathodoluminescence: (a) excitation across the band gap; (b) radiative recombination involving (1) intrinsic transitions and (2, 3) transitions involving extrinsic (defect) states.***

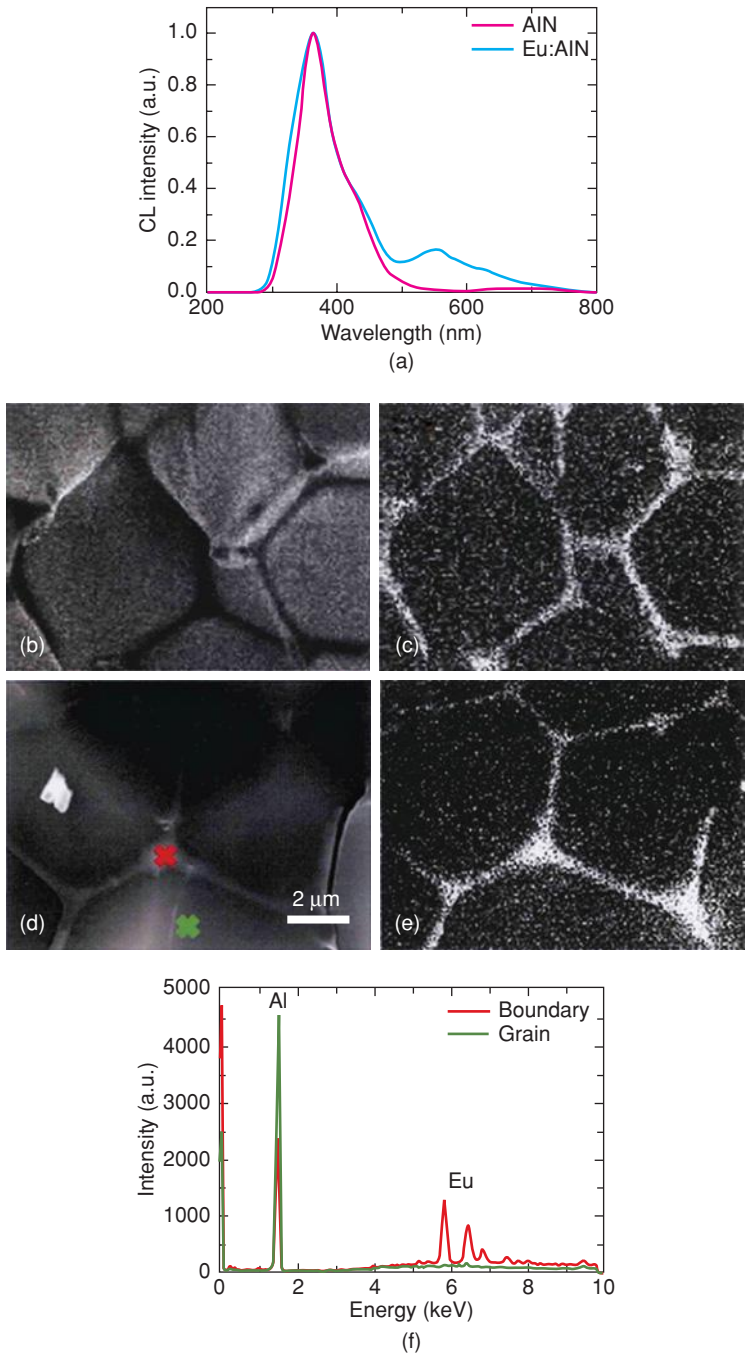


Figure 6.13 ***(a) Normalised CL spectra for undoped AlN and AlN doped with 0.2% Eu. CL images using (b) 360 nm and (c) 550 nm spectral peaks in (a); (d) SEM and (e) EDS map of Eu distribution and (f) Eu, Al spectra. Reproduced from B. Dierre* et al.*,* Sci. Technol. Adv. Mater. 11 *© 2010 IOP Publishing.***

in the parent structure, whereas the spectrum of the doped sample shows an additional weak band at 550 nm, Fig. 6.13(a). CL images using the two spectral peaks are shown in Fig. 6.13(b) and (c), from which it may be deduced that the Eu is essentially concentrated at the grain boundary regions (c). This is further confirmed by SEM (d) and EDS images of grain and grain boundary regions showing the regions of high Eu content in (e). Parts of the EDS spectra in (f) for the two regions show a high concentration of Al in the grains with a much smaller concentration of Al in the grain boundaries but a high Eu concentration. These results collectively show that Eu has very little solubility in the AlN crystal lattice and, instead, the Eu dopant segregates to the grain boundary where it probably forms a separate crystalline phase. Without the high resolution capabilities of SEM, EDS and CL, such information could not be obtained.

6.2.2.5 Transmission electron microscopy, TEM, and scanning transmission electron microscopy, STEM

TEM and related techniques detect transmitted electrons and radiation, in contrast to SEM, EPMA, EDS, AES and CL, which are reflection-based techniques. With TEM, both diffraction patterns and magnified images can be obtained from the same sample area; diffraction patterns give unit cell and space group information and, using HREM, can be used for lattice imaging purposes. In imaging mode, TEM gives morphological information on the sample.

There are two categories of diffraction techniques. In *selected area diffraction*, SAD, a parallel beam of electrons strikes the sample over an area with micron dimensions; electrons are diffracted strongly by the sample but are also absorbed and, therefore, the resulting intensities of the diffracted beams depend on sample thickness and cannot realistically be used for crystal structure determination in the same way that X-rays and neutrons are used, Chapter 5. The SAD technique, although rather specialised, is invaluable for obtaining unit cell information and especially for obtaining space group data and detecting the presence of superlattices, which may be hard to detect by XRD. It is particularly valuable when the correct unit cell is not known or cannot be determined from powder XRD and the crystal size is too small for single-crystal XRD techniques. Further discussion on SAD requires an appreciation of the reciprocal lattice concept and is beyond the scope of this chapter.

In *convergent beam diffraction*, CBD, a convergent beam is focused on the sample with a spot size of $\sim$10 nm, which gives a far higher resolution than obtained with SAD. The nature of the diffraction patterns is rather different to those obtained by SAD, but similar information can be obtained on crystal symmetry and lattice parameters. With high-resolution TEM, elemental information at the 1–10 nm level can also be obtained using EDS and EELS; these techniques are known collectively as *analytical electron microscopy*, AEM.

In favourable circumstances, HREM may give direct information on crystal structures by *lattice imaging*. A certain number of the diffracted beams are allowed to recombine and form a direct image of the crystal structure, giving detail at the atomic level. Difficulties and ambiguities arise because first, the images are only a projection of the structure and second, the image obtained varies greatly with the instrumental focusing conditions. It is possible to calculate images for different models of the crystal structure and, by comparison with the observed images, one can find structural models that are at least consistent with the experimental data. This is clearly not as useful as solving crystal structures *ab initio*; nevertheless, it is valuable for studying complex inorganic materials which have pronounced layer, shear or block structures. Images may be obtained from which, for instance, layer spacings or block widths may be determined. An example of an HREM image is shown in Fig. 6.14.

Although it is impressive that individual atoms can be effectively ‘seen’ by HREM, the technique is not working at the theoretically possible resolution limits since the associated electron wavelength, given by equation (6.2), is 2–3 orders of magnitude smaller, especially for high-voltage, 1 MeV, instruments.

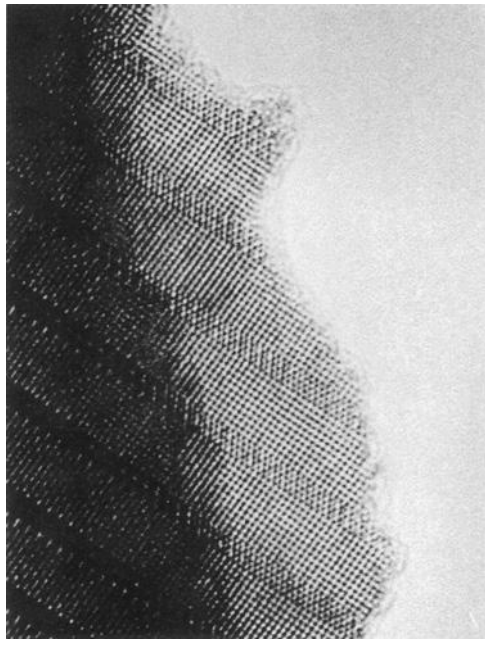


Figure 6.14 ***High-resolution electron micrograph of an intergrowth tungsten bronze, $Rb_{0.1}WO_3$. Black dots represent WO_6 octahedra. The structure may be regarded as an intergrowth of primitive cubic WO_3 (strips of black dots based on a square grid) and hexagonal WO_3 containing Rb (narrow strips of black dots on a hexagonal grid). Photograph courtesy of Dr M. Sundberg, University of Stockholm, Department of Chemistry.***

The problem that arises at these high magnifications concerns the difficulty of focusing electrons using the electromagnetic lenses in the microscopes. There are three kinds of lens aberrations, which can be corrected to varying extents. *Spherical aberration* is a basic focusing problem and is corrected by introducing a negative spherical aberration into the lenses to cancel out the positive observed spherical aberration. *Astigmatism* arises because the cross-section of the incident electron beam is elliptical rather than circular. *Chromatic aberration* arises because the incident electrons have a range of energies and therefore range of wavelengths.

6.2.2.6 Electron energy loss spectroscopy, EELS

An exciting recent advance in TEM-related techniques is the ability not only to image individual atoms but also to identify them using the EELS technique. EELS detects and analyses those electrons which were responsible for the initial ionisation and excitation processes that occurred during electron bombardment of the sample. These electrons have an energy reduced by amounts equivalent to the ionisation potential of the secondary emitted electrons together with any kinetic energy imparted to them or to the energy used in promoting electrons to higher atomic levels. By scanning the energies and wavelengths of these electrons of reduced energy, an absorption spectrum is obtained, which is analogous in many ways to that obtained by the X-ray XANES technique, Section 6.3.5. Thus, the *energy loss near edge structure*, ELNES, technique is the electron analogue of XANES and the extended *energy loss fine structure*, EXELFS, technique is the electron analogue of EXAFS.

The ELNES technique can be used to determine element oxidation states, as shown by a mineral example in Fig. 6.15 of thin-layer precipitates, known as exsolution lamellae, of haematite, Fe_2O_3, within a matrix of

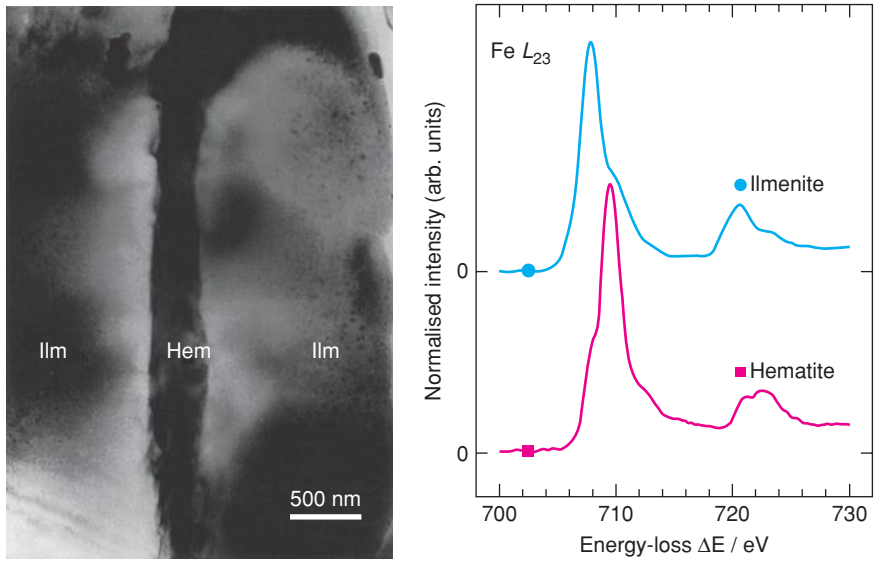


Figure 6.15 ***(a) TEM bright field micrograph of an ilmenite (Ilm) sample containing haematite (Hem) exsolution lamellae with (b) Fe L-edge ELNES of the two components. Reproduced with permission from P. A. van Aken, B. Liebscher and V. J. Styrsa,*** **Phys. Chem. Minerals**, 25, ***323, © 1998 Springer Berlin / Heidelberg.***

ilmenite, $FeTiO_3$. The haematite lamellae are ~100 nm thick, as shown by the TEM bright field micrograph in (a). The oxidation state of Fe in the two components was shown by the EELS spectrum (b) in the region of the L absorption edge ($3s \rightarrow \infty$ ionisation transition). Using other Fe-containing samples with known oxidation states as standards, the absorption peak at 708 eV is characteristic of Fe^{2+} and the peak at 710 eV is characteristic of Fe^{3+}, which allows assignment of the two peaks to the ilmenite (containing Fe^{2+}) and the haematite (containing Fe^{3+}) components. This demonstrates the usefulness of EELS, and in this particular case of L-edge ELNES spectroscopy, as a valence-selective probe at the nanometre level of materials, whether minerals or synthetic, that contain variable valence transition metal cations.

EELS is particularly useful for detecting elements of low atomic number, which absorb electrons less strongly than elements of high atomic number. The energy resolution is good, 0.3–2.0 eV, and the technique is very useful for nanometre-level mapping of element distributions; data collection is rapid but data processing is complex. EELS therefore complements EDS, which is more suited to heavy element analysis, but at a much lower energy resolution (>100 eV). The EDS technique is rather slow for mapping purposes but data processing is simple.

In addition to the use of EELS for compositional analysis associated with discrete electronic transitions and ionisations, the technique also detects excitation of electrons within the conduction band. Transitions to empty states above the Fermi level give information on the band structure and the degree of hybridisation between atomic orbitals, which therefore relates to local bonding details.

6.2.2.7 High-angle annular dark field, HAADF/Z-contrast STEM

The technique designed to allow atomic level resolution by EELS is STEM. The instrument has a very fine focus with a spot size <0.1–10 nm, which can be rastered across the sample. In addition to the transmitted

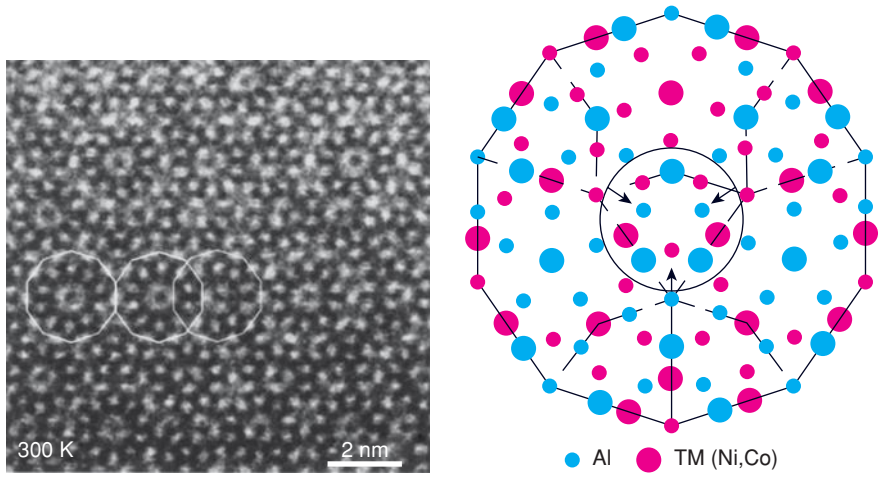


Figure 6.16 ***Atomic resolution HAADF–STEM image of $Al_{17}Ni_{20}CoO_8$ along the 10-fold symmetry axis (a) with an atomic model of the local structural regions (b). Reproduced from E. Abe,*** **Materials Transactions,** 44 ***(10), 2035, © 2003 The Japan Institute of Metals.***

EELS electrons, other incoherently scattered X-rays can be detected. Those scattered at high angle form the basis of the high-angle annular dark field, HAADF, or *Z*-contrast, imaging technique. By collecting such scattered electrons, coherence and diffraction effects between different atoms in the structure are reduced and effectively, each atom becomes an independent scatterer of electrons. With thin, suitably oriented crystals, individual atomic columns are illuminated sequentially as the probe scans across the specimen, giving a map in which the intensity depends on the square of the average atomic number, Z^2, of the atoms in the columns, giving good chemical contrast and without the need for image simulations.

Using this technique, many of the phase contrast and focusing problems associated with more conventional HREM are eliminated. In *Z*-contrast imaging, there is a correlation between the focus and size of the incident electron probe and the information that can be obtained; for defocus less than the optimum, the probe broadens, which causes the individual atomic columns to be unresolved; for higher defocus, with a narrowed probe, sharper image features are obtained but the compositional data are averaged over several columns, reducing the accuracy of the analyses. The optimum focus condition therefore represents a compromise in the images obtained between high-resolution sharp images and accurate local compositional information.

An example of *Z*-contrast imaging, which shows beautifully the 10-fold symmetry of the quasi-crystalline alloy $Al_{72}Ni_{20}Co_8$, is given in Fig. 6.16(a), with an atomic model of the local clusters in (b). Recalling that this technique is a transmission technique that represents projections in 2D of the atoms in the sample, then the bright spots are dominated by the heaviest atoms, which form approximate columns. The local 10-fold symmetry of three decagons can be seen within the three regions marked but, as is a characteristic of the quasi-crystalline state, the 10-fold symmetry does not extend to the structure as a whole but is confined to local cluster regions.

The two techniques, *Z*-contrast imaging and EELS, operate in a complementary manner, but independently: *Z*-contrast imaging uses high-angle scattering whereas EELS collects the low-angle transmitted beams,

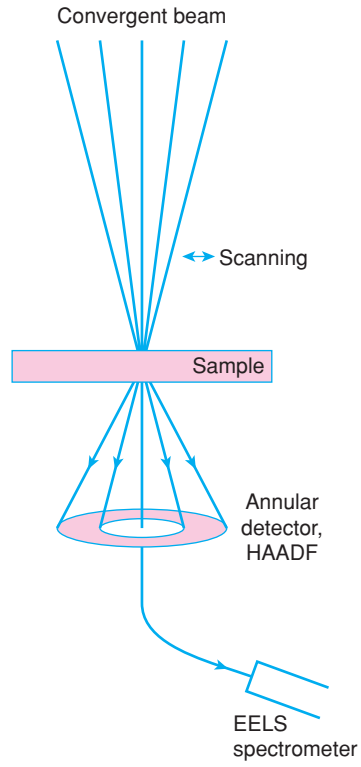


Figure 6.17 ***Schematic illustration of atomic resolution imaging and analysis using the annular detector for Z-contrast imaging and the modified primary transmitted beam for EELS.***

Fig. 6.17. In analysing samples, therefore, *Z*-contrast imaging may be used to locate particular regions of interest, which can then be further characterised using EELS.

6.3 Spectroscopic Techniques

There are many different spectroscopic techniques and all work on the principle that, under certain conditions, materials absorb or emit energy. The energy can take various forms. Usually it is electromagnetic radiation but it can also be sound waves, particles of matter, etc. The spectra are usually a plot of intensity of absorption or emission (y axis) as a function of energy (x axis). The energy axis is usually expressed as either frequency, f, or wavelength, λ, of the radiation. The various terms are interrelated by the classic equation

$$E = hf = hc\lambda^{-1} \tag{6.3}$$

where h is Planck's constant (6.6×10^{-34} J s), c is the velocity of light (2.998×10^{10} cm s^{-1}), f is frequency (in hertz, cycles s^{-1}) and λ is the wavelength (in cm). Units of λ^{-1} are cm^{-1} or wavenumbers and E is in joules. Chemists usually prefer joules per mole for units of E, in which case equation (6.3) is multiplied

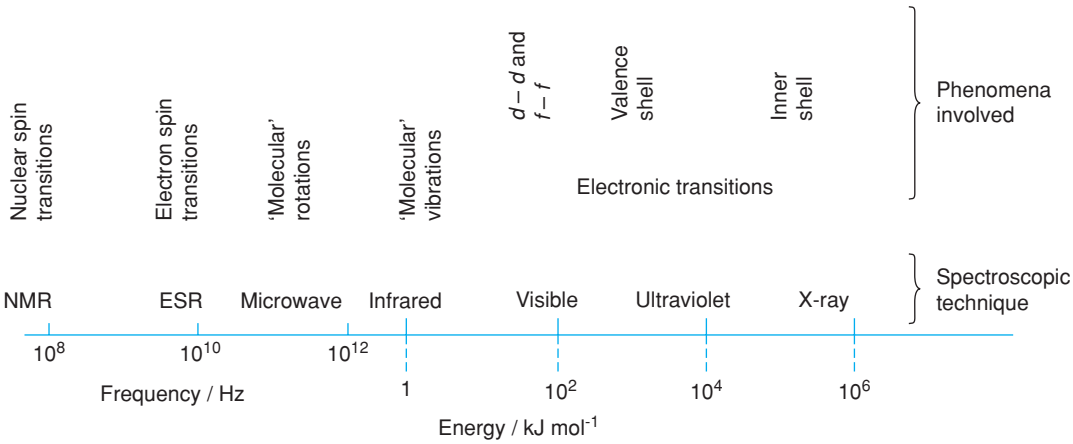


Figure 6.18 ***Principal regions of the electromagnetic spectrum and the associated spectroscopic techniques.***

by Avogadro's number, N. Some useful interconversions obtained by substituting for the constants are as follows:

$$\begin{aligned} E\ (\text{J mol}^{-1}) &= 3.991 \times 10^{-10} f \approx 4 \times 10^{-10} f\ (f \text{ in s}^{-1}) \\ E\ (\text{J mol}^{-1}) &= 11.97\lambda^{-1} \approx 12\lambda^{-1}\ (\lambda \text{ in cm}) \\ E\ (\text{s}^{-1}) &= 3 \times 10^{10}\lambda^{-1}\ (\lambda \text{ in cm}) \\ E\ (\text{eV}) &\approx 96E\ (\text{kJ mol}^{-1}) \end{aligned} \tag{6.4}$$

The electromagnetic spectrum covers an enormous span of frequency, wavelength and, therefore, energy. Different spectroscopic techniques operate over different, limited frequency ranges within this broad spectrum, depending on the processes and magnitudes of the energy changes involved, Fig. 6.18. At the low-frequency, long-wavelength end, the energy changes are small, <1 J mol^{-1}, but sufficient to cause the reversal of spins of nuclei or electrons in an applied magnetic field. Thus, the nuclear magnetic resonance (NMR) technique operates in the radiofrequency region at, for example, 400 MHz (4×10^8 Hz) and detects changes in nuclear spin state.

At higher frequencies and shorter wavelengths, the energy increases and, for instance, vibrational motions of atoms in molecules or solids may be altered by absorption or emission of infrared (IR) radiation. At still higher frequencies, electronic transitions within atoms may occur. For transitions involving outer (valence) shells, the associated energy usually lies in the visible and ultraviolet (UV) regions; for inner shell transitions, much larger energies are involved and fall in the X-ray region.

Spectroscopic measurements on solids complement well the results obtained from (X-ray) diffraction since spectroscopy gives information on *local* structure whereas diffraction is concerned primarily with *long-range* order. Spectroscopic techniques may give coordination numbers and site symmetries; impurities and imperfections can be detected; amorphous materials, such as glasses and gels, can be studied just as easily as crystalline materials. By contrast, the long-range periodic structures of crystals can be determined only by diffraction techniques (and occasionally by HREM); however, this results in an *average* picture of the local structure in which information on defects, impurities and subtle variations in local order may be lost. Also, diffraction techniques are of limited use for studying amorphous materials.

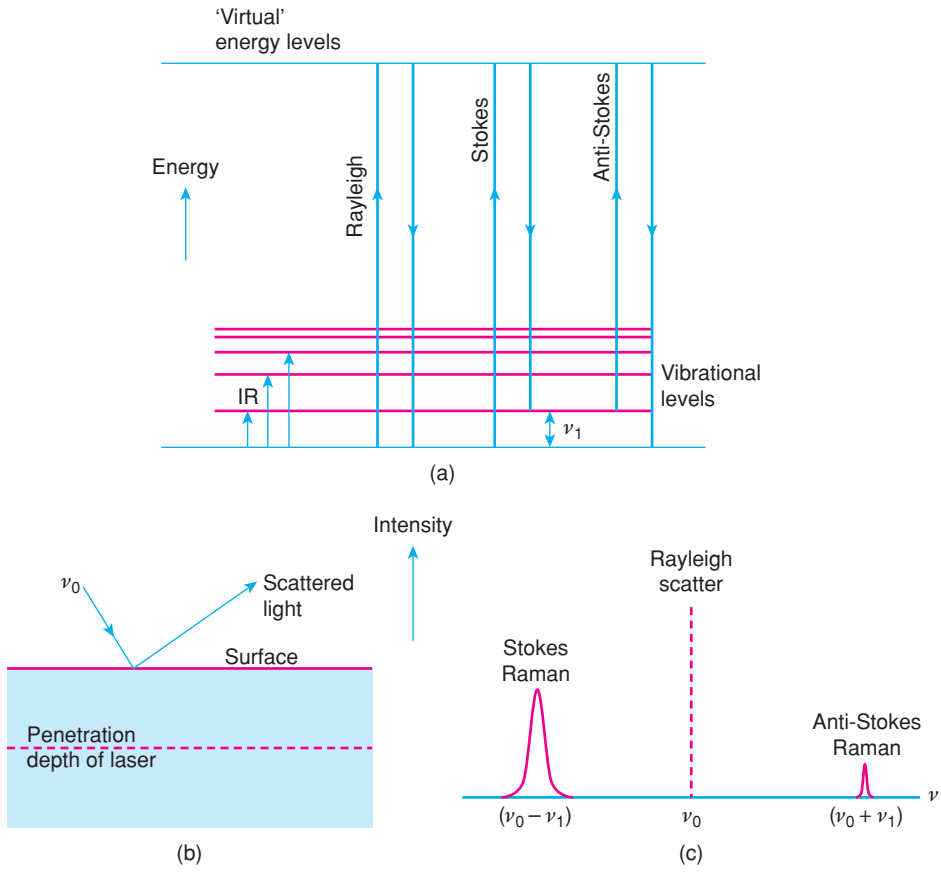


Figure 6.19 ***Vibrational transitions involved in IR and Raman spectroscopy.***

Several spectroscopic techniques associated with electron microscopes have been covered in Section 6.2.2. Some of the other main spectroscopic techniques and their uses in solid state chemistry are as follows.

6.3.1 Vibrational spectroscopy: IR and Raman

Atoms in solids vibrate at frequencies of 10^{12}–10^{13} Hz. Vibrational modes can be excited to higher energy states by absorption of radiation of appropriate frequency, Fig. 6.19(a). IR spectra and closely related Raman spectra are plots of intensity of absorption (IR) or scattering (Raman) as a function of frequency or wavenumber. Although these techniques are finding increasing application for the characterisation of non-molecular solids, they have been developed and used primarily for the study of molecular substances by measuring the resonant frequencies of vibrating bonds in specific functional groups. In the IR technique, the frequency of the incident radiation is varied and the radiation absorbed or transmitted by the sample is obtained. The process is basically simple since the local functional group or covalent bond responsible for the absorption is promoted to a higher vibrational level by the amount of absorbed energy.

Raman spectroscopy is named after the famous Indian scientist C. V. Raman, who carried out the key experiments in establishing the phenomenon of Raman scattering for which he was awarded the Nobel Prize

in Physics in 1930. In one experiment, he used sunlight, which was passed through a monochromatic filter and focused on samples by means of a telescope; in addition to the intense direct transmitted light, he detected very weak (Raman) scattering of different colour coming off in other directions. This showed that light of different wavelength was produced by a process of inelastic scattering.

In the Raman technique, the sample is illuminated with monochromatic light, usually generated by a laser. Two types of scattered light are produced. *Rayleigh scatter* emerges with exactly the same energy and wavelength as the incident light. *Raman scatter*, which is much less intense than Rayleigh scatter, emerges at either longer or shorter wavelength. The Raman process is intrinsically more complex than IR absorption; first, the irradiation source is of much higher energy and is usually a laser such as Nd:YAG with a fixed wavelength of 532 nm, an argon ion laser at 514 nm or an He–Ne laser at 633 nm. Second, in the absorption process the functional groups are promoted to a much higher energy state, a so-called virtual energy state, and immediately relax back, giving either Rayleigh scatter if the functional group returns to its initial vibrational state or Stokes Raman scatter if it returns to a vibrational state of higher energy, Fig. 6.19(a). In this case, for a difference in vibrational frequency ν_1 between the two vibrational states, the emitted Raman signal has frequency $\nu_0 - \nu_1$, where ν_0 is the frequency of the incident radiation.

A second possibility is that the functional group is already in an excited state, the incident radiation causes promotion to the virtual energy level and decay to the ground state occurs. In this case, the emitted radiation has frequency $\nu_0 + \nu_1$, giving so-called anti-Stokes scattering, Fig. 6.19(c). Anti-Stokes scattering is much weaker than Stokes scattering for the simple reason that there are many fewer functional groups in excited vibrational states than in ground states and therefore far fewer possibilities of making the anti-Stokes transition. However, both Stokes and anti-Stokes signals are many orders of magnitude weaker than the Rayleigh scattering. The net effect is that photons of light from the laser, of frequency ν_0, induce transitions in the sample and the photons gain or lose energy as a consequence. For a vibrational transition of frequency ν_1, associated Raman lines of frequency $\nu_0 \pm \nu_1$ appear in the scattered beam. This scattered light is detected in a direction perpendicular to the incident beam.

IR and Raman spectra of solids are usually complex, with a large number of peaks, each corresponding to a particular vibrational transition. A complete assignment of all the peaks to specific vibrational modes is possible with molecular materials and, in favourable cases, with non-molecular solids. IR and Raman spectra are usually completely different since the two techniques are governed by different selection rules. The number of observed peaks is considerably less than the total number of vibrational modes and different modes may be active in the two techniques. For a mode to be *IR active*, the associated dipole moment must vary during the vibrational cycle. Consequently, centrosymmetric vibrational modes are IR inactive. For a vibrational mode to be *Raman active*, the nuclear motions involved must produce a change in polarisability.

IR and Raman spectra are much used to identify specific functional groups, especially in organic molecules. In inorganic solids, covalently bonded linkages such as hydroxyl groups, trapped water and oxyanions give rise to intense IR and Raman peaks. Some examples shown in Fig. 6.20 differ in the region 300–1500 cm^{-1}, due largely to different oxyanions: (a) carbonate in $CaCO_3$, (b) nitrate in $NaNO_3$ and (c) sulfate in $CaSO_4$. Spectrum (c) contains an additional absorption doublet at ~3500 cm^{-1} caused by the H_2O molecules in the gypsum. Peaks that occur in the region ~3000–3500 cm^{-1} are usually characteristic of OH groups. The peak frequencies vary significantly depending on the O–H bond strength and give information on, for instance, the location of the OH group, whether it belongs to a water molecule and whether or not hydrogen bonding is present.

Peaks associated with covalently bonded groups, such as oxyanions, usually occur at frequencies above ~300 cm^{-1}. At lower frequencies, in the *far-IR region*, lattice vibrations give rise to peaks, e.g. alkali halides give broad lattice absorption bands at 100–300 cm^{-1}. The peak positions depend inversely on the mass of the ions as shown in the following sequences (positions in cm^{-1}): LiF (307), NaF (246), KF (190), RbF (156), CsF (127); and LiCl (191), NaCl (164), KCl (141), RbCl (118), CsCl (99).

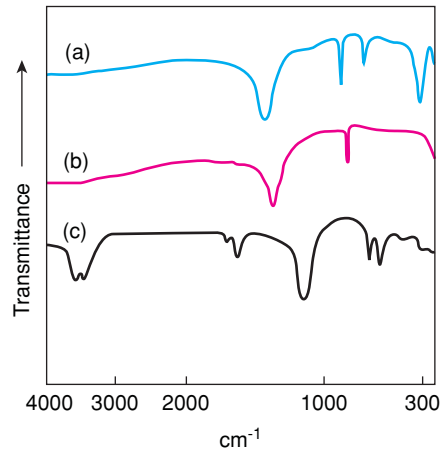


Figure 6.20 ***IR spectra of (a) calcite, $CaCO_3$, (b) $NaNO_3$ and (c) gypsum, $CaSO_4 \cdot 2H_2O$. Data reproduced with permission from P. A. Estep-Barnes, ed. Zussman,* Physical Methods in Determinative Mineralogy, *© 1977 Elsevier.***

Since inorganic solids give characteristic vibrational spectra, these may be used for identification purposes; files of reference spectra exist from which it is possible to identify unknowns. This is well established for organic functional groups but has not yet been adequately achieved for inorganic solids. Perhaps, with the comprehensive coverage and success of the ICDD X-ray powder diffraction file, there is less demand for a second method of fingerprinting inorganic solids.

An interesting example of the use of laser Raman spectroscopy for fingerprinting crystalline solids is in the distinction between two of the polymorphs of silica, quartz and cristobalite, Fig. 6.21. In the region 100–500 cm^{-1}, each polymorph gives three sharp peaks, but at different positions. Use was made of these differences in the analysis of ash from the Mount St Helens volcanic eruption in the USA.

A second example of the use of laser Raman spectroscopy for fingerprinting is in the distinction between diamond and graphite polymorphs of carbon, Fig. 4.16, which is used for assessing the quality of diamond

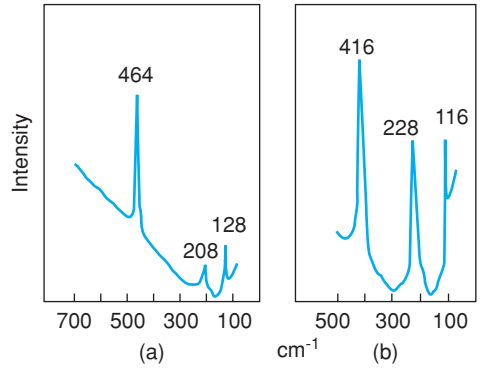


Figure 6.21 ***Laser Raman spectra of (a) quartz and (b) cristobalite. Reproduced with permission from Farwell and Gage, Crystalline silica in Mount St. Helens ash,* J. Anal. Chem., *53, 1529, © 1981 American Chemical Society.***

films. The vibrational frequencies of materials such as diamond are sensitive to pressure and with suitable calibration can be used as a pressure sensor or to study stress effects in, for example, diamond films. Tensile stress shifts Raman peaks to lower frequency or can even cause peak splitting due to loss of degeneracy; vice versa, compressive stress causes Raman peaks to shift to higher frequency.

Sample preparation for Raman spectroscopy is easy. Spectra are recorded in reflection mode from solid surfaces, Fig. 6.19(b). The penetration depth of the incident laser beam depends on laser wavelength and if it is desired to monitor spectra from different depths then different laser sources can be used. Good spectral resolution, 3–4 cm^{-1}, is obtained and also good spatial resolution, 0.5–2 μm, which allows spectra to be recorded from micron-sized regions of a solid sample. Raman spectra are symmetric on a frequency scale to either side of the incident frequency, but usually only the more intense Stokes scattering is recorded.

Apart from uses in identification, vibrational spectra may be used to provide structural information. However, a deeper understanding of the spectra is necessary and the assignment of peaks to specific vibrational modes is almost essential. Methods to carry out this assignment are not discussed here since they are rather complicated and as yet have been used only for simple crystal structures.

6.3.2 Visible and ultraviolet (UV) spectroscopy

Transitions of electrons between outermost energy levels are associated with energy changes in the range $\sim 10^4$–10^5 cm^{-1} or $\sim 10^2$–10^3 kJ mol^{-1}. These energies span the near-IR through the visible to the UV region, Fig. 6.18, and are often associated with colour. Various types of electronic transition occur, Fig. 6.22. The two atoms A and B are neighbouring atoms in some kind of solid structure; they may be an anion and a cation in an ionic crystal. The inner electron shells are localised on the individual atoms. The outermost shells may overlap to form delocalised bands of energy levels.

Four basic types of transition are indicated in Fig. 6.22:

(i) Promotion of an electron from a localised orbital on one atom to a higher energy localised orbital on the same atom. The associated absorption band is sometimes known as an *exciton band*. Examples include (a) *d*–*d* and *f*–*f* transitions in transition metal compounds, (b) outer shell transitions in heavy metal compounds, e.g. $6s \rightarrow 6p$ in lead(II) compounds, (c) transitions associated with defects such as trapped

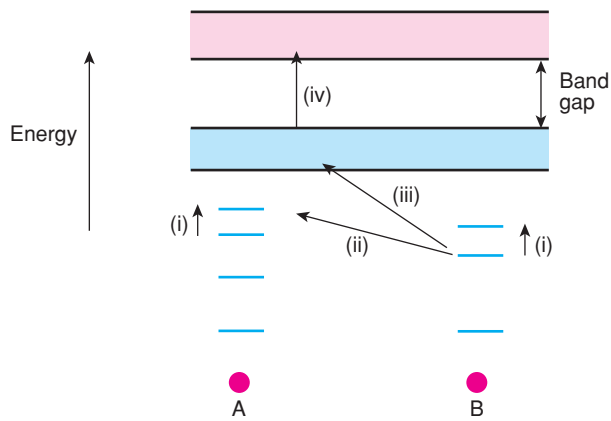


Figure 6.22 ***Possible electronic transitions in a solid. These involve electrons in localised orbitals and/or delocalised bands.***

electrons or holes, e.g. colour centres (F, H, etc.) in alkali halides, and (d) transitions involving Ag atoms in photochromic glasses: colloidal Ag is precipitated initially on photoirradiation and subsequent electronic transitions occur within the Ag particles.

(ii) Promotion of an electron from a localised orbital on one atom to a higher energy localised orbital on an adjacent atom. The associated absorption bands are known as *charge-transfer spectra*. The transitions are usually 'allowed transitions' according to the spectroscopic selection rules and hence the absorption bands are intense. Charge-transfer processes are responsible for the intense yellow colour of chromates; an electron is transferred from an oxygen atom in a $(CrO_4)^{2-}$ tetrahedral anion to the central chromium atom. In mixed valence transition metal compounds, e.g. in magnetite, Fe_3O_4, charge-transfer processes also occur.

(iii) Promotion of an electron from a localised orbital on one atom to a delocalised energy band, the *conduction band*, which is characteristic of the entire solid. In many solids, the energy required to cause such a transition is very high but in others, especially containing heavy elements, the transition occurs in the visible/UV region and the materials are photoconductive, e.g. some chalcogenide glasses are photoconductive.

(iv) Promotion of an electron from one energy band, the valence band, to another band of higher energy, the conduction band. The magnitude of the band gap in semiconductors (Si, Ge, etc.) may be determined spectroscopically; a typical value is 1 eV, 96 kJ mol^{-1}, which lies between the visible and UV regions.

The appearance of a typical UV and visible absorption spectrum is shown schematically in Fig. 6.23. It contains two principal features. Above a certain energy or frequency known as the *absorption edge*, intense absorption occurs; this places a high-frequency limit on the spectral range that can be investigated. To access frequencies above the absorption edge, reflectance techniques must be used. Transitions of types (iii) and (iv) involve an absorption edge. In electronically insulating ionic solids, the absorption edge may occur in the UV region, but in photoconducting and semiconducting materials it may occur in the visible or near-IR region.

The second feature is the appearance of broad absorption peaks or bands at frequencies below that of the absorption cut-off. These are generally associated with type (i) and (ii) transitions.

Visible and UV spectroscopy has a variety of applications associated with the local structure of materials. This is because the positions of the absorption bands are sensitive to coordination environment and bond character. Information on the local structure of amorphous materials may be obtained by adding a small quantity of a spectroscopically active species to the raw materials. This may be a transition metal compound,

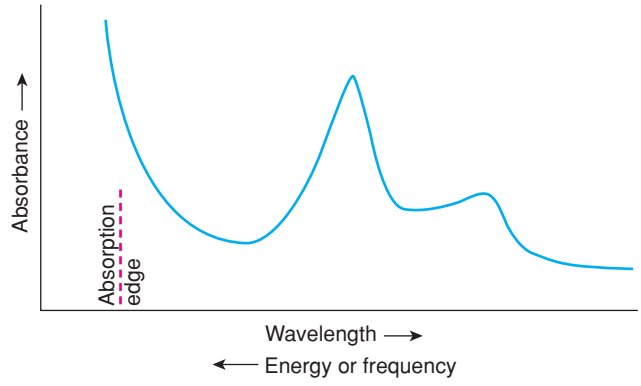


Figure 6.23 ***Schematic of a typical UV/visible absorption spectrum.***

in which case the *d–d* spectra of the transition metal ion are recorded or a heavy *p*-block metal cation such as Tl^{+}, Pb^{2+} for which the $s \rightarrow p$ transitions may be recorded. From the nature of the spectra, coordination numbers can be deduced and information obtained on the availability of such sites in the glass structure.

The spectra of ions such as Pb^{2+} are sensitive to the degree of covalent character in the bonds between Pb and the neighbouring anions. This provides a means for studying the basicity of glasses. Basicity is associated with electron-donating ability. Non-bridging oxide ions which contain a net unit negative charge are highly basic in a silicate glass structure, whereas bridging oxide ions, which link two silicon atoms, contain very little excess negative charge and are not basic. Pb^{2+} ions may therefore be added to a glass in order to probe the basicity of the sites available to it.

Redox equilibria in glasses may be studied. One of the most important is the Fe^{2+}–Fe^{3+} couple since the Fe^{3+} ion is responsible for the green/brown colour of many glasses. For applications such as optical fibre communications, the presence of these ions and their associated spectra is highly undesirable.

Laser materials often contain a transition metal ion as the active species: the ruby laser is essentially Al_2O_3 doped with a small amount of Cr^{3+}; the neodymium laser consists of a glass doped with Nd^{3+} ions. Information on the Cr^{3+} and Nd^{3+} ions and their coordination environment in the host structure may be obtained from UV and visible spectra. Laser action results from a population inversion in which a large number of electrons are promoted into a higher energy level. These subsequently drop back into a lower level and, at the same time, emit the laser beam. It is therefore important to have detailed knowledge of the energy levels and possible electronic transitions. Further details are given in Chapter 10.

UV/visible spectroscopy is widely used in analytical chemistry to determine the concentrations of ions/molecules/functional groups that are responsible for a particular absorption peak. The Beer–Lambert law describes the relationship between the absorbance, A, the molar absorptivity or extinction coefficient, ε, of the species and wavelength responsible for the absorption and the concentration, c, of the absorbing species as follows:

$$\begin{aligned} \mathrm{A} &= \varepsilon c l \\ &= \log(I_0/I) \end{aligned} \tag{6.5}$$

where l is the path length or thickness of the sample, I_0 is the intensity of the incident beam and I is the intensity of the transmitted beam. Hence with suitable calibration, particularly to determine the value of ε that is appropriate for the samples and experimental conditions used, concentrations of species may be determined.

6.3.3 Nuclear magnetic resonance (NMR) spectroscopy

NMR spectroscopy has had an enormous impact on the determination of molecular structure and in recent years has become extremely valuable for providing structural information on solids. Molecular liquids give NMR spectra composed of a number of sharp peaks. From their positions and intensities, it is possible to tell which atoms are bonded together, coordination numbers, next nearest neighbours, and so on. In the solid state, early NMR spectra gave broad, featureless peaks from which little direct structural information was obtained. In the more recent *magic angle spinning* (MAS) technique, the sample is rotated at a high velocity at a critical angle of 54.74° to the applied magnetic field; resulting spectra are composed of sharp peaks yielding much structural information.

NMR spectroscopy utilises the magnetic spin energy of atomic nuclei. For elements with non-zero nuclear spin, such as ^{1}H, ^{2}H, ^{6}Li, ^{7}Li, ^{13}C and ^{29}Si, but not ^{12}C, ^{16}O or ^{28}Si, an applied magnetic field influences the energy of the nuclei. The magnetic energy levels split into two groups, depending on whether the nuclear spins are aligned parallel or antiparallel to the applied magnetic field. The magnitude of the energy difference

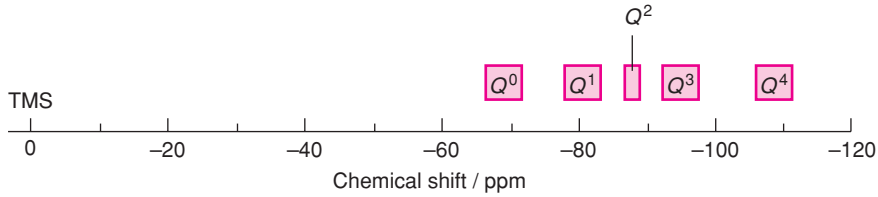


Figure 6.24 ***Positions of ^{29}Si NMR peaks in silicates as a function of the degree of condensation, Q, of the silicate anion. Reproduced with permission from Lippmaa* et al.*, Structural studies of silicates by solid-state high-resolution silicon-29 NMR,* J. Amer. Chem. Soc.*,* 102*, 4889, © 1980 American Chemical Society.***

between parallel and antiparallel spin states is small, ~0.01 J mol^{-1} for an applied magnetic field of 10^4 G (1 T). This amount of energy is associated with the radiofrequency region of the electromagnetic spectrum. NMR spectrometers operating at, for example, 500 MHz can therefore induce nuclear spin transitions by absorption of r.f. energy. The magnitude of the energy change and the associated frequency of absorption depends on the element and its chemical environment. Thus, hydrogen atoms bonded to different types of carbon atom or different functional groups may be distinguished since they absorb at slightly different frequencies. (Note that in practice, NMR instruments usually operate at fixed frequency and the magnitude of the energy difference is varied by varying the applied magnetic field strength.)

Conventional NMR measurements on solids give broad, featureless bands which are of little use for structural work. In the MAS technique, the broad bands collapse to reveal a fine structure. Lippmaa and others first applied this technique to crystalline silicates and found that the ^{29}Si NMR spectra give peaks whose positions depend on the nature of the silicate anion. In particular, the spectra can distinguish between isolated SiO_4 tetrahedra and SiO_4 tetrahedra linked at their corners (oxygen atoms) to one, two, three or four other tetrahedra. It is customary to assign to each silicon (or its SiO_4 tetrahedron) a 'Q value', which represents the number of adjacent SiO_4 tetrahedra to which it is directly bonded. Q values range from zero (as in orthosilicates such as Mg_2SiO_4 with isolated tetrahedra) to four (as in 3D framework structures such as SiO_2, in which all four corners are shared). The positions of the ^{29}Si NMR peaks, i.e. the *chemical shifts* relative to the internal standard [tetramethylsilane, $(CH_3)_4Si$, or TMS], depend approximately on the Q value, as indicated in Fig. 6.24. For each Q value, a range of chemical shifts is observed, depending on other features of the crystal structure.

The ^{29}Si NMR spectrum of a calcium silicate, xonotlite, is shown in Fig. 6.25. The silicate anion is an infinite double chain or ladder with a cross-link or rung at every third tetrahedron in each chain. Two types of silicon are present, therefore, namely Q^2 and Q^3, in the ratio 2:1. The NMR peaks appear in the appropriate positions for Q^2 and Q^3 silicon atoms and their intensities are in the ratio of 2:1, as expected.

NMR spectroscopy can be used to probe the fine structure of aluminosilicate structures. In these, Al atoms play two possible roles: either they occupy octahedral sites and do not really form part of the aluminosilicate framework of linked tetrahedra, or they occupy tetrahedral sites, similar to Si, and thus form part of the framework. In the latter case, the chemical shift of a particular Si depends on how many Al atoms are in its second coordination sphere, Fig. 6.26. For instance, a Q^4 Si in an aluminosilicate framework is surrounded by four other tetrahedral atoms and, of these, any number between zero and four may be Al (i.e. AlO_4 tetrahedra). The chemical shift increases with decreasing number of Al neighbours, from –84 ppm with four Al neighbours, as in nepheline, $KNa_3(AlSiO_4)_4$, to ~−108 ppm with no Al neighbours (as in SiO_2).

The crystal structure of natrolite, $Na_2(Al_2Si_3O_{10})\cdot 2H_2O$, contains two types of Q^4 Si with three and two neighbouring Al atoms, respectively: the ^{29}Si NMR spectrum shows two peaks at –87.7 and –95.4 ppm,

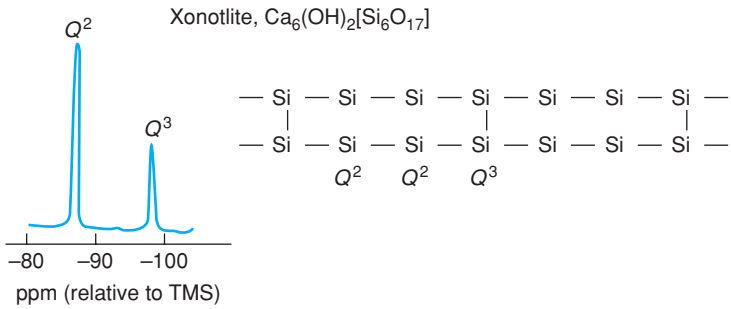


Figure 6.25 ***^{29}Si NMR spectrum of xonotlite. Reproduced with permission from Lippmaa* et al.*, Structural studies of silicates by solid-state high-resolution silicon-29 NMR,* J. Amer. Chem. Soc.*, 102, 4889 © 1980, American Chemical Society. Also shown schematically is the silicate double chain anion and the Q values of the Si atoms involved. Oxygen atoms are omitted.***

Fig. 6.26(c). In some aluminosilicates, the Al, Si distribution is disordered over the tetrahedral sites, as shown by a broadening of the NMR lines. An extreme example is sanidine, $K(AlSi_3O_8)$, in which the Al, Si positions are completely disordered; its NMR signal gives a very broad line, ~15–20 ppm wide at half-height (d).

The information obtained by MAS NMR complements very well the 'average' structural information obtained by XRD. Thus the nature of the Al, Si distribution in silicate structures has long been a thorny

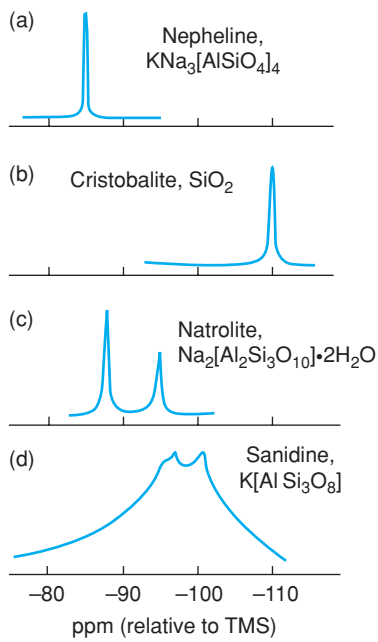


Figure 6.26 ***Schematic ^{29}Si NMR spectra of silicates containing different Q^4 Si atoms. The number of Al atoms in the second coordination sphere are (a) 4, (b) 0, (c) 3 and 2, (d) 2 and 1. Reproduced with permission from Lippmaa* et al.*, Structural studies of silicates by solid-state high-resolution silicon-29 NMR,* J. Amer. Chem. Soc.*, 102, 4889, © 1980 American Chemical Society.***

problem for mineralogists and crystallographers and only in favourable cases have X-ray measurements fully resolved the structure.

6.3.4 Electron spin resonance (ESR) spectroscopy

The ESR technique, also known as *electron paramagnetic resonance*, EPR, is closely related to NMR; it detects changes in electron spin configuration. ESR depends on the presence of permanent magnetic dipoles, i.e. unpaired electrons, such as those which occur in many transition metal ions. The reversal of spin of these unpaired electrons in an applied magnetic field is recorded. The magnitude of the energy change is again small, $\sim$1.0 J mol^{-1}, although somewhat larger than the corresponding NMR energy changes, Fig. 6.19. ESR spectrometers operate at microwave frequencies, e.g. at 2.8×10^{10} Hz (28 GHz), with an applied magnetic field of e.g. 3000 G. In practice, the spectra are obtained by varying the magnetic field at constant frequency. Absorption of energy occurs at the resonant condition:

$$\Delta E = hf = g\beta_c H \tag{6.6}$$

where β_c is a constant, the *Bohr magneton* ($\beta_c = eh/4\pi mc = 9.723 \times 10^{-12}$ J G^{-1}), and H is the strength of the applied magnetic field. The factor g, the *gyromagnetic ratio*, has a value of 2.0023 for a free electron but varies significantly for paramagnetic ions in the solid state. The value of g depends on the particular ion and its oxidation state and coordination number. It is analogous to the 'chemical shift' of NMR spectra.

ESR spectra of solids often show broad peaks (as do NMR spectra) and, therefore, certain conditions must be met in order to obtain sharp peaks from which useful information may be obtained. One source of line broadening is spin–spin interactions between neighbouring unpaired electrons. This is overcome by having a low concentration of unpaired electrons, e.g. 0.1–1% of a paramagnetic transition metal ion dissolved in a diamagnetic host structure. A second source is low-lying excited states near the ground state energy of the paramagnetic ion. This leads to frequent electron transitions, short relaxation times and broad peaks. To overcome this, and to reduce the frequency of the transitions, the spectra are recorded at low temperatures, often at the liquid He temperature, 4.2 K. The interpretation of ESR spectra is greatly simplified if the paramagnetic species has only one unpaired electron, as in, e.g., d^1 transition metal ions, V^{4+} and Cr^{5+}.

ESR spectra are usually presented as the first derivative of the absorption, Fig. 6.27(b), rather than as absorption itself, (a). ESR spectra often comprise a set of closely spaced peaks. The 'hyperfine splitting' which causes multiple peaks arises because, in addition to the applied external magnetic field, there are internal nuclear magnetic fields associated with the transition metal ion itself or with surrounding ligands.

For instance, ^{53}Cr has a nuclear spin moment, I, of 3/2 and the spectrum of $^{53}Cr^{5+}$ (d^1 ion) is split into $2I + 1 = 4$ hyperfine lines. An example is shown in Fig. 6.27(c) for the ESR spectrum of Cr^{5+}, as CrO_4^{3-}, dissolved in apatite, Ca_2PO_4Cl, at 77 K. Naturally occurring Cr is a mixture of isotopes, the main one being ^{52}Cr, $I = 0$. Unpaired electrons on the $^{52}Cr^{5+}$ ions are responsible for the very intense central line, 3, in the spectrum. The four small, equally spaced lines 1, 2, 4 and 5 are associated with unpaired electrons on the small percentage of $^{53}Cr^{5+}$ ions which have $I = 3/2$.

In some cases, additional line structure, 'superhyperfine splitting', may occur due to interactions of an unpaired electron with the nuclear moment of neighbouring ions.

From ESR spectra such as Fig. 6.27, one can (with appropriate experience!) obtain information on the paramagnetic ion and its immediate environment in the host structure. Specifically, one may determine:

(a) the oxidation state, electronic configuration and coordination number of the paramagnetic ion;
(b) structural distortions arising from, for example, a Jahn–Teller effect;
(c) the extent of any covalency in bonds between the paramagnetic ion and its surrounding anions or ligands.

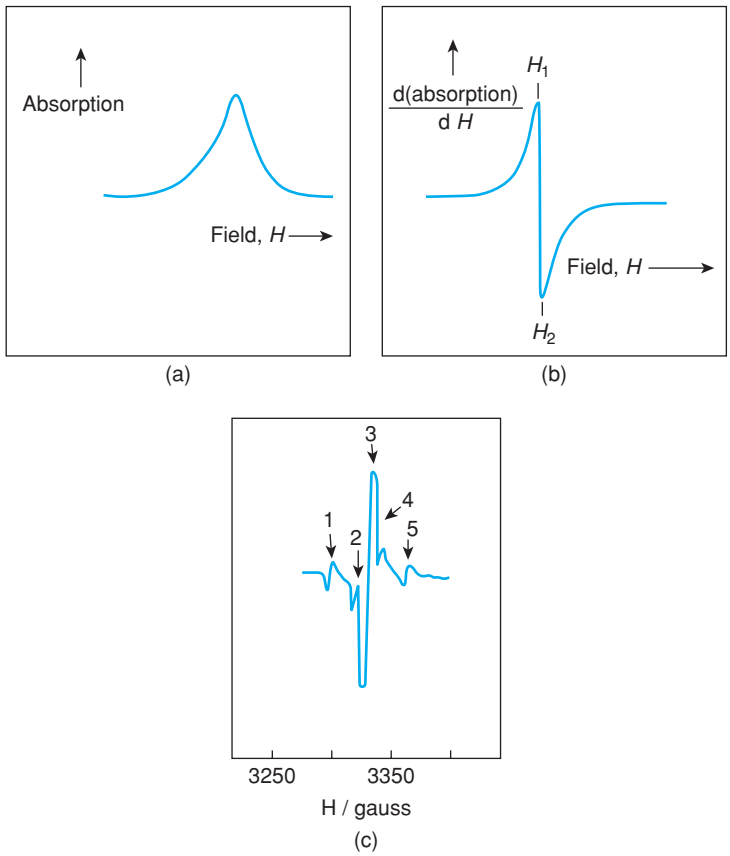


Figure 6.27 ***(a) Schematic ESR absorption peak, (b) the first derivative of the absorption and (c) the ESR spectrum of*** CrO_4^{3-} ***in*** Ca_2PO_4Cl ***at 77 K. Reproduced with permission from Greenblatt, M., Electron spin resonance of tetrahedral transition metal oxyanions (*** MO_4^{n-}***) in solids,* J. Chem. Educ.*, 57, 546, © 1980 American Chemical Society.***

Since the paramagnetic ion is present in dilute amounts, it is assumed to have the same site symmetry as the host ion that it substitutionally replaces. For instance, in Cr-doped apatites, $M_5(PO_4)_3X$: M = Ca, Sr, Ba; X = Cl, F, the ESR spectra show that the CrO_4^{3-} tetrahedron is compressed along one of the fourfold inversion axes of the tetrahedron. It is very likely, therefore, that the PO_4 tetrahedra have a similar distortion. By contrast, in $BaSO_4$, the SO_4 tetrahedra are distorted by stretching along a fourfold inversion axis as shown by the ESR spectra of $BaSO_4$ doped with MnO_4^{2-} containing Mn^{6+}, also a d^1 ion.

A technique related to ESR, which is useful for observing hyperfine and superhyperfine splittings, is *electron nuclear double resonance* (ENDOR). Basically, it is a combination of NMR and ESR spectroscopy. In ENDOR, the NMR frequencies of nuclei adjacent to a paramagnetic centre are scanned. The resulting NMR spectrum shows fine structure due to interactions with the paramagnetic ion. By scanning in turn each of the nuclei in the region of the paramagnetic species, a detailed plan of the local atomic structure is obtained. The method has been applied successfully to the study of crystal defects, such as colour centres (trapped electrons in alkali halide crystals), radiation damage effects and doping in, for example, phosphors and semiconductors.

6.3.5 X-ray spectroscopies: XRF, AEFS, EXAFS

The X-ray region of the electromagnetic spectrum is extremely useful for structural studies, analysis and characterisation of solids. There are three main ways in which X-rays are used: for diffraction, emission spectroscopies and absorption spectroscopies, Fig. 6.28(a).

Diffraction techniques utilise a monochromatic beam of X-rays which is diffracted by the sample and used for both crystal structure determination and phase identification (see Chapter 5).

Emission techniques utilise the characteristic X-ray emission spectra of elements generated by, for instance, bombardment with high-energy electrons. The spectra are used for chemical analysis of both bulk samples (*X-ray fluorescence*) and micron- or submicron-sized particles (*analytical electron microscopy*, AEM, *electron probe microanalysis*, EPMA, etc.). They also have some uses in determination of local structure and coordination numbers.

Absorption techniques measure the absorption of X-rays by samples, especially at energies in the region of *absorption edges*. They are powerful techniques for studying local structure but usually require a high-energy X-ray source such as a synchrotron and are mainly confined to national facilities.

6.3.5.1 Emission techniques

On bombardment of matter with high-energy electrons, inner shell electrons may be ejected from atoms. Outer shell electrons then drop into the vacancies in the inner levels and the excess energy is released in the form of electromagnetic radiation, often X-radiation (Fig. 5.3a and Fig. 6.5). Each element gives a characteristic X-ray emission spectrum (Fig. 5.3b) composed of a set of sharp peaks. The spectra are different for each element since the peak positions depend on the difference in energy between electron levels, e.g. $2p$ and $1s$, which, in turn, depend on atomic number [Moseley's law, equation (5.1)]. X-ray emission spectra may be used for elemental analysis, both qualitatively by looking for peaks at certain positions or wavelengths, and quantitatively by measuring peak intensities and comparing them against a calibration chart. *X-ray fluorescence* (XRF) does just this: a solid sample is bombarded with high-energy electrons and the resulting emission spectrum is recorded. From the spectral peak positions, the elements present can be identified, and from their intensities, a quantitative analysis can be made. XRF is an important analytical technique, especially in industry. With fully automated modern instruments, large numbers of solid samples can be analysed.

Analytical electron microscopy techniques such as EPMA and EDAX (or EDS) operate in the same way as XRF. The sample is bombarded with high-energy electrons; the diffracted electrons are used to record both electron micrographs and electron diffraction patterns, whereas the emitted X-rays are used for elemental analysis. AEM techniques are invaluable for chemical analysis of solids, in particular because very small particles or regions of a solid may be characterised. The heterogeneous nature of many solids, such as cements, steels and catalysts, may therefore be studied.

X-ray emission spectra also have uses in determining local structure such as coordination numbers and bond distances. This is because peak positions vary slightly with local environment of the atoms in question. As an example, the Al $K\beta$ emission bands (corresponding to the $3p \rightarrow 1s$ transition) for three widely different materials, Al metal, α-Al_2O_3 and sanidine, $KAlSi_3O_8$, are shown in Fig. 6.29. The peaks occur in different positions and for Al_2O_3 and sanidine appear to be poorly resolved doublets. By comparing the spectra for a range of Al-containing oxides, correlations between (a) peak position and coordination number of Al (it is usually 4 or 6) and (b) peak position and Al–O bond distance for a given coordination number have been found. Similar studies on silicates (in which Si is tetrahedrally coordinated) have shown a correlation between the Si $K\beta$ peak position and Si–O bond length; since the Si–O length varies according to whether the oxygen bridges two silicate tetrahedra or is non-bridging, this may be used to study polymerisation of silicate anions in, for example, glasses, gels and crystals.

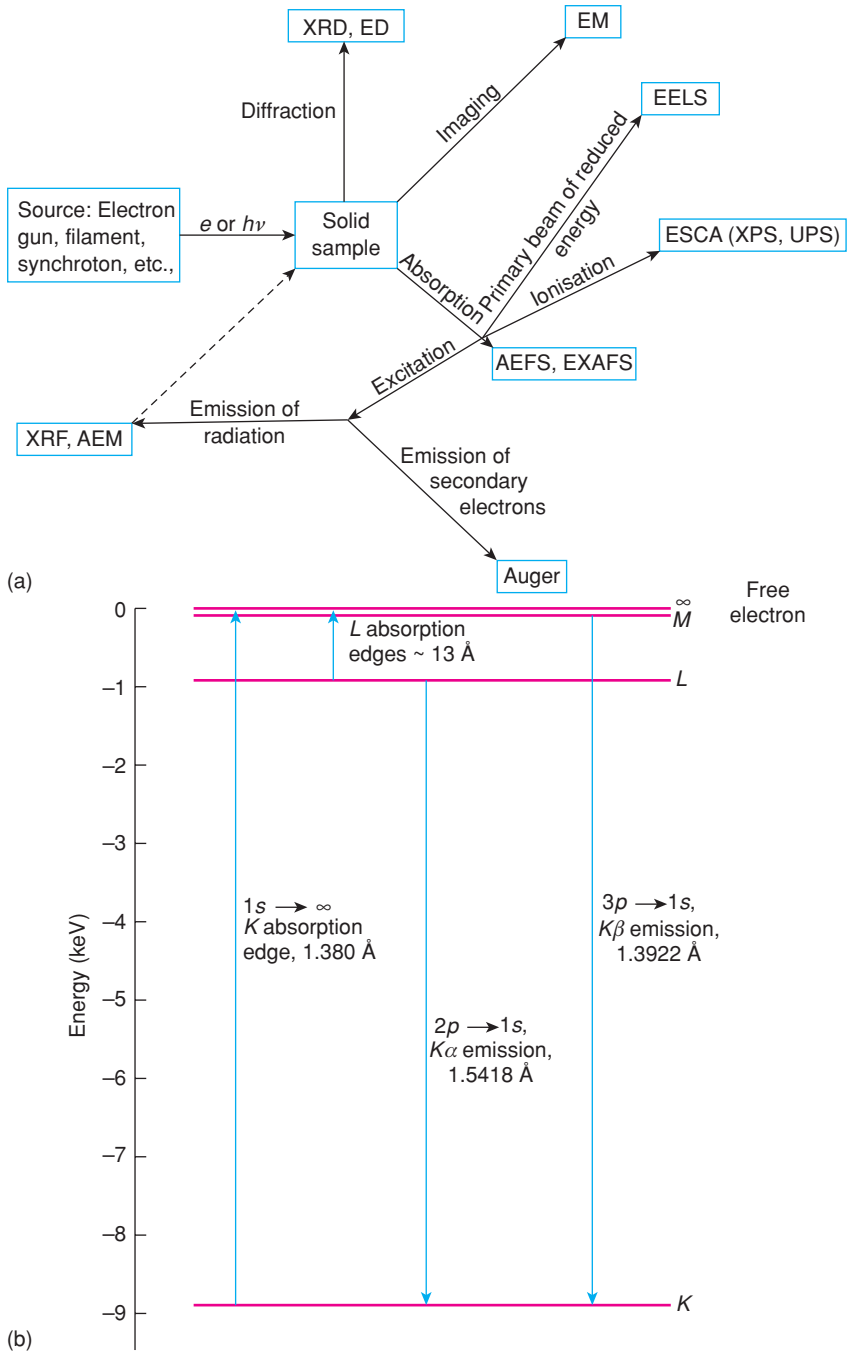


Figure 6.28 ***(a) Flowsheet showing relations between various diffraction, microscopic and spectroscopic techniques associated with electrons and X-rays. (b) Electronic transitions responsible for emission and absorption X-ray spectra. Wavelength values are for Cu.***

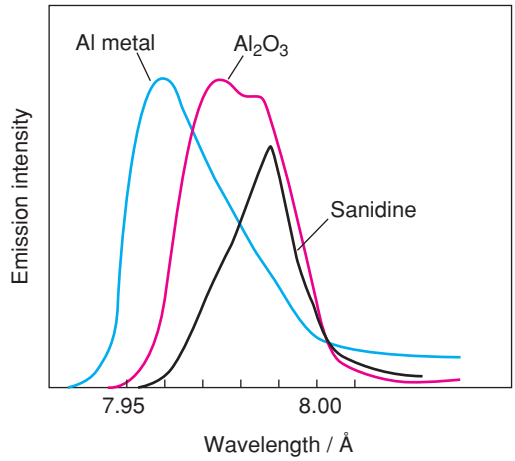


Figure 6.29 ***Al Kβ emission spectra of three Al-containing materials. Reproduced with permission from White and Gibbs,*** **Amer. Mineral.,** *54, 931,* ***© 1969 The Mineralogical Society of America.***

For elements that can exist in more than one valence state or oxidation state, a correlation usually exists between peak position and oxidation state. Thus, XRF may distinguish the different oxidation states of, for example, S (–II to +VI) in a variety of S-containing compounds.

6.3.5.2 Absorption techniques

Atoms give characteristic X-ray absorption spectra in addition to characteristic emission spectra. These arise from various ionisation and intershell transitions, as shown in Fig. 6.28(b). In order to ionise a 1*s* electron in Cu metal, an energy of almost 9 keV, corresponding to a wavelength of 1.380 Å, is needed. Much less energy (∼1 keV) is needed to ionise *L* shell (2*s*, 2*p*) electrons. The X-ray absorption spectrum of Cu is shown in Fig. 6.30. It takes the form of a smooth curve which increases rapidly at low X-ray energies and superposed on which are *K*, *L*, etc., absorption edges. The *K* edge represents the minimum energy required to ionise a 1*s* electron in Cu. The absorption coefficient therefore undergoes an abrupt increase as the energy is increased to this value. The *L* edge (in fact, three closely spaced *L* edges are usually seen) represents ionisation of 2*s*, 2*p* electrons. At intermediate energies of the incident X-ray beam, e.g. at 4 keV, electrons may be ionised from *L* or outer shells and leave the atom with a net kinetic energy, *E*, given by

$$E = h\nu - E_0 \tag{6.7}$$

where $h\nu$ is the energy of the incident X-ray photon (4 keV in this case) and E_0 is the critical ionisation energy required to free the electron from the atom. The wavelengths at which absorption edges occur depend on the relative separation of the atomic energy levels in atoms which, in turn, depend on atomic number as given by Moseley's law. They are, therefore, characteristic for each element and may be used for identification purposes in a similar manner to emission spectra.

Although X-ray absorption techniques have been used since the 1930s, recent developments using synchrotrons and storage rings as a source of X-rays have given them a new impetus. Synchrotron radiation, produced when electrons are accelerated in a magnetic field, is a very intense spectrum of continuous X-ray wavelengths. Using it, much higher sensitivity is achieved and more information can be obtained from the

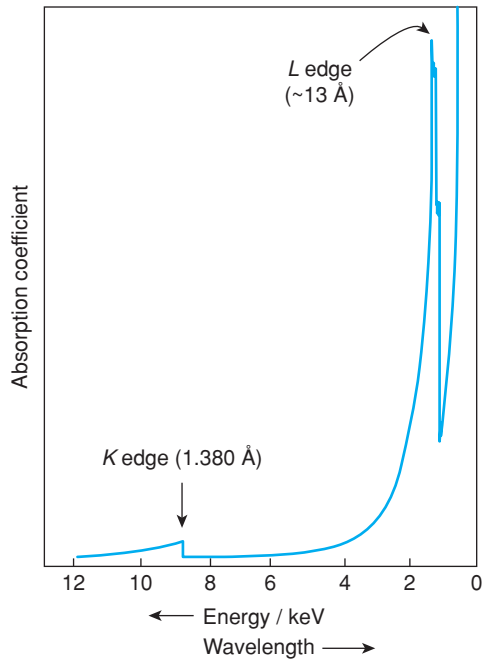


Figure 6.30 ***Variation of X-ray absorption coefficient with wavelength for Cu. Reproduced with permission from E. A. Stern 'The Analysis of Materials by X-Ray Absorption'* Scientific American*, 234(4):96 © 1976 Scientific American, a division of Nature America, Inc.***

absorption spectra. The number of laboratories that have a particle accelerator suitable for producing synchrotron radiation is, of course, very limited; there is only one, DIAMOND, in the UK. Others are at Orsay, France, Stanford, USA, and the Tsukuba Photon Factory, Japan. The main techniques are *absorption edge fine structure* (AEFS), also called *X-ray absorption near edge structure* (XANES), and *extended X-ray absorption fine structure* (EXAFS), both of which monitor the absorption spectra in the region of absorption edges.

XANES

In the region of an absorption edge, fine structure associated with inner shell transitions is often seen, e.g. the *K* edge of Cu may show additional peaks due to transitions $1s \rightarrow 3d$ (but not in Cu^{1+} compounds), $1s \rightarrow 4s$, $1s \rightarrow 4p$. The exact peak positions depend on details of oxidation state, site symmetry, surrounding ligands and the nature of the bonding. The absorption spectra may therefore be used to probe local structure. Examples are shown in Fig. 6.31, on an expanded energy or wavelength scale (compared to Fig. 6.30), for two Cu compounds, CuCl and $CuCl_2{\cdot}2H_2O$. For each, the *K* edge, with superposed fine structure peaks ($1s \rightarrow 4p$, etc.) is seen. The entire spectrum is displaced to higher energies in $CuCl_2{\cdot}2H_2O$, reflecting the higher oxidation state of Cu (+2) in $CuCl_2{\cdot}2H_2O$ compared with that (+1) in CuCl and, therefore, the increased difficulty in ionising *K* shell electrons.

EXAFS

Whereas XANES examines at high resolution the fine structure in the region of an absorption edge, EXAFS examines the variation of absorption with energy (or wavelength) over a much wider range, extending out

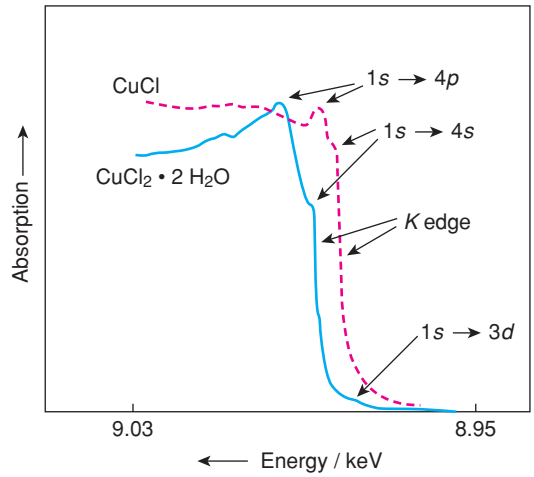


Figure 6.31 ***AEFS spectra of CuCl and $CuCl_2{\cdot}2H_2O$. Reproduced with permission from Chan, Hu and Gamble, Application of synchrotron radiation to the structural studies of metalloproteins.,* J. Mol. Str.*, 45, 239, © 1978 Elsevier.***

from the absorption edge to higher energies by up to ~1 keV. The absorption usually shows a ripple, known also as the *Kronig fine structure*, Fig. 6.32, from which information on local structure and, especially, bond distances may be obtained. Explanations of the origin of the ripple will not be attempted here. Suffice it to say that it is related to the wave properties of the ionised photoelectrons which interact with neighbouring atoms in the solid; these then act as secondary sources of scattering for the photoelectrons. Interference between adjacent scattered waves may occur and this influences the probability of absorption of an incident X-ray photon. The degree of interference depends on the wavelength of the photoelectron (and hence on the wavelength of the incident X-ray photons) and the local structure, including interatomic distances, in the region of the emitting atom. EXAFS is therefore a kind of *in situ* electron diffraction in which the source of the electron is the actual atom which participates in the X-ray absorption event. Using Fourier transform

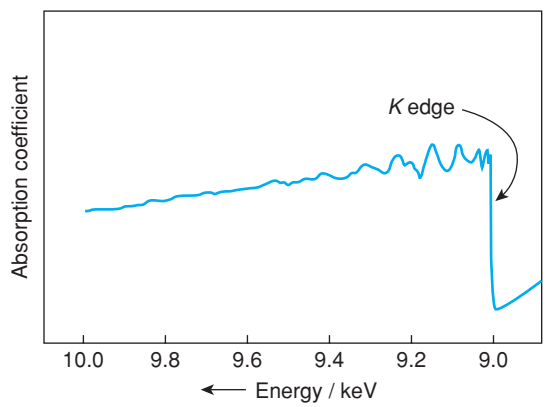


Figure 6.32 ***EXAFS spectrum of Cu. Reproduced with permission from E. A. Stern 'The Analysis of Materials by X-Ray Absorption'* Scientific American*, 234(4):96 © 1976 Scientific American, a division of Nature America, Inc.***

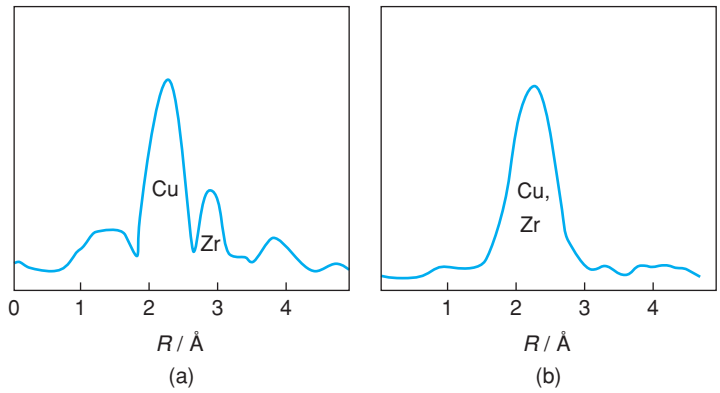


Figure 6.33 ***EXAFS-derived partial RDFs for an amorphous $Cu_{46}Zr_{54}$ alloy: (a) Zr K edge; (b) Cu K edge. Reproduced with permission from Gurman, S. J., EXAFS studies in materials science,* J. Mat. Sci., *17, 1541, © 1982 Springer Netherlands.***

techniques, it is possible to analyse the ripple pattern and obtain information on coordination numbers and bond distances.

EXAFS is equally suitable for non-crystalline and crystalline materials since it is a technique for probing local structure. It is particularly valuable for studying glasses, gels and amorphous metals for which structural information is generally hard to obtain. This is because EXAFS has one great advantage: by tuning in to the absorption edge of each element present in the material in turn, the local structure around each element may be determined. By contrast, conventional diffraction techniques give only a single averaged coordination environment for all the elements present.

An example is shown in Fig. 6.33 for the alloy $Cu_{46}Zr_{54}$. The plots are Fourier transforms derived from (a) the Zr *K* edge at 18 keV and (b) the Cu *K* edge at 9 keV. The positions of the peaks are related to, but not directly equal to, interatomic distances. From these plots, it was shown that each Zr atom is surrounded by an average of 4.6 Cu atoms at 2.74 Å and 5.1 Zr atoms at 3.14 Å; Cu–Cu distances are 2.47 Å.

6.3.6 Electron spectroscopies: ESCA, XPS, UPS, AES, EELS

Electron spectroscopy techniques measure the kinetic energy of electrons that are emitted from matter as a consequence of bombarding it with ionising radiation or high-energy particles. Various processes take place when atoms are exposed to ionising radiation, Fig. 6.28(a). The simplest is ionisation of an electron from either a valence or an inner shell. The kinetic energy, E, of the ionised electron is equal to the difference between the energy, $h\nu$, of the incident radiation and the binding energy or ionisation potential, E_0, of the electron, equation (6.7). For a given atom, a range of E_0 values is possible, corresponding to the ionisation of electrons from different inner and outer valence shells, and these E_0 values are characteristic for each element.

Measurement of E, and therefore E_0, provides a means of identification of atoms and forms the basis of the ESCA (*electron spectroscopy for chemical analysis*) technique developed by Siegbahn and co-workers in Uppsala (1967). The ionising radiation is usually either X-rays (Mg $K\alpha$, 1254 eV, or Al $K\alpha$, 1487 eV) or UV light (He discharge, 21.4 and 40.8 eV for the $2p \rightarrow 1s$ transitions in He and He^+, respectively) and the techniques are then also known as XPS (*X-ray photoelectron spectroscopy*) and UPS (*ultraviolet photoelectron spectroscopy*), respectively. The main difference between XPS and UPS is in the electron shells that are accessible for ionisation. Inner shell electrons may be ionised in XPS but only outer electrons – in valence shells, molecular orbitals or energy bands – may be ejected in UPS.

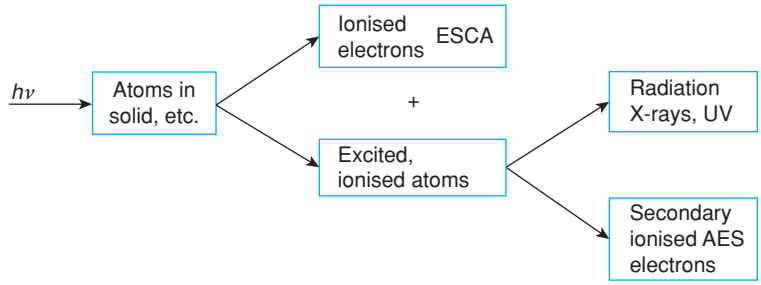


Figure 6.34 ***Origins of ESCA and Auger spectra.***

A related technique is *Auger electron spectroscopy* (AES), as described in Section 6.2.2.3. In this, the electrons that are ejected and detected are not the primary ionised electrons, as in ESCA, but are produced by secondary processes involving the decay of ionised atoms from excited states to lower energy states, Fig. 6.34.

The ESCA and Auger processes may be visualised as involving

$$\text{Atom A} \xrightarrow{h\nu} \text{A}^{+*} + \text{e}^{-}$$

where A^{+*} refers to an ionised atom which is in an excited state and e^{-} is the ionised electron which is detected in ESCA. The excited state condition arises either if the electron is ejected from an inner shell, leaving a vacancy, or if other electrons in the atom have been promoted to higher, normally empty levels, during irradiation. Either way, the excited atom decays when electrons drop into vacancies in lower energy levels. Energy is consequently released by one of two methods:

$$\text{A}^{+*} \rightarrow \text{A}^{+} + h\nu \text{ (X-rays, UV)}$$

or

$$\text{A}^{+*} \rightarrow \text{A}^{2+} + \text{e}^{-} \text{ (Auger electrons)}$$

The energy may be emitted as electromagnetic radiation: this is, in fact, the normal method by which X-rays are produced, although for lighter atoms, UV photons are generated instead. Alternatively, the energy may be transferred to another (outer shell) electron in the same atom which is then ejected. Such secondary ionised electrons are known as Auger electrons.

AES spectra are usually observed at the same time as ESCA spectra. A schematic example is shown in Fig. 6.35 as a plot of intensity of ionised electrons against energy. Often, AES spectra are complex and difficult to interpret. Although AES is not, as yet, a widely used technique, this may change in the future.

XPS and UPS are powerful techniques for determining energy levels in atoms and molecules. They are particularly useful for studying surfaces because the electrons that are produced are not very energetic (usually their energy is $\ll 1$ keV) and are rapidly absorbed by solid matter. Consequently, they cannot escape from solids unless they are ejected within ~20–50 Å (2–5 nm) of the surface.

Some success has been achieved in using XPS to probe the local structure of solid surfaces. This is possible because the binding energies of electrons may vary depending on the immediate environment of the particular atom and its charge or oxidation state. The idea is to measure the 'chemical shift' of an atom, relative to a standard, and thereby obtain information on local structure (as in NMR). Examples of chemical shifts

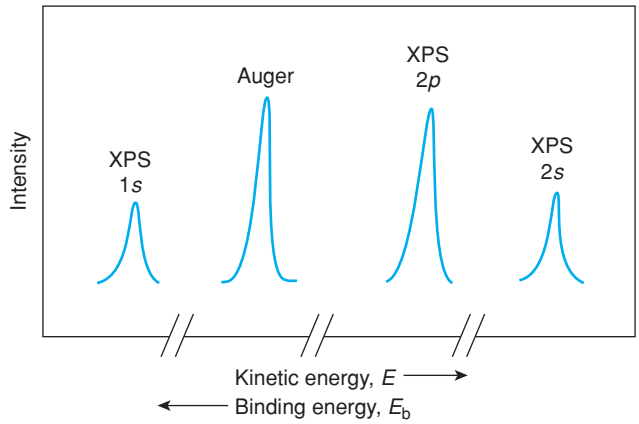


Figure 6.35 ***Schematic XPS and AES spectra of Na^+ in a sodium-containing solid. The Auger peak arises from an initial 1s vacancy which is filled by a 2p electron, causing ionisation of another 2p (Auger) electron.***

are shown in Fig. 6.36 and Fig. 6.37. In sodium thiosulfate, $Na_2S_2O_3$, the two types of S atom may be distinguished. Peaks are of equal height, indicating equal numbers of each. Assignment of the peaks of higher kinetic energy to the terminal S atom is made on the basis that this atom carries more negative charge than the central atom and is, therefore, easier to ionise. Hence E_0 is less and $h\nu - E_0$ is greater for the terminal S atom than for the central S atom. By comparison, sodium sulfate, Na_2SO_4, shows a single S 2*p* peak at the same energy as that for the central S atom in $Na_2S_2O_3$.

KCr_3O_8 is a mixed valence compound better written as $KCr^{III}(Cr^{VI}O_4)_2$. Its XPS spectrum, Fig. 6.37, shows doublets for Cr 3*s* and 3*p* electrons. The intensities are in the ratio 2:1 and the peaks are assigned to the oxidation states Cr^{VI} and Cr^{III}. This fits with the formula and with the expectation that E_0 is greater for Cr^{VI} than for Cr^{III}.

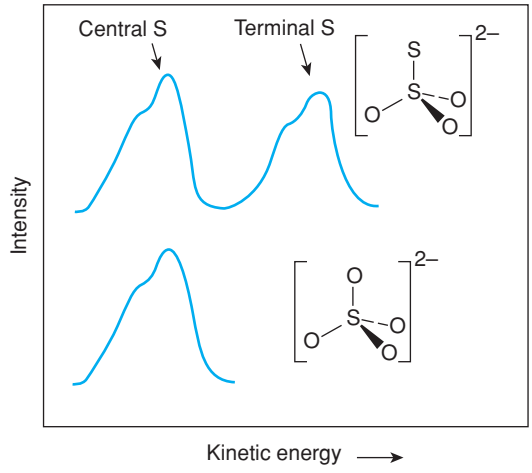


Figure 6.36 ***Schematic XPS 2p spectra of $Na_2S_2O_3$ and Na_2SO_4. Each peak is a doublet because of the two spin–orbit states, 1/2 and 3/2, of the 2p electron.***

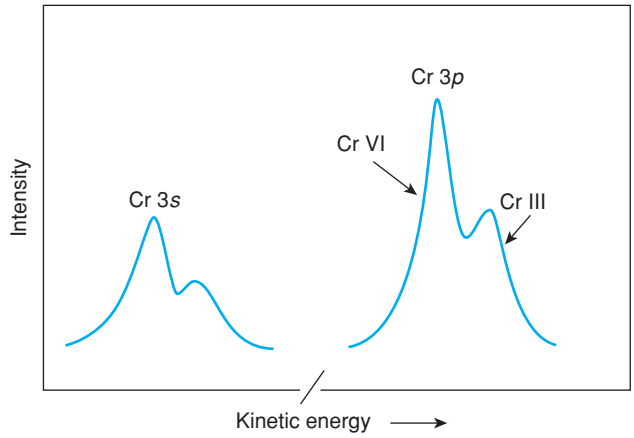


Figure 6.37 ***XPS spectrum of Cr 3s, 3p electrons in KCr_3O_8.***

These two examples show clearly the influence of local structure on the ESCA spectra. In many other cases, the chemical shifts associated with different oxidation states or local environments are small and insensitive to local structure. The promise of ESCA as a local structure probe, applicable to virtually all elements of the Periodic Table, is therefore as yet only partly fulfilled.

ESCA is widely used to investigate the band structures of solids, especially metals and semiconductors. An interesting example is shown in Fig. 6.38 for a series of sodium tungsten bronzes. These have the ReO_3

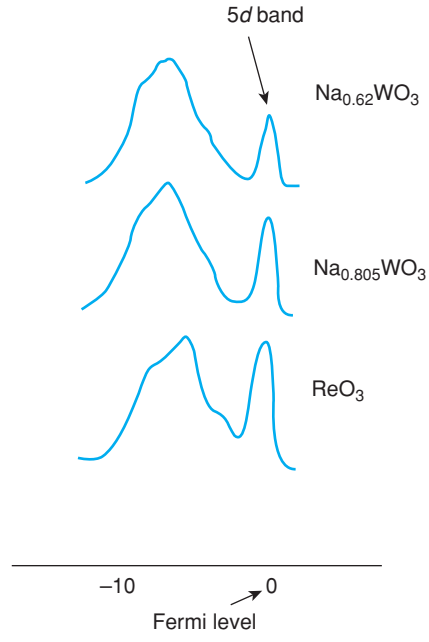


Figure 6.38 ***XPS spectra for tungsten bronzes. Reproduced with permission from Campagna* et al.*, Phys. Rev. Lett.*, 34*, 738, © 1975 American Physical Society.***

structure but contain variable amounts of intercalated Na. W^{VI} is d^0 and hence, ideally, WO_3 should have no 5*d* electrons. On addition of Na to WO_3, to form Na_xWO_3, electrons from Na enter the W 5*d* band. This shows up in the XPS spectrum as a peak representing ionisation from the 5*d* band and whose intensity increases with increasing Na content. The spectrum of ReO_3 is shown for comparison. Re^{VI} is d^1 and hence its spectrum is similar to that which might be expected for $Na_{1.0}WO_3$. Sometimes, localised surface states in the electron band structure of a metal can be detected. These appear as electron orbitals localised at or near the surface, rather than delocalised throughout the bulk of the metal. In order to detect them, very clean surfaces are needed.

EELS or ELS is a technique associated with analytical electron microscopy and is discussed in Section 6.2.2.6. It can be used for elemental analysis, for light elements such as C and N and for studying energy levels in surface layers of solids. In EELS, a monochromatic incident beam of electrons causes ionisation of inner shell electrons in atoms of the sample. This ionisation is a necessary precursor to the generation of X-rays. The electrons that are responsible for the ionisation suffer an energy loss as a consequence. An EELS spectrum is a plot of intensity of these electrons against energy loss. An intense peak occurs at zero energy loss. This corresponds to electrons that are either scattered elastically or do not interact with the sample. The other peaks in the EELS spectrum correspond to the electrons responsible for inner shell ionisations. The peaks are usually weak and broad and the spectra increase in complexity with increasing atomic number. EELS is particularly useful for analysing light atoms and therefore complements X-ray fluorescence which is more useful with heavier elements.

6.3.7 Mössbauer spectroscopy

Mössbauer or γ-ray spectroscopy is akin to NMR spectroscopy in that it is concerned with transitions that take place inside atomic nuclei. The incident radiation is a highly monochromatic beam of γ-rays whose energy may be varied by making use of the Doppler effect. Absorption is monitored as a function of energy and a spectrum obtained which usually consists of a number of poorly resolved peaks, from which information on local structure – oxidation states, coordination numbers and bond character – may be obtained.

The γ-rays are produced by decay of radioactive elements such as $^{57}Fe^*_{29}$ or $^{119}Sn^*_{50}$. The γ-rays appear in the electromagnetic spectrum, Fig. 6.18, to the high-energy side of X-rays. The γ-emission is associated with a change in population of energy levels in the nuclei responsible, rather than with a change in atomic mass or number. Under certain conditions of 'recoilless emission', all of this energy change is transmitted to the emitted γ-rays. This gives rise to a highly monochromatic beam of radiation which may be absorbed by a sample that contains similar atoms to those responsible for the emission. The nuclear energy levels of the absorbing atoms vary somewhat, depending on oxidation state, coordination number, etc., and some means of modulating either the energy of the incident radiation or the energy levels within the nuclei of the sample is required.

In practice, the energy of the incident γ-rays is modified by making use of the Doppler effect. With the sample placed in a fixed position, the γ-ray source is moved at a constant velocity either towards or away from the sample. This has the effect of increasing or decreasing the energy of the γ-rays incident upon the sample. In this way, the γ-ray absorption spectrum of the sample is determined. Few isotopes can be induced to emit γ-rays suitable for Mössbauer work. The most widely used are ^{57}Fe and ^{119}Sn and, hence, Mössbauer spectroscopy has been most used on iron- and tin-containing substances. Other isotopes which have been profitably studied include ^{129}I, ^{99}Ru and ^{121}Sb. A schematic illustration of the technique is given in Fig. 6.39.

Several types of information may be obtained from Mössbauer spectra. In the simplest case, in which emitter and sample are identical, a resonant absorption peak occurs when the source is stationary, Fig. 6.39(b). When

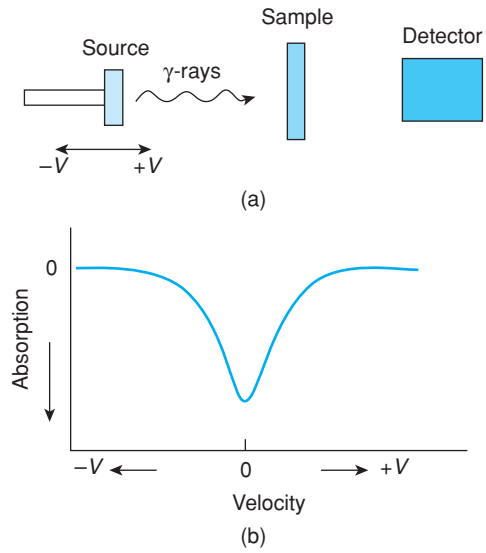


Figure 6.39 ***(a) The Mössbauer effect. The energy of the γ-rays is modified when emitted from a moving source (Doppler effect). (b) Typical single line spectrum obtained when source and sample are identical.***

the emitter and sample are not identical, the absorption peak is shifted. This *chemical shift*, δ, arises because the nuclear energy levels have been modified by changes in the extra-nuclear electron density distribution in the atoms. In particular, chemical shifts may correlate with the density at the nucleus of the outer shell *s* electrons. A summary given in Fig. 6.40 of the chemical shifts in various Fe-containing solids shows the influence of both charge and coordination number on the chemical shift. Use may be made of this for diagnostic purpose, e.g. to determine the nature of Fe in inorganic compounds.

For nuclei that have a nuclear spin quantum number $I > {}^1/_2$, the distribution of positive charge inside the nucleus is non-spherical and a quadrupole moment, Q, results. The net effect of this is to cause splitting of the nuclear energy levels and hence splitting of the peaks in the Mössbauer spectrum. For both ^{57}Fe and ^{119}Sn,

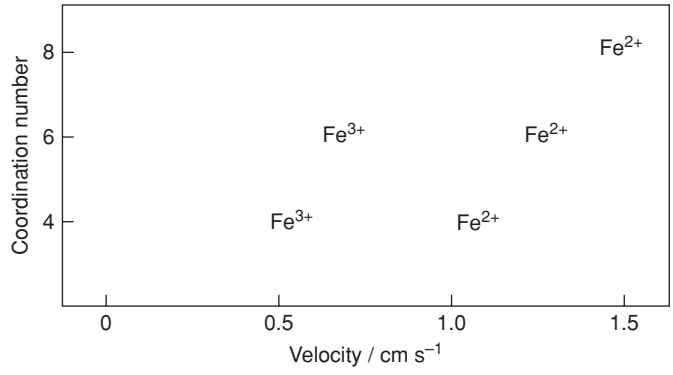


Figure 6.40 ***Chemical shifts in iron-containing compounds. Reproduced with permission from Bancroft,* Mössbauer Spectroscopy, *© 1973 McGraw-Hill.***

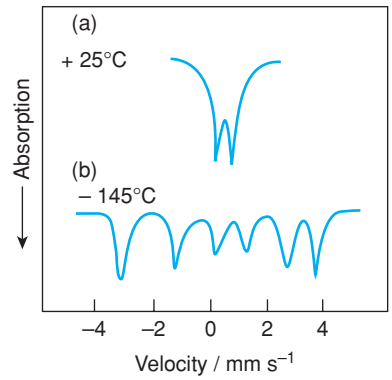


Figure 6.41 ***Mössbauer spectrum of $KFeS_2$ (a) above and (b) below the Néel temperature. Reproduced with permission from Greenwood,* Chem. Brit.*, 3, 56, © 1967 Royal Society of Chemistry.***

the peaks are split into doublets. The separation of the doublets, known as the *quadrupole splitting*, Δ, is sensitive to local structure and oxidation state.

A second type of line splitting is the *magnetic hyperfine Zeeman splitting*. This arises when a nucleus, of spin I, is placed in a magnetic field; each nuclear energy level splits into $2I + 1$ sublevels. The magnetic field arises either from magnetic exchange effects in ferro-, antiferro- and paramagnetic samples or from an externally applied field. The study of hyperfine splitting and, especially, its temperature dependence provides information on magnetic ordering. For instance, $KFeS_2$ is antiferromagnetic below 245 K and hyperfine splitting results in a spectrum containing six peaks. Above 245 K, only quadrupole coupling occurs and the spectrum reduces to a doublet, Fig. 6.41.

6.4 Thermal Analysis (TA)

Thermal analysis, TA, involves the measurement of certain physical and chemical properties as a function of temperature. These are mainly enthalpy, heat capacity, mass and coefficient of thermal expansion. Measurement of the coefficient of thermal expansion of metal bars is a simple example of TA. Another is measurement of the change in weight of oxysalts or hydrates as they decompose on heating. Uses of TA in solid state science are many and varied and include the study of solid state reactions, thermal decompositions, phase transitions and the determination of phase diagrams. Most solids are 'thermally active' in one way or another and may be studied profitably by TA.

The main TA techniques are *thermogravimetry* (TG), which records the mass of a sample as a function of temperature or time, and *differential thermal analysis* (DTA), which measures the difference in temperature, ΔT, between a sample and an inert reference material as a function of temperature; DTA therefore detects changes in heat content. A technique closely related to DTA is *differential scanning calorimetry* (DSC). In DSC, the equipment is designed to allow a quantitative measure of the enthalpy changes that occur. In *dilatometry*, the change in linear dimension of a sample with temperature is recorded. It has long been used to measure the coefficients of thermal expansion of metals; recently, it has acquired a new name, *thermomechanical analysis* (TMA), and has been applied to more diverse materials and problems, e.g. to the quality control of polymers.

Here, the basic principles of TG, DTA and DSC are described, together with applications of each; instrumental descriptions are omitted.

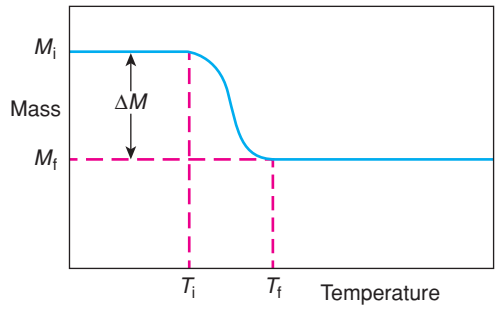


Figure 6.42 ***Schematic TG curve for a single-step decomposition reaction.***

6.4.1 Thermogravimetry (TG)

TG measures the change in mass of a substance as a function of temperature or time. The results appear as a continuous record as shown in Fig. 6.42. The sample is heated at constant rate and has a constant mass, M_i, until it begins to decompose at temperature T_i. Under conditions of dynamic heating, decomposition usually takes place over a range of temperatures, T_i to T_f, and a second constant-mass plateau is observed above T_f, which corresponds to the mass of the residue M_f. The masses M_i and M_f and the difference ΔM are fundamental properties of the sample and can be used for quantitative calculations of compositional changes, etc. By contrast, the temperatures T_i and T_f depend on variables such as heating rate, nature of the solid (e.g. its particle size) and atmosphere above the sample. The effect of atmosphere can be dramatic, as shown in Fig. 6.43 for the decomposition of $CaCO_3$: in vacuum, decomposition is complete by ~500 °C, but in CO_2 at atmospheric pressure, decomposition does not even commence until above 900 °C. T_i and T_f pertain to the particular experimental conditions, therefore, and do not necessarily represent equilibrium decomposition temperatures.

6.4.2 Differential thermal analysis (DTA) and differential scanning calorimetry (DSC)

In DTA, the temperature of a sample is compared with that of an inert reference material during a programmed change of temperature. The temperatures should be the same until some thermal event, such as melting, decomposition or change in crystal structure, occurs, in which case the sample temperature either lags behind (if the change is endothermic) or leads (if the change is exothermic) the reference temperature.

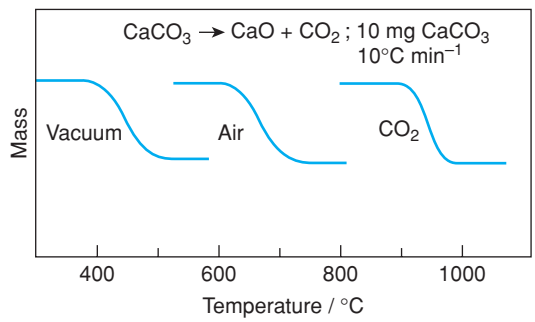


Figure 6.43 ***Decomposition of $CaCO_3$ in different atmospheres.***

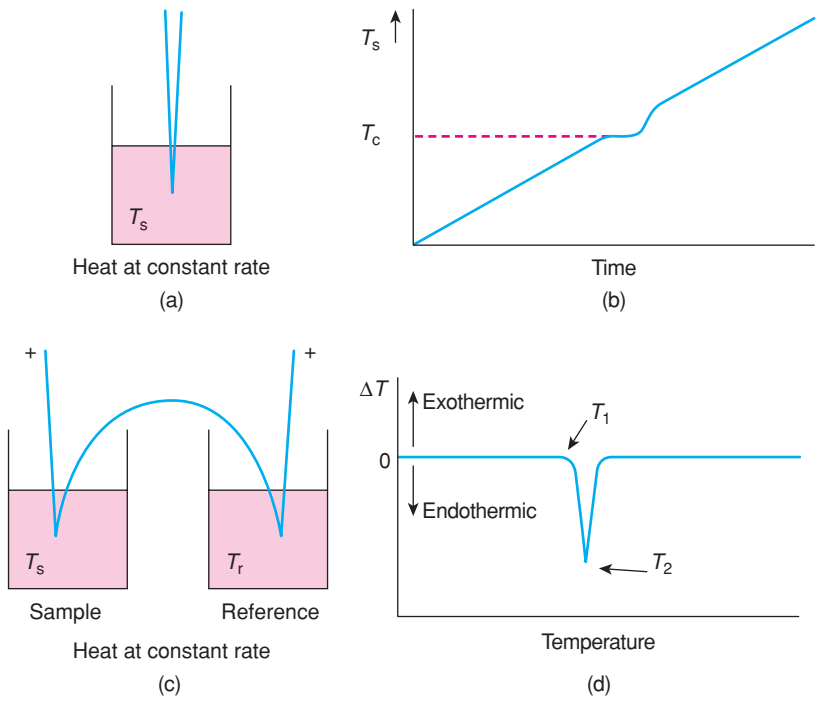


Figure 6.44 ***The DTA method. Graph (b) results from the set-up shown in (a) and graph (d), a typical DTA trace, results from the arrangement shown in (c).***

The reason for having both a sample and a reference is shown in Fig. 6.44. In (a), a sample is heated at a constant rate; its temperature, T_s, monitored with a thermocouple, varies linearly with time (b) until an endothermic event occurs, e.g. melting at temperature T_c. The sample temperature remains constant at T_c until the event is completed; it then increases rapidly to catch up with the temperature required by the programmer. The thermal event at T_c therefore appears as a rather broad deviation from the sloping baseline in (b).

Such a plot is insensitive to small heat effects. Further, any spurious baseline variations, caused by, for example, fluctuations in the heating rate, would appear as apparent thermal events. Because of its insensitivity, this technique has limited applications; its main use historically has been in the 'method of cooling curves', which was used to determine phase diagrams: the sample temperature was recorded on cooling rather than on heating and since heat effects associated with solidification and crystallisation are usually large, they could be detected by this method.

In Fig. 6.44(c), the arrangement used in DTA is shown. Sample and reference are placed side-by-side in a heating block which is heated or cooled at a constant rate; identical thermocouples are placed in each and are connected 'back-to-back'. When the sample and reference are at the same temperature, the net output of this pair of thermocouples is zero. When a thermal event occurs in the sample, a temperature difference, ΔT, exists between sample and reference, which is detected by the net voltage of the thermocouples. A third thermocouple (not shown) is used to monitor the temperature of the heating block and the results are presented as ΔT versus temperature (d). A horizontal baseline, corresponding to $\Delta T = 0$, occurs and superposed on this is a sharp peak due to the thermal event in the sample. The temperature of the peak is taken either as the temperature at which deviation from the baseline begins, T_1, or as the peak temperature, T_2. Although it is probably more correct to use T_1, it is often not clear where the peak begins and, therefore, it is more common

to use T_2. The size of the ΔT peak may be amplified so that events with very small enthalpy changes may be detected. Figure 6.44(d) is clearly a much more sensitive and accurate way of presenting data than (b) and is the normal method for presenting DTA (and DSC) results.

DTA cells are designed for maximum sensitivity to thermal changes, but this can be at the expense of losing a calorimetric response in which peak areas or peak heights are related only qualitatively to the magnitude of the enthalpy changes occurring. It is possible to calibrate DTA equipment so that quantitative enthalpy values can be obtained, but it is better to use DSC for quantitative measurements.

DSC is very similar to DTA. A sample and an inert reference are also used but the cell design is different. In some DSC cells, the sample and reference are maintained at the same temperature during heating and the extra heat input to the sample (or to the reference if the sample undergoes an exothermic change) required in order to maintain this balance is measured. Enthalpy changes are therefore measured directly. In other DSC cells, the difference in temperature between sample and reference is measured, as in DTA, but by careful attention to cell design, the response is calorimetric.

6.4.3 Applications

Uses of TA are many and varied. Generally, DTA/DSC is more versatile than TG: TG detects effects which involve mass changes only. DTA/DSC also detects such effects but, in addition, detects other effects, such as polymorphic transitions, which do not involve changes in mass. For many problems, it is advantageous to use both DTA/DSC and TG together, because the DTA/DSC events can then be classified into those which do and those which do not involve mass change. An example is the decomposition of kaolin, $Al_4(Si_4O_{10})(OH)_8$, Fig. 6.45. By TG, a change in mass occurs at ~500–600 °C, which corresponds to dehydration of the sample; this dehydration shows up on DTA as an endotherm. A second DTA effect occurs at 950–980 °C, which has no counterpart in the TG trace: it corresponds to recrystallisation of the dehydrated kaolin. This effect is exothermic, which means that the structure obtained between ~600 and 950 °C is metastable; the exotherm signals a decrease in enthalpy of the sample and therefore a change to a more stable structure. The details of the structural changes that occur at this transition have still not been fully resolved.

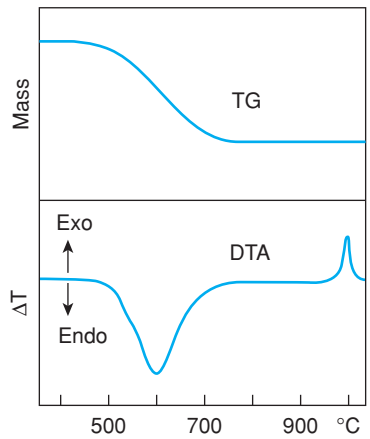


Figure 6.45 ***Schematic TG and DTA curves for kaolin minerals. Curves vary depending on the sample structure and composition, e.g. the TG weight loss and associated DTA endotherm can occur anywhere in the range 450–750 °C.***

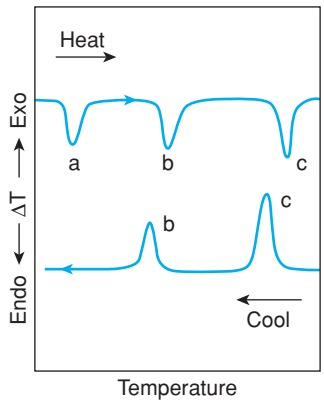


Figure 6.46 ***Some schematic reversible and irreversible changes: (a) dehydration, (b) polymorphic transition and (c) melting/solidification.***

Another useful ploy is to follow the thermal changes on cooling as well as on heating. This allows a separation of reversible changes, such as melting/solidification, from irreversible changes, such as most decomposition reactions. A schematic DTA sequence illustrating reversible and irreversible changes is shown in Fig. 6.46. Starting with a hydrated material, the first event on heating is dehydration (a), which requires heat input and is detected as an endotherm. Subsequently, the dehydrated material undergoes a polymorphic transition, which is also endothermic (b), at some higher temperature. Finally, the sample melts, giving a third endotherm (c). On cooling, the melt crystallises, as shown by an exotherm (c); the polymorphic change also occurs, exothermically (b), but rehydration does not occur. The diagram therefore shows two reversible and one irreversible process.

On studying reversible processes, which are detected on both heating and cooling, it is common to observe *hysteresis*; the exotherm on cooling may be displaced to lower temperatures than the corresponding endotherm on heating. Ideally, the two processes should occur at the same temperature, but hystereses ranging from a few degrees to several hundred degrees are commonplace. The two reversible changes in Fig. 6.46 are shown with small hystereses.

Hysteresis depends on the nature of the material and the structural changes involved; transitions involving the breaking of strong bonds are likely to exhibit much hysteresis. Hysteresis occurs particularly on rapid cooling; in some cases, if the cooling rate is sufficiently fast, the change can be suppressed completely. The change is then effectively irreversible under those particular conditions. An example of great industrial importance is associated with glass formation, as shown schematically in Fig. 6.47. Beginning with a crystalline substance, e.g. silica, an endotherm appears when the substance melts. On cooling, the liquid does not crystallise but becomes supercooled; as the temperature drops, its viscosity increases until eventually it becomes a glass. Hence crystallisation has been suppressed entirely. In the case of SiO_2, the liquid is very viscous even above the melting point, $\sim$1700 °C, and crystallisation is slow, even at slow cooling rates.

An important use of DTA and DSC in glass science is to measure the glass transition temperature, T_g. This appears as a broad anomaly in the baseline of the DTA curve, as shown in Fig. 6.47; T_g represents the temperature at which the glass transforms from a rigid solid to a supercooled, very viscous, liquid. The glass transition represents the upper temperature limit at which the glass can be used and also provides a readily measurable parameter for characterising glasses. For glasses which kinetically are very stable, such as silica glass, the glass transition is the only thermal event observed on DTA since crystallisation is usually too sluggish to occur. For other glasses, however, crystallisation or devitrification may occur at some temperature

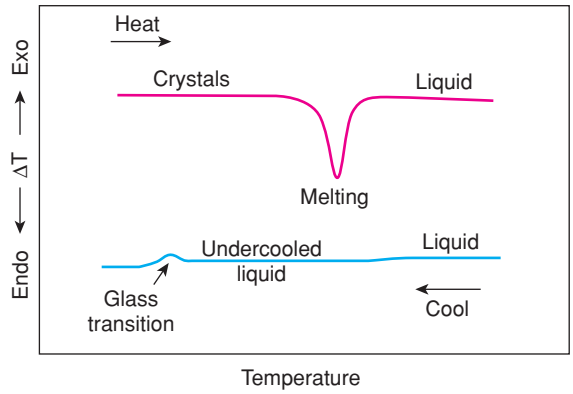


Figure 6.47 ***Schematic DTA curves showing melting of crystals on heating and large hysteresis on cooling, which gives rise to glass formation.***

above T_g and below the melting point, T_f. On heating, devitrification appears as an exotherm and is followed by an endotherm at a higher temperature that corresponds to melting of these same crystals. Examples of glasses which readily devitrify are metallic glasses; they may be prepared as thin films by rapidly quenching certain liquid alloy compositions. Other important glass-forming materials are amorphous polymers and amorphous chalcogenide semiconductors.

Polymorphic phase transitions may be studied easily and accurately by DTA or DSC; since many properties may be modified or changed completely as a consequence of a phase transition, their study is extremely important and useful. Examples are as follows:

(a) Ferroelectric $BaTiO_3$ has a Curie temperature, T_c, of $\sim$120 °C, which may be determined by DTA; substitution of other ions for Ba^{2+} or Ti^{4+} causes the Curie temperature to vary. For many substitutions, T_c may vary by tens, if not hundreds, of degrees and therefore measurement of T_c can be an accurate indicator of composition.
(b) In cements, the β-polymorph of Ca_2SiO_4 has superior cementitious properties to the γ-polymorph. The effect of different additives on the transitions can be studied by DTA.
(c) In refractories, transitions such as $\alpha \rightarrow \beta$ quartz or quartz $\rightarrow$ cristobalite have a deleterious effect on silica refractories because volume changes associated with each transition reduce the mechanical strength of the refractory. These transitions, which should be prevented from occurring if possible, may be monitored by DTA.

DTA/DSC is a powerful method for the determination of phase diagrams (see Chapter 7), especially in conjunction with other techniques, such as X-ray diffraction for phase identification. Its use is illustrated in Fig. 6.48(b) for two compositions in the simple binary eutectic system shown in (a). On heating composition A, melting begins at the eutectic temperature, T_2, and gives an endothermic peak. However, this is superposed on a much broader endotherm which terminates approximately at temperature T_1, and which is due to the continued melting that occurs over the temperature range T_2 to T_1. For this composition, an estimate of both solidus, T_2, and liquidus, T_1, temperatures may therefore be made. Composition B corresponds to the eutectic composition. On heating, this transforms completely to liquid at the eutectic temperature, T_2, and gives a single, large endotherm, peaking at T_2, on DTA.

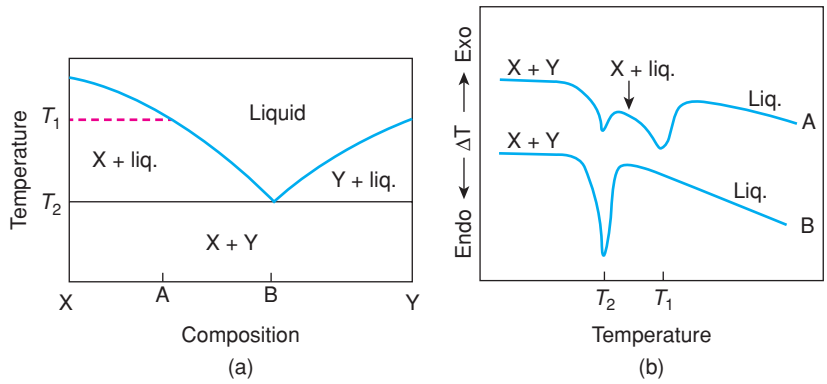


Figure 6.48 ***Use of DTA for phase diagram determination: (a) a simple binary eutectic system; (b) schematic DTA traces for two compositions, A and B, on heating.***

Thus, if DTA traces for a range of mixtures between X and Y are compared, all should give an endotherm at T_2, the magnitude of which depends on the degree of melting at T_2 and hence on the closeness of the sample composition to the eutectic composition, B. In addition, all compositions, apart from B, should give a broad endotherm at some temperature above T_2 due to the completion of melting. The temperature of this peak should vary with composition. Polymorphic phase transitions also appear on phase diagrams, at subsolidus temperatures, and these may be determined very readily by DTA, especially if solid solutions form and the transition temperature is composition dependent.

In multistage decomposition processes, TG, either alone or in conjunction with DTA, may be used to separate and determine the individual steps. A good example, the decomposition of calcium oxalate monohydrate, is illustrated in Fig. 6.49; decomposition occurs in three stages giving, as intermediates, anhydrous calcium oxalate and calcium carbonate. Many other examples of multistage decompositions occur with hydrates, hydroxides, oxysalts and minerals.

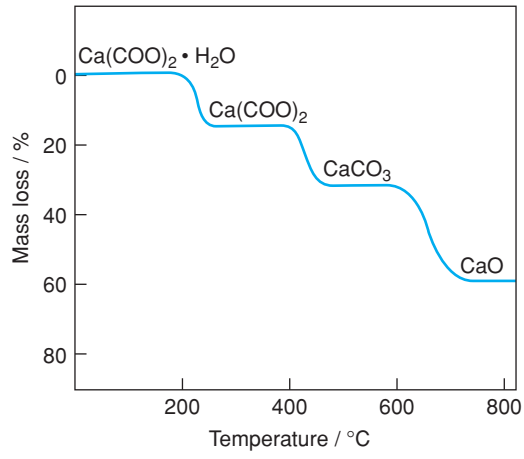


Figure 6.49 ***Schematic, stepwise decomposition of calcium oxalate hydrate by TG.***

6.5 Strategy to Identify, Analyse and Characterise 'Unknown' Solids

You have a solid material that you wish to identify, analyse and characterise; let us assume that it is a white powder, but it could be a solid piece or a deposit on a surface. Let us also assume that you have all the analytical techniques available and help with running the equipment and interpreting results.

The first point to make is that there is no single technique that can answer all your questions. Powder XRD is undoubtedly the most generally useful technique to obtain a preliminary identification but is unable to answer many follow-up questions about composition, structure, defects and their distribution.

The main techniques for the task in hand have been covered in Chapter 5 and the present chapter. In Table 6.1, an overview of the techniques that can be used to answer specific questions about our 'unknown' solid is presented. Even though one person cannot be an expert in more than a small number of the wide range of possible techniques, it is, nevertheless, important to appreciate what each technique can offer and be able to ask the right questions.

Table 6.1 is not all-encompassing. It does not include any property measurement techniques and their uses for the characterisation of solids; for instance, magnetic measurements can very usefully complement the results obtained by Mössbauer spectroscopy on unpaired electron spin configurations in transition metal-containing materials; similarly, electrical properties of materials may be very sensitive to the effect of dopants and variable oxidation state of transition metal-containing compounds. Optical property measurements can also give information on oxidation states and d,f electron configurations. These various property measurements also come into the category of structure–property relations, which would usually be outside the scope of our target to identify and characterise an unknown solid.

Table 6.1 ***Strategy and overview for identification, analysis and structure determination of solids***

Technique	Information sought	Notes
	1. What is it (in general terms?)	
Optical microscopy	• Is it amorphous/glassy or crystalline? • Are the particles at least micron-size or submicron/nano?	An invaluable, quick, first-look technique. If the solid is crystalline, follow-up characterisation of crystal symmetry is possible: are all the particles similar or is there a phase mixture?; are the crystals twinned?; is the crystal quality good (sharp extinction)?. If particles are submicron size, move on to SEM
Powder X-ray diffraction	• Crystalline phase identification; phase purity	The most important general technique for fingerprinting solids. Should tell what crystalline phase(s) is present, but not its chemical composition. Can identify individual phases in a mixture. Requires access to database and assumes that all possible phases are in the database
	2. What is its chemical composition?	
Atomic absorption spectroscopy; inductively-coupled plasma mass spectrometry	• Bulk chemical analysis of dissolved sample	Gives overall chemical composition but no information on phase identity or purity. Dissolution prior to analysis may be easy or may require fusion in e.g. Na_2CO_3 or B_2O_3, followed by dissolution
X-ray fluorescence	• Bulk chemical analysis of solids	A non-destructive technique although the sample may be embedded within a glass matrix. Also, only gives overall composition. Not suitable for light elements
C, H, N analysis	• Composition of organic solids or organic component of organometallic compounds	Relevant to molecular organic-based materials. Linked to spectroscopic techniques (NMR, mass spectrometry, IR) for structure determination. Separate branch of chemistry; not considered further here
Electron probe microanalysis; energy-dispersive analysis of X-rays; high-angle, annular darkfield/*Z*-contrast imaging	• Elemental analysis of individual particles from micron to ångström-level	For micron-sized particles, use EPMA for quantitative analysis (excluding H, Li). For semiquantitative analysis to nanometre level, use SEM–EDX/EDS. For atomic-level analysis, use *Z*-contrast STEM
	3. What is its structure?	
Single-crystal X-ray diffraction	• Crystal structure: atomic coordinates of unit cell contents	The standard technique for structure determination. For molecular materials, gives both solid state packing arrangement and molecular configuration. Crystals as small as 1–10 μm can be studied with synchrotron XRD
Powder X-ray and neutron diffraction (ND)	• Crystal structure: atomic coordinates of unit cell contents	An increasingly powerful technique using synchrotron XRD but one needs access to national facilities for this and for ND. Assumes either prior knowledge of unit cell and space group or determination of these during earlier stages of data processing; not always possible, therefore. ND also gives magnetic structure
Electron diffraction: selected area and convergent beam	• Unit cell and space group	Invaluable aid to structure determination if unit cell cannot be determined *ab initio* by powder XRD/ND. Full structure determination and refinement not usually possible with ED as diffraction intensities are unreliable

High-resolution and *Z*-contrast electron microscopy	• Lattice imaging	Under favourable circumstances, lattice planes can be imaged by HREM. Invaluable for shear structures, layered intergrowths and homologous series of phases. *Z*-contrast imaging gives 2D structural projections directly, without focusing problems, at atomic level. Crystallites need to be oriented correctly
X-ray spectroscopies	• Local structure information on bulk materials	Diffraction techniques give crystal structure information averaged over many unit cells. This is complemented by local structure information given by a wide range of spectroscopic techniques including XANES (for oxidation states) and EXAFS (for coordination numbers of specific elements and bond lengths)
Mössbauer spectroscopy	• Coordination numbers and oxidation states of selected elements	Very powerful local structure probe but limited to a few elements: Fe, Sn, I
Electron spectroscopies	• Local structure information especially at surfaces	The shallow escape depth for low-energy electrons gives surface-specific information. XPS provides rapid overview data for solid samples. EELS and Auger are electron microscope-based. EELS gives atomic-level bonding information when used in conjunction with *Z*-contrast imaging. Auger is particularly useful for light elements which give simpler spectra
NMR spectroscopy	• Element-specific local structure information	Provides bulk information using the solid state, MAS technique. Particularly useful for amorphous materials such as aluminosilicate and borosilicate glasses and for characterising intermediate-range structural features (such as next nearest neighbours)
ESR spectroscopy	• Element-specific local structure information	Specific to paramagnetic species; very high sensitivity
IR and Raman spectroscopy	• Identification of functional groups	For non-carbon-containing materials, IR is useful for identifying functional groups such as oxyanions and H-species. Raman has uses for phase fingerprinting of e.g. C polymorphs
	4. Is it stable to temperature/atmosphere?	
Thermogravimetry	• Determine thermal stability, susceptibility to atmospheric attack	TG detects weight loss from specific compounds such as hydrates or carbonates or the effects of atmospheric attack, mainly by H_2O, CO_2. Link TG to MS for evolved gas analysis. Drying conditions for chemical reagents can be established. Avoid possible instrument damage and contamination by separate prior heating tests on samples to maximum TG temperature. Use of TG to determine oxygen stoichiometries by H-reduction TG
Differential thermal analysis and differential scanning calorimetry	• Determine melting points, phase transitions and their modification with dopants/impurities	DSC/DTA detect thermal events on heat/cool cycles. Combine with TG to separate effects involving weight loss. Determine reversibility of changes on heat/cool cycles.

7

Phase Diagrams and their Interpretation

Phase diagrams summarise in graphical form the ranges of temperature (and sometimes pressure) and composition over which phases or mixtures of phases are stable under conditions of thermodynamic equilibrium. They therefore contain information on the compositions of compounds and solid solution series, phase transitions and melting or decomposition temperatures. For inorganic solids that have practical or potential applications, it is vital to know whether they are stable in contact with their environment and whether they react or transform into another solid(s) during use. Thus, phase diagrams have been key to the use and development of inorganic materials in technologies such as glass, refractories, cement and ceramics. This applies not only to the conditions under which these materials are used but also to their synthesis and fabrication.

Phase diagrams tell us the final equilibrium products of a particular reaction, although they give no information on reaction kinetics. Nevertheless, it can be useful to know in which direction a reaction should proceed or whether a particular reaction product has reached equilibrium. Knowledge of an equilibrium phase diagram enables us to understand the effects of temperature, pressure and composition on reaction conditions. If phase diagram information is unavailable, for instance, if we have prepared a new material for the first time, we may wish to start collecting information that would enable us to construct a phase diagram and devise tests to establish whether or not a reaction is complete. The main purpose of this chapter is to explain the use and interpretation of phase diagrams. In the last section, some guidelines for constructing new phase diagrams are given.

The fundamental rule upon which phase diagrams are based is the *phase rule* of W. J. Gibbs. Its derivation is given later but first the phase rule is presented and the terminology explained.

7.1 The Phase Rule, the Condensed Phase Rule and Some Definitions

The *phase rule* is given by the equation

$$P + F = C + 2 \tag{7.1}$$

where P is the number of phases present in equilibrium, C is the number of components needed to describe the sample or system under consideration and F is the number of degrees of freedom or independent variables

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

taken from temperature, pressure and composition of the phases present. In many cases, pressure is not a variable and then the *condensed phase rule* is used:

$$P + F = C + 1 \tag{7.2}$$

Each term is now explained more fully.

The word *system* is used to specify the range of compositions of interest. Thus, when referring to the 'system MgO–SiO_2', for example, we are interested in the end-member phases MgO and SiO_2, in addition to all intermediate compositions including the magnesium silicate phases $MgSiO_3$ and Mg_2SiO_4. The system must be completely self-contained and all the phases that are encountered in the various single- and multi-phase regions and also on changing temperature or pressure must have compositions that can be described in terms of their MgO and SiO_2 contents.

If, instead, we are interested in iron silicates such as Fe_2SiO_4 rather than magnesium silicates, we may find that attention to the oxidation state of Fe is necessary during sample synthesis. Oxygen partial pressure then becomes important and we may refer to the system as FeO–SiO_2–O or Fe–SiO_2–O. This then allows us to include oxides such as Fe_2O_3 and Fe_3O_4, and also FeO (which is, itself, a complex material: Chapter 2).

The number of *phases*, *P*, in a sample at equilibrium is the number of physically distinct and mechanically separable (in principle) portions, each phase being itself *homogeneous*. The distinction between different *crystalline* phases is usually clear. For example, the differences between chalk, $CaCO_3$, and sand, SiO_2, are obvious. The distinction between crystalline phases made from the same components but of different composition is also usually clear. Thus, the magnesium silicate minerals enstatite, $MgSiO_3$, and forsterite, Mg_2SiO_4, are different phases with different compositions, crystal structures and properties.

With solids it is also possible to obtain different crystalline phases having the *same* chemical composition. This is known as *polymorphism*. For example, two polymorphs of Ca_2SiO_4 can be prepared at room temperature, the stable γ form and the metastable β form, but these have distinct physical and chemical properties and crystal structures. This example underpins the entire construction industry, since the β form is a key component of cement, which reacts with water to form concrete. The γ form, by contrast, shows little reactivity towards water and therefore care is needed in cement manufacture to avoid the formation of the unreactive γ polymorph. A topical example of polymorphism is shown by carbon, which, depending on the experimental conditions, can be prepared as diamond, graphite, C_{60}, carbon nanotubes and, recently, graphene.

One complicating but very important factor in classifying solid phases is the occurrence of solid solutions (see Chapter 2): a solid solution is a single phase that has variable composition. For example, α-Al_2O_3 and Cr_2O_3 have the same crystal structure (corundum) and form a continuous range of solid solutions at high temperature. In these, a mixture of Al_2O_3 and Cr_2O_3 forms a single, homogeneous phase, the structure of which is the same as that of the Al_2O_3 and Cr_2O_3 end members but in which the cation sites contain a random mixture of Cr^{3+} and Al^{3+}.

Sometimes, as with crystallographic shear structures (see Chapter 2), it may be difficult to decide exactly what constitutes a separate phase. This is because a small change in composition can lead to a different, non-random arrangement of defects in a structure. In the oxygen-deficient tungsten oxides, WO_{3-x}, what was previously thought to be a range of homogeneous solid solutions, with a random distribution of oxygen vacancies, is now known to be a large number of phases that are very close in composition and similar, but distinct, in structure. Some have formulae belonging to the homologous series W_nO_{3n-1}. Thus $W_{20}O_{59}$ and $W_{19}O_{56}$ are physically distinct phases! With the majority of solids, however, such complexities do not occur and there is no problem in deciding exactly what constitutes a phase.

In the *liquid* state, the number of possible, separate homogeneous phases that can exist is much more limited. This is because single-phase liquid solutions form very readily and over wider compositional ranges

than do single-phase solid solutions. Take the Na_2O–SiO_2 system, for instance. In the liquid state at high temperatures, Na_2O and SiO_2 are completely miscible to give a single, liquid, sodium silicate phase of variable composition. In the crystalline state, however, the number of phases is quite large: crystalline sodium silicate phases form at five different compositions and at least one of these shows polymorphism. In other systems, such as MgO–SiO_2, mixtures of two liquid phases, or *liquid immiscibility*, can exist at high temperatures and over a certain range of compositions.

In the *gaseous* state, the maximum number of possible phases appears always to be one; there are no known cases of immiscibility between two gases, if the effects of gravity are ignored.

The number of *components*, *C*, of a system is the number of constituents that can undergo independent variation in the different phases; alternatively, it is the *minimum* number of constituents needed in order to describe completely the compositions of the phases present. This is best understood with the aid of examples:

(a) All of the calcium silicates can be considered as built from CaO and SiO_2 in varying proportions. CaO–SiO_2 is therefore a *two-component system* even though there are three elements present, Ca, Si and O. Alternatively, we can say that CaO and SiO_2 form a *binary* (i.e. two-component) *join* in the *ternary system* Ca–Si–O, Fig. 7.1. Ternary phase diagrams are beyond the scope of this chapter and we simply note that compositionally, they can be represented as equilateral triangles as shown in Fig. 7.1.

(b) The system MgO is *unary* (*one-component*), at least up to the melting point 2700 °C, because the composition of MgO is always fixed; we do not need to consider this as the binary system Mg–O if we are only interested in the phase MgO. Of course, if we are interested in the reaction of Mg metal with oxygen, and the effects of temperature/pressure/composition, then we may consider the binary system Mg–O.

(c) The composition 'FeO' is part of the two-component system, Fe–O, because wüstite is actually a non-stoichiometric, iron-deficient phase, $Fe_{1-x}O$, caused by having some Fe^{3+} present (see Chapter 2). The bulk composition 'FeO' in fact contains a mixture of two phases at equilibrium: $Fe_{1-x}O$ and Fe metal; proper consideration of 'FeO' requires us to think in terms of the binary system Fe–O.

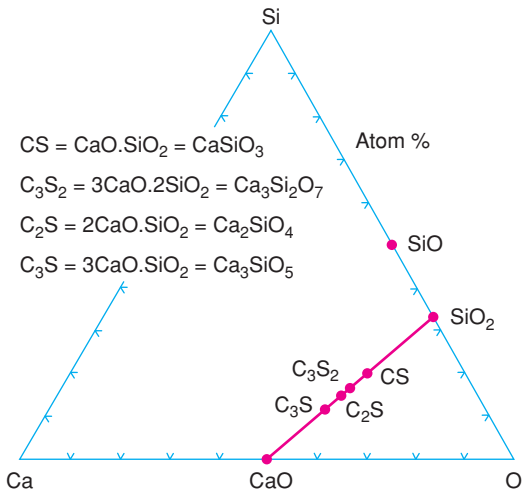


Figure 7.1 ***Binary join CaO–SiO_2 in the ternary system Ca–Si–O. Note the method used for the labelling of phases, C = CaO, etc. This type of abbreviation is widely used in oxide chemistry.***

The number of *degrees of freedom*, F, is the number of independent variables taken from the temperature, pressure and composition of phases, i.e. it is the *number* of variables that must be specified for the system to be completely defined. Again, let us see some examples:

(a) A solid solution in the system Al_2O_3–Cr_2O_3 has one composition variable because the Al_2O_3:Cr_2O_3 ratio can be varied. Thus, if a particular sample has composition x% Al_2O_3, then the Cr_2O_3 content is fixed as $(100 - x)$%. The temperature of these single-phase solid solutions can also be varied. Two degrees of freedom (or three, if pressure is included) are therefore needed in order to characterise a solid solution fully, namely its composition and temperature (and pressure), i.e. $F = 2$ (or 3).

(b) At the other extreme, a partially melted solid in a one-component system, in equilibrium at its melting point, has zero degrees of freedom, i.e. $F = 0$. For instance, a two-phase mixture of ice and water can coexist in equilibrium at only one temperature, 0 °C [ignoring for the moment, the effect of pressure on the melting point of ice]; also, there is no change in the composition of the water; it is fixed at H_2O and therefore $C = 1$.

(c) A good example of a system that is intermediate between the above two cases is boiling water, i.e. water and steam in equilibrium. Again, this does not have a composition variable since both water and steam contain molecules of the same fixed formula, H_2O; it is a *unary* system since $C = 1$. To describe fully a water–steam mixture, that is in equilibrium at the boiling temperature, it is necessary to specify either the steam pressure because then the temperature of boiling is automatically fixed, or vice versa. Therefore, $F = 1$. A practical example of the dependence of boiling point on vapour pressure is that, at sea level, water boils at 100 °C, but at the high altitude of Mexico City, it boils at 92 °C since atmospheric pressure is only 580 mmHg. Application of the phase rule to the boiling water system gives

$$P + F = C + 2$$

$C = 1$ (i.e. H_2O), $P = 2$ (vapour and liquid) and so $F = 1$ (either temperature or pressure, but not both).

The system water–steam is therefore *univariant* because one degree of freedom, either pressure or temperature, is needed to describe completely the system at equilibrium. It should be emphasised that the relative amounts of water and steam are not given by the phase rule. Provided that there is sufficient steam present to maintain the equilibrium pressure, the volume of vapour is unimportant. This illustrates an important general point that the amount of a phase never constitutes a degree of freedom although the composition of a phase, which may be variable, may constitute a degree of freedom.

In many solids with high melting temperatures, the vapour pressure of the solids and even that of the liquid may be negligible in comparison with atmospheric pressure. The vapour phase is effectively non-existent, therefore, and need not be regarded as a possible phase for work at atmospheric pressure. For such systems, and those in which pressure is not a variable, the condensed phase rule, equation (7.2), is used. We use the condensed phase rule for the above cases of melting ice ($P = 2$, $C = 1$, hence $F = 0$ at 0 °C) and Al_2O_3–Cr_2O_3 solid solutions ($P = 1$, $C = 2$, hence F = 2) with temperature and composition as the degrees of freedom for a particular member of the solid solution series.

In using phase diagrams, it is important to understand what is meant by *equilibrium*. The equilibrium state is always that which has the lowest free energy. It may be thought of as lying at the bottom of a free energy well, Fig. 7.2. The problem in determining whether equilibrium has been reached is that other free energy minima may exist but not be as deep as the equilibrium well. There is no simple way of determining whether a particular reaction product, or mixture, represents the equilibrium state. Various tests can be applied (see section 7.4), but if there is a considerable energy barrier involved in moving from a *metastable* state to the *stable* state, this barrier may be prohibitively high and the reaction product may be frozen-in in its metastable state.

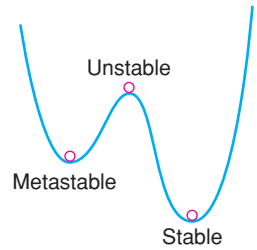


Figure 7.2 ***Schematic diagram showing stable, unstable and metastable conditions.***

In addition to various experimental tests to determine whether equilibrium has been reached, thermodynamic measurements may be used to determine enthalpies and entropies of formation of phases, and using such information, calculated phase diagrams may be obtained and compared with those obtained experimentally. A non-experimental approach which is becoming increasingly useful is to perform lattice energy calculations, which can help to establish the preferred crystal structure for a particular composition. Various modelling codes and strategies are increasingly available and are especially useful as a complement to experimental studies of compound formation, phase equilibria and crystal structures.

A good example of a kinetically stable but thermodynamically metastable state is the metastability of diamond relative to graphite at room temperature. The energy barrier or activation energy for the diamond → graphite reaction is so high that, once formed, diamond is kinetically stable although thermodynamically metastable. There are no known examples of spontaneous transformation of diamond to graphite, although from thermodynamics this should happen; as the James Bond theme song says: 'diamonds are forever'!

This example does, however, bring us to the important point that phase diagrams tell us only the identity of the thermodynamically stable phase or phase mixture at a given temperature. Phase diagrams give no information concerning kinetics of reactions or transformations. For example, the phases $MgSiO_3$ and Mg_2SiO_4 appear on the $MgO–SiO_2$ phase diagram for temperatures at and above room temperature, but this does not mean that MgO and SiO_2 will spontaneously react when they come into contact to form one or other of these magnesium silicate phases. The phase diagram tells us what should happen, eventually, but gives no information at all on reaction rate. One reason why phase diagrams for oxide systems are often limited to the phases and phase mixtures that form at high temperatures is because it can be extremely difficult to know which phases and phase assemblages are genuinely stable thermodynamically at lower temperatures whereas they may be kinetically stable indefinitely.

The thermodynamic meaning of the term *unstable* is also shown in Fig. 7.2. If a ball is perched on a hill-top, the slightest movement is sufficient to cause it to start rolling down one side or the other. In the same way, there is no activation energy involved in changing from a thermodynamically unstable state to either a stable or a metastable state. Examples of unstable equilibrium are difficult to find (because of their instability!) but one can point to their would-be occurrence. For example, inside a region of liquid immiscibility exists an area bounded by a dome called a *spinodal*. Within the spinodal, a homogeneous liquid is unstable and spontaneously separates into two liquids by the process known as *spinodal decomposition*.

We are now ready to derive the phase rule. There is a close analogy between the phase rule and simultaneous equations. To solve simultaneous equations fully, it is necessary to have as many equations as unknowns. If there are insufficient equations to determine values for the unknowns, the shortage is known as the variance.

The unknowns in the phase rule are the total number of variables, which are temperature, pressure and the compositions of all of the phases present. The number of equations relate to the *chemical potentials* of each component; at equilibrium, by definition, the chemical potential of each component is the same in all of

the phases present. The variance is the number of undetermined variables which is the number of degrees of freedom, F, in the phase rule. F is therefore defined as follows:

$$F = \text{total number of variables} - \text{total number of equations relating the variables} \tag{7.3}$$

In order to specify the composition of each phase, for a system of C components, it is necessary to specify $(C - 1)$ variables since the overall composition is unity (if we are using mole fractions) or 100 (using percentages). Hence,

$$\begin{aligned} \text{total number of variables} &= \text{number of phases} \times (C - 1) + \text{pressure} + \text{temperature} \\ &= P(C - 1) + 2 \end{aligned} \tag{7.4}$$

For each component, i, at equilibrium, the chemical potential μ_i is the same for all phases, i.e.

$$\mu_i(\text{phase 1}) = \mu_i(\text{phase 2}), \text{ etc.} \tag{7.5}$$

For P phases, there are $(P - 1)$ equations of this type; therefore,

$$\text{total number of equations} = C(P - 1) \tag{7.6}$$

Combining equations (7.4) and (7.6) gives the phase rule, i.e.

$$\begin{aligned} F &= P(C - 1) + 2 - \mathrm{C}(P - 1) \\ &= 2 - P + C \end{aligned} \tag{7.1}$$

The condensed phase rule, referred to above, equation (7.2), has one less variable if pressure is constant.

7.2 One-Component Systems

The independent variables in a one-component system are temperature, T, and pressure, P, because the composition is fixed. From the phase rule, $P + F = C + 2 = 3$. The system is *bivariant* ($F = 2$) if one phase is present, *univariant* ($F = 1$) if two are present and *invariant* ($F = 0$) if three are present. Schematic phase relations are given in Fig. 7.3 for a one-component system in which the axes are the independent variables, P and T. Possible phases are two crystalline polymorphs (sometimes called *allotropes*), X and Y, liquid and vapour. Each phase occupies an area or *field* on the diagram since $F = 2$ when $P = 1$ (both P and T are needed to describe a point in one of these fields). Each single-phase region is separated from neighbouring single-phase regions by univariant curves (for which $P = 2$ and so $F = 1$). On these curves, if one variable, say P, is fixed, then the other, T, is automatically fixed. The univariant curves on the diagram represent the following equilibria:

(a) BE – transition temperature between polymorphs X and Y; it gives the change of transition temperature with pressure.
(b) FC – change of melting point of polymorph Y with pressure.
(c) AB, BC – sublimation curves for X and Y, respectively.
(d) CD – change of boiling point of polymorph Y with pressure.

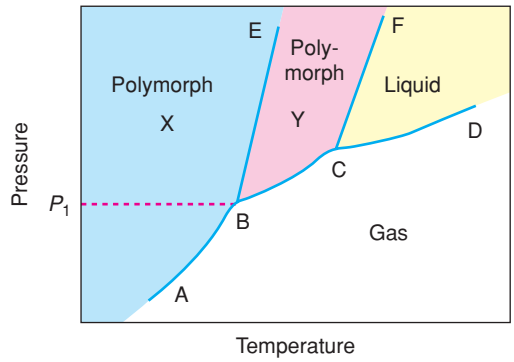


Figure 7.3 ***Schematic* P *versus* T *phase diagram of a one-component system.***

From Fig. 7.3, polymorph X can never melt directly under equilibrium conditions because the fields of X and liquid never meet on the diagram.

On heating, X can either sublime at a pressure below P_1 or transform to polymorph Y at pressures above P_1. It cannot melt directly. Also present in Fig. 7.3 are two invariant points B and C for which $P = 3$ and $F = 0$. The three phases that coexist at point B are: polymorph X, polymorph Y and vapour. Points B and C are also called *triple points*.

7.2.1 The system H_2O

This important one-component system, shown in Fig. 7.4, gives examples of solid–solid and solid–liquid transitions. The polymorph ice I is stable at atmospheric pressure; several high-pressure polymorphs are known – ice II to ice VI. At first sight, there is little similarity between the schematic diagram, Fig. 7.3, and that for water, Fig. 7.4, but this is mainly because of the location of the univariant curve XY that separates the fields of ice I and water. It is well known that ice I has the unusual property of being less dense than liquid water at 0 °C. The effect of pressure on the ice I–water transition temperature can be understood from Le Chatelier's principle: 'When a constraint is applied to a system in equilibrium, the system adjusts itself to

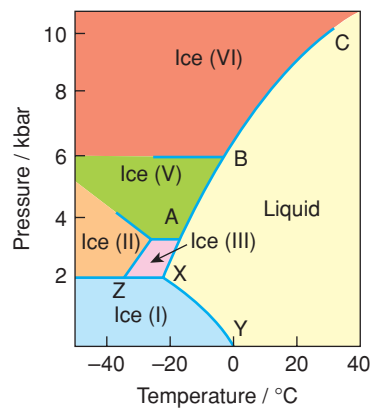


Figure 7.4 ***The system H_2O.***

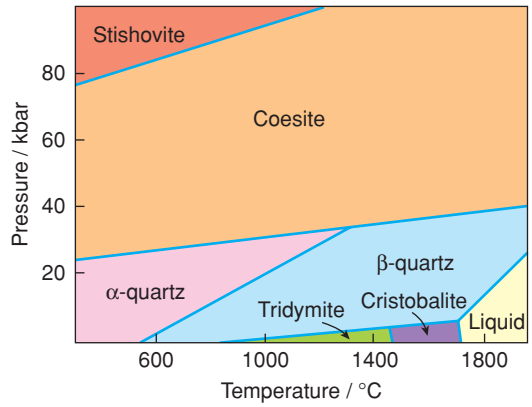


Figure 7.5 ***The system SiO_2. Adapted with permission from Griffin,* Silicate Crystal Chemistry, *© 1992 Oxford University Press.***

nullify the effects of this constraint.' The melting of ice I is accompanied by a decrease in volume; increased pressure makes melting easier, therefore, and so the melting temperature *decreases* with increase in pressure, in the direction YX.

The water system is also more complex than Fig. 7.3 since additional invariant points exist which correspond to three solid phases in equilibrium (e.g. point Z). The rest of the diagram should be self-explanatory; thus the curves YXABC give the variation of melting point with pressure for some of the ice polymorphs. Liquid–vapour equilibria are omitted from Fig. 7.4 because, with the pressure scale used, these equilibria lie very close to the temperature axis and in the high-temperature corner.

7.2.2 The system SiO_2

Silica is the main component of many ceramic materials in addition to being the most common oxide, apart from H_2O, in the Earth's crust. The polymorphism of SiO_2 is complex, with major first-order phase changes such as quartz–tridymite and minor changes such as α(low)–β(high) quartz. The polymorphism at atmospheric pressure can be summarised by the following sequence of changes on heating:

$$\alpha\text{-quartz} \xrightarrow{573\,^\circ\text{C}} \beta\text{-quartz} \xrightarrow{870\,^\circ\text{C}} \beta\text{-tridymite} \xrightarrow{1470\,^\circ\text{C}} \beta\text{-cristobalite} \xrightarrow{1710\,^\circ\text{C}} \text{liquid}$$

With increasing pressure, two main changes are observed, Fig. 7.5; first, the contraction of the field of tridymite and its eventual disappearance, at ~5 kbar; second, the disappearance of the field of cristobalite at ~8 kbar. The disappearance of tridymite and cristobalite under pressure can be correlated with the lesser density of these phases relative to that of quartz, Table 7.1; the effect of pressure generally is to produce polymorphs that have higher density and therefore smaller volume. Above 20–40 kbar (depending on temperature), quartz transforms to coesite, and above 80–100 kbar, stishovite is stable.

It may also be noted that there are many metastable polymorphs of SiO_2 which are absent from Fig. 7.5; for example, it is easy to undercool cristobalite and to observe a reversible β (high) $\rightarrow$ α (low) transformation at ~270 °C. However, at these temperatures, cristobalite is metastable relative to quartz and so this transformation is omitted from Fig. 7.5.

Table 7.1 ***Densities of SiO_2 polymorphs***

Polymorph	Density/g cm^{-3}
Low tridymite	2.265
Low cristobalite	2.334
Low quartz	2.647
Coesite	3.00
Stishovite	4.40

7.2.3 Condensed one-component systems

For many systems and applications of interest in solid state chemistry, the condensed phase rule, equation (7.2), is applicable and pressure is not a variable. If the vapour phase is not important, the phase diagram for a condensed, one-component system then reduces to a line since temperature is the only degree of freedom. For instance, the condensed phase diagram at 1 atm pressure for SiO_2 would simply be a line showing the polymorphic changes that occur with change in temperature. In such cases, it is easier to represent the changes as a 'flow diagram', as indicated above for SiO_2.

7.3 Two-Component Condensed Systems

Two-component systems have three independent variables: P, T and composition. If pressure need not be considered as a variable, the condensed phase rule $P + F = C + 1$ is used. In binary systems ($C = 2$) under these conditions, an invariant point ($F = 0$) occurs when three phases coexist, a univariant curve for two phases and a bivariant condition for one phase. Conventionally, temperature is the vertical axis and composition is the horizontal axis in binary phase diagrams. In most of what follows, we confine discussion to solid–liquid phase diagrams and assume that the vapour pressure of the phases present is unimportant. Referring to such systems as condensed systems often implies both that pressure is not a variable and that the vapour phase is unimportant, although strictly, the condensed phase rule refers only to the first of these conditions.

7.3.1 A simple eutectic system

The simplest two-component condensed system is the eutectic system shown in Fig. 7.6(a). It occurs whenever two non-interacting solids, A and B, that can melt without decomposition, are mixed. No compounds or solid solutions are formed but the mixtures melt at lower temperatures than either pure solid.

In the solid state, therefore, there are no intermediate compounds or solid solutions but only a mixture of the end-member phases, A and B. In the liquid state, at higher temperatures, a complete range of single-phase, liquid solutions occurs. At intermediate temperatures, regions of partial melting appear on the diagram. These regions contain a mixture of a crystalline phase and a liquid of different composition to the crystalline phase.

The phase diagram shows several regions or areas which contain either one or two phases; for consistency throughout this chapter, all single-phase regions are shown shaded. These single-phase areas are separated from two-phase areas, unshaded, by solid curves or lines. The area 'liquid' is single phase and bivariant. Every point within this area represents a different state for the liquid, i.e. a different temperature and composition. In this region, $P = 1$ and $F = 2$.

The other three areas each contain two phases: A + B; A + liquid; B + liquid. They are univariant since $P = 2$ and $F = 1$. Let us consider one of these, the region B + liquid. For a mixture of B and liquid of overall composition f [as marked on the composition axis, Fig. 7.6 (b)] and temperature T_2, this mixture is

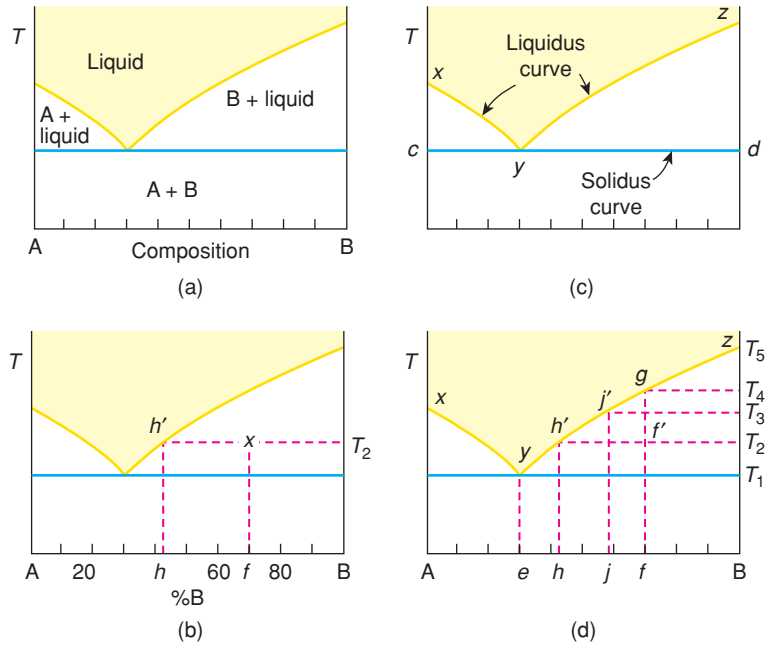


Figure 7.6 ***Simple eutectic binary system.***

represented by point x. In this region, there is only one degree of freedom, either temperature or composition, but not both, and we have decided to fix the temperature at T_2.

In order to determine the compositions of the two phases, B and liquid, in equilibrium at point x, construction lines, dashed in (b), are drawn. First, an *isotherm* is drawn at the temperature of interest, T_2. This is the horizontal dashed line that terminates on the *liquidus* at point h'. This point, h', represents the liquid that is present in the mixture of B + liquid at temperature T_2. The composition of this liquid is given by drawing a vertical line or *isopleth* which intersects the composition (horizontal) axis at h; in this case, the liquid composition is 43% B, 57% A. The other phase that is present in the mixture is B, whose composition is fixed, as pure B, in this example.

An important distinction to be made here is between different meanings of the word 'composition'. It has at least three meanings:

(a) *The composition of a particular phase*. In the above example, the liquid phase has composition h, 43% B, 57% A.
(b) *The relative amounts of the different phases present in a mixture*. This is often referred to as the *phase composition*, although it would be better to refer to it as the phase content. In the above example, B and liquid are present in the ratio ~1:1 (see later for an explanation of the lever rule used to determine phase compositions or contents).
(c) *The overall composition of a mixture*, in terms of the components and irrespective of the phases present. In the above example, the overall composition of mixture f is 30% A, 70% B.

Since there is no universally adopted convention over the use of the word 'composition', one can only say, be careful!

In the sense to which the phase rule applies, composition is a degree of freedom only when it refers to the actual compositions of the phases involved, category (a). The amounts of phases in a mixture, category (b), do not constitute a degree of freedom. Thus, along the isotherm passing through h' and x at temperature T_2, the relative amounts of B and liquid vary, but their compositions do not. The overall composition, category (c), is included in the phase rule, not as a degree of freedom but as the number of components.

7.3.1.1 Liquidus and solidus

The *liquidus* curve, *xyz* in (c), gives the highest temperatures at which crystals can exist. Liquids that cross the liquidus curve between points *x* and *y* on cooling enter the two-phase region: A + liquid. For these compositions, A is the *primary phase*, because it is the first phase to crystallise on cooling. Similarly, for liquids between *y* and *z*, B is the primary phase. The line *cyd* is the *solidus* and gives the lowest temperatures at which liquids can exist, i.e. above the solidus melting commences. In summary, above the liquidus, mixtures are completely molten, below the solidus they are completely solid, and in between partial melting occurs.

7.3.1.2 Eutectic

Point *y* is an *invariant point* at which three phases, A, B and liquid, coexist, i.e. $C = 2$, $P = 3$, $F = 0$. It is a *eutectic* and its temperature is the lowest temperature at which a composition (here 70% A, 30% B) can be completely liquid. It is also a minimum on the liquidus curve, *xyz*. In simple eutectic systems such as this, the eutectic and solidus temperatures are the same.

7.3.1.3 Lever rule

In order to determine the relative amounts of two phases in a mixture [the phase composition, category (b) above], the *lever rule* is used. This is the same as the 'principle of moments' which operates when children balance on a see-saw. The pivot point or fulcrum of the see-saw is equivalent to the overall composition of a mixture and the two phases are equivalent to the two children. In the same way that children of different sizes sit at different distances from the centre of the see-saw to achieve balance, so the compositions and amounts of two phases in a mixture are interrelated. For the see-saw, the balance condition, Fig. 7.7(a), is given by

$$m_1 \times (\text{distance } xy) = m_2 \times (\text{distance } yz) \tag{7.7}$$

This may be reorganised to give the ratio m_1/m_2 or the ratio $m_1/(m_1 + m_2)$.

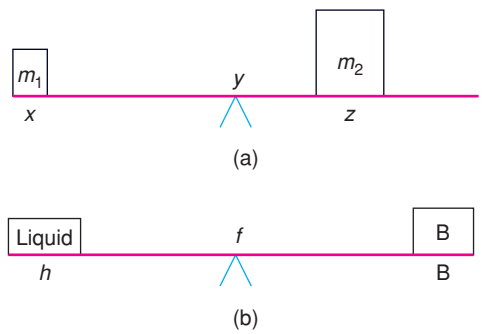


Figure 7.7 Principle of moments.

Returning now to the lever rule and mixtures of phases, the liquid in Fig. 7.6(b) that is present in composition f at temperature T_2 has composition h on the A–B axis; the crystals are pure B and the overall composition is f. Crystals B and liquid h' are therefore in equilibrium. The line joining these two compositions at T_2 is known as a *tie line*: as a general rule, a tie line connects two phases that are in thermodynamic equilibrium. This equilibrium is reproduced schematically in Fig. 7.7(b). Applying the principle of moments,

$$(\text{amount of liquid}) \times (\text{distance } hf) = (\text{amount of B}) \times (\text{distance } f\text{B}) \tag{7.8}$$

Since

$$(\text{amount of B}) + (\text{amount of liquid}) = 1 \tag{7.9}$$

equation (7.8) may be rearranged to give

$$\frac{(\text{amount of B})}{(\text{amount of liquid})} = \frac{(\text{amount of B})}{1 - (\text{amount of B})} = \frac{hf}{\text{B}f} = \frac{hf}{\text{B}h - fh}$$

$$(\text{amount of B}) = \frac{hf}{\text{B}h} \tag{7.10}$$

$$(\text{amount of liquid}) = \frac{\text{B}f}{\text{B}h} \tag{7.11}$$

The lever rule may be used to determine how the compositions of phase mixtures change with temperature, Fig. 7.6(d). At temperature T_2, the amount of liquid in composition f is given by Bf/Bh, i.e. 0.53. At a higher temperature T_3, the amount of liquid is Bf/Bj, i.e. 0.71. At a lower temperature, just above the solidus T_1, the amount of liquid is Bf/Be, i.e. 0.43. Clearly, therefore, the effect of raising the temperature above T_1 is to increase the degree of melting from 43% at T_1 to 71% at T_3. The limit is reached at T_4 where the fraction of liquid is 1 and melting is complete.

As melting increases with temperature, the composition of the liquid must also change: since crystals of B disappear into the liquid on melting, the liquid must become richer in B. Thus the first liquid that appears on heating, at temperature T_1, has composition e, i.e. 30% B, 70% A. As melting continues, the liquid follows the liquidus curve $yh'j'g$ until, when melting is complete at T_4, the liquid has composition f, i.e. 70% B, 30% A. On cooling liquid of composition f, the reverse process should be observed under equilibrium conditions. At T_4, crystals of B begin to form and with decrease in temperature the liquid composition moves from g to y as more crystals of B precipitate.

7.3.1.4 Eutectic reaction

The *eutectic reaction* which occurs on cooling through temperature T_1 gives a good example of the use of the lever rule. Just above T_1, the fraction of B present in composition f is fe/Be and is roughly 0.57. Just below T_1 the fraction of B is fA/BA and is roughly 0.70. Thus, the residual liquid, of composition e, has crystallised to a mixture of A and B, i.e. the quantity of B present has increased from 0.57 to 0.70 and crystals of A form for the first time. A solid mixture of A and B of eutectic composition e melts completely at temperature T_1 and conversely, on cooling, a homogeneous liquid of this composition completely crystallises to a mixture of A and B at T_1.

The reactions described above are those that should occur under equilibrium conditions. This usually means that slow rates of heating and, especially, cooling are necessary. Rapid cooling often leads to different results,

especially in systems with more complicated phase diagrams. However, the equilibrium diagram can often be very useful in rationalising these non-equilibrium results (see later).

7.3.1.5 The liquidus, saturation solubilities and freezing point depression

The liquidus curve *xyz* may be regarded in various ways. In addition to giving the maximum temperature at which crystals can exist, it is also a *saturation solubility curve*. Thus, curve *yz* gives the solubility limit with temperature for crystals of B dissolved in liquid. Above *yz* a homogeneous solution occurs but below this curve undissolved crystals of B are present. On cooling, crystals of B precipitate below the temperatures of curve *yz*; otherwise a metastable undercooled, supersaturated solution would be present.

Another interpretation of the liquidus is that it shows the effect of soluble impurities on the melting points of pure compounds. If a specimen of B is held at temperature T_4 and a small amount of A is added, then some liquid of composition *g* will form. As the amount of added A increases, so the amount of liquid *g* increases until, when sufficient *flux*, A, has been added to bring the overall composition to *g*, the solid phase disappears and the sample is all liquid. Therefore, a small amount of soluble impurity A has lowered the melting point of B from T_5 to T_4. A familiar practical example of this is the addition of salt to icy roads. In the binary system H_2O–NaCl, addition of NaCl lowers the melting point of ice below 0 °C; the system contains a low-temperature eutectic at –21 °C.

7.3.2 Binary systems with compounds

Three types of binary system with a compound AB are shown in Fig. 7.8 and Fig. 7.9. A stoichiometric binary compound such as AB is represented by a vertical line. This shows the range of temperatures over which it is stable.

7.3.2.1 Congruent melting

Compound AB melts *congruently* in Fig. 7.8(a) because it changes directly from solid AB to liquid of the same composition at temperature T_3. This figure may be conveniently divided into two parts, given by the composition ranges A–AB and AB–B as shown in (b); each part may be treated as a simple eutectic system in exactly the same manner as Fig. 7.6. Although the horizontal lines at T_1 and T_2, corresponding to the two eutectic temperatures, meet the vertical line representing crystalline AB, no changes would be observed at T_1 and T_2 on heating pure AB. This is because these horizontal lines should peter out as composition AB is approached; composition AB, by itself, is an independent one-component system and only when another component, A or B, is added are changes observed at T_1 or T_2.

A characteristic feature of congruent melting is that it represents a maximum in melting or liquidus curves; thus, either side of point *y* in (a), the liquidus temperature drops.

7.3.2.2 Incongruent melting, peritectic point, peritectic reaction

In Fig. 7.8(c), compound AB melts *incongruently* at T_2 to give a mixture of A and liquid of composition *x*. The relative amounts of liquid and A just above T_2 are given by the lever rule; the fraction of liquid = *yz*/*xz*. On further heating, A gradually dissolves as the liquid becomes richer in A and moves along the liquidus in the direction *xn*. At T_3, the liquid has reached *n* and the last A crystals should disappear.

Point *x* is an invariant point at which three phases coexist: A, AB and liquid; in fact, these three phases coexist anywhere along the line *xyz* of constant temperature T_2 but the invariant point is usually taken as point *x* on the liquidus surface. This invariant point is known as a *peritectic* point; it can be distinguished

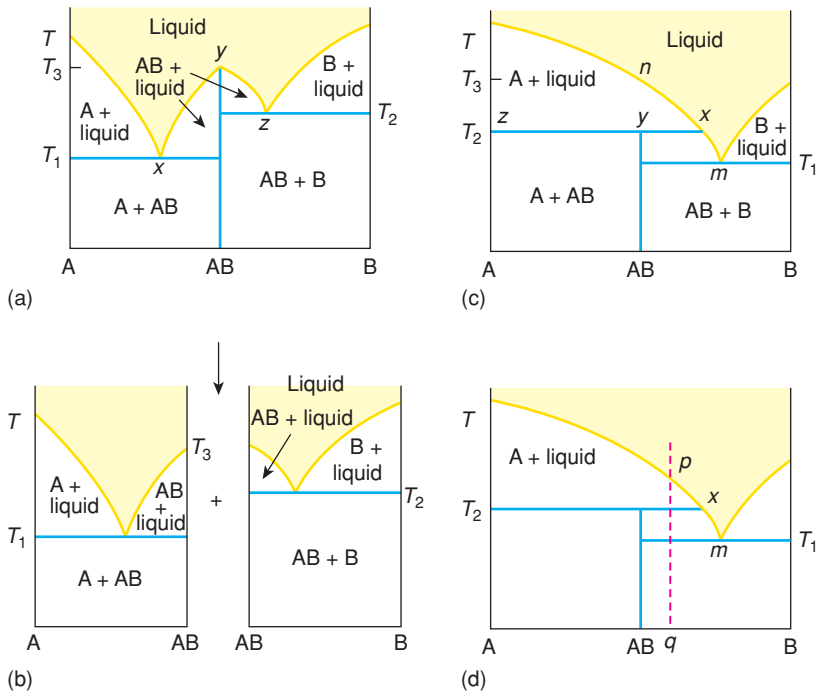


Figure 7.8 ***Binary systems showing a compound AB melting congruently (a) and incongruently (c), (d). In (b), the diagram in (a) is separated into two self-contained, simple eutectic systems.***

from a eutectic point in two ways. First, unlike a eutectic, it is not a minimum on the liquidus curve. Second, the composition of the liquid cannot be represented by positive quantities of the two coexisting solids, i.e. composition x does not lie between A and AB, as is the case for a eutectic. In addition to the peritectic point x, the phase diagram also has a eutectic at point m. This is a minimum on the liquidus surface and the composition of point m lies between the compositions AB and B of the two solid phases that coexist with liquid at the eutectic temperature.

Phase AB has a primary phase field. It is the first phase to crystallise from liquids in the range x–m. However, the composition of AB is separated from its primary phase field. This is different to the case for congruently melting AB, Fig. 7.8(a), where the composition AB lies *within* the range xyz over which AB is the primary phase.

Let us now describe the behaviour of different liquid compositions on cooling, beginning with liquid of composition q in Fig. 7.8(d). At point p, solid A starts to crystallise; more A crystals form as the temperature drops and the liquid composition moves from p to x. At temperature T_2, the complex *peritectic reaction*: liquid (x) + A $\rightarrow$ liquid (x) + AB occurs. Thus, the crystalline phase changes from A to AB and the amount of liquid present must diminish. From the lever rule, just above T_2, the mixture is ~85% liquid and just below T_2 only ~50% liquid. Therefore, during the peritectic reaction, *all* of phase A has reacted with *some* of the liquid to give AB. On further cooling from T_2 to T_1, more AB crystallises as the liquid composition moves from x to m; finally, at T_1, the residual liquid of composition m crystallises to a mixture of AB and B. Just above T_1, the mixture is ~40% liquid and 60% AB; just below T_1, it is ~20% B and 80% AB. The behaviour that has just been described represents crystallisation under conditions of thermodynamic equilibrium in which at each temperature during cooling the system adjusts to give the phases expected from the phase diagram. As

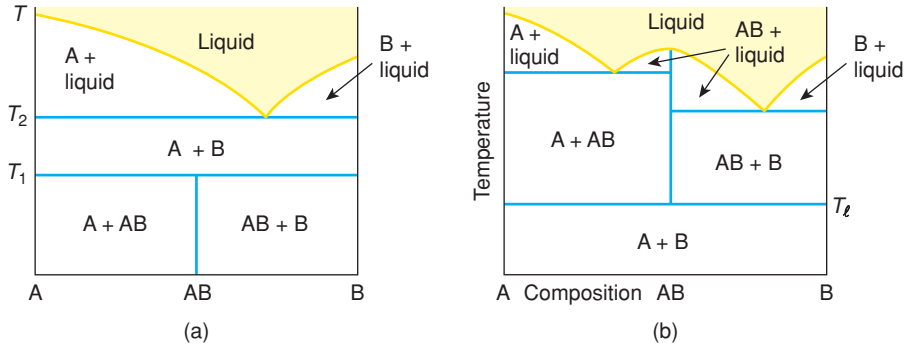


Figure 7.9 ***Binary system showing compound AB with (a) an upper limit of stability and (b) a lower limit of stability.***

discussed in the next section, however, attainment of equilibrium can be difficult, especially when peritectic reactions are involved.

The behaviour on cooling any liquid of composition between A and AB under equilibrium conditions is similar but with one important difference. Just above T_2, a mixture of A and liquid is present; just below T_2, a solid mixture of A and AB is present. At T_2, therefore, the peritectic reaction for these compositions involves *some* of A reacting with *all* the liquid to give AB. Below T_2, a mixture of A and AB coexists and no further changes occur on cooling.

The behaviour of liquid of composition *n*, Fig. 7.8(c), on cooling is a special case. Between temperatures T_3 and T_2, solid A begins to crystallise. At T_2 the peritectic reaction occurs but in this case *all* of A reacts with *all* of liquid to give single phase AB. This peritectic reaction, which seems simple in theory, is complex in practice and in general it will be difficult to obtain a phase-pure, AB product unless samples are cooled very slowly through T_2 and held for a long time at temperatures just below T_2 to allow enough time for complete reaction to occur.

7.3.2.3 Non-equilibrium effects

In systems that contain incongruently melting compounds, such as in Fig. 7.8(c), it is easy to obtain non-equilibrium products on cooling. This is because the peritectic reaction that *should* occur between A and liquid on cooling is slow, especially if the crystals of A are much denser than the liquid and have settled to the bottom of the sample container. What commonly happens in practice is that the crystals of A which form are effectively lost to the system and there is not time for much peritectic reaction to occur at T_2. The liquid of composition *x* then effectively begins to crystallise again below T_2, but this time crystals of AB form; at the eutectic temperature T_1, the residual liquid *m* crystallises to a mixture of AB and B, as usual. Thus, it is fairly common to obtain a mixture of *three* crystalline phases on cooling: A, AB and B, at least one of which should be absent under equilibrium conditions.

Another common type of non-equilibrium assemblage occurs when the intermediate, incongruently melting compound AB fails completely to crystallise on cooling. In this case, the liquidus surface changes from showing two invariant points, the peritectic *x* and the eutectic *m*, to a single invariant point with a eutectic similar to, but probably not identical with, *m* and temperature T_1. If this happened on cooling liquids, e.g. *n* in Fig. 7.8(c) or *p* in (d), the products would be A and B with no AB. Hence the peritectic reaction A + liquid $\rightarrow$ AB would have been suppressed entirely.

7.3.2.4 *Upper and lower limits of stability*

Sometimes, compounds decompose before melting, as shown for AB in Fig. 7.9(a). In this case, compound AB has an *upper limit of stability* and at temperature T_1 it disproportionates into a mixture of A and B; at higher temperatures the system is simple eutectic in character.

There are also many examples of systems which contain compounds with a *lower limit of stability*, i.e. below a certain temperature, AB decomposes into a mixture of A and B, Fig. 7.9(b). The behaviour of AB at higher temperatures can then be any of the three types described above.

7.3.3 Binary systems with solid solutions

7.3.3.1 *Complete solid solution*

The simplest solid solution system is one that shows complete miscibility in both solid and liquid states, Fig. 7.10. The melting point of one end member, A, is depressed by addition of the other end member, B, and vice versa, that of B is increased on addition of A. The liquidus and solidus are smooth curves which meet only at the end-member compositions A and B. At low temperatures, a single-phase solid solution exists and is bivariant ($C = 2$, $P = 1$, and so $F = 2$). At high temperatures, a single-phase liquid solution exists and is similarly bivariant. At intermediate temperatures, a two-phase region of solid solution + liquid exists ($C = 2$, $P = 2$, therefore $F = 1$). Both the solidus and liquidus curves are smooth curves whose temperature changes with composition and which meet at the two end members A and B. Within this two-phase region, the compositions of the two phases in equilibrium are found by drawing isotherms or tie-lines at the temperature of interest, e.g. T_1. The intersection of the tie-line and the solidus gives the composition of the solid solution, a, and the intersection of the tie-line and the liquidus gives the liquid composition, b. The compositions of the solid solution and liquid that are in equilibrium at T_1 are found from the intersection of the isopleths at a and b on the A–B composition axis, as shown.

On cooling liquids in a system such as this, the crystallisation pathways are complex, Fig. 7.10(b). A liquid of composition b begins to crystallise a solid solution of composition a at temperature T_1. At a lower temperature, T_2, the amount of solid solution present increases but also its composition changes to a'. The fraction of solid solution a', given by the lever rule, is equal to $bb'/a'b'$, i.e. the equilibrium mixture is

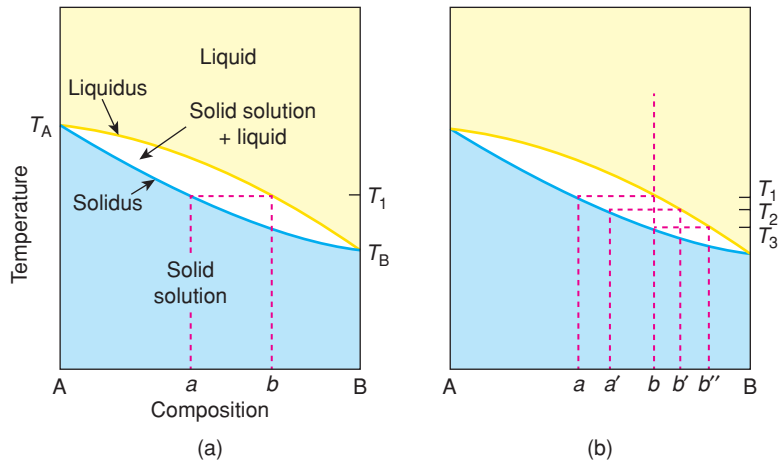


Figure 7.10 ***Binary system with a complete range of solid solutions.***

approximately one-third solid solution and two-thirds liquid at T_2. Crystallisation with decreasing temperature under equilibrium conditions requires the composition of the solid solution to change continuously, which is a difficult process once crystallisation has already occurred. With falling temperature, under equilibrium conditions, both crystals and liquid become progressively richer in B but their quantities change in accord with the lever rule; the overall composition must obviously always be b. Finally, at T_3, the solid solution composition reaches the bulk composition b and the remaining liquid, of composition b'', disappears.

7.3.3.2 Fractional crystallisation

In systems such as this, metastable or non-equilibrium products are often produced by a process of *fractional crystallisation*. This occurs unless cooling takes place sufficiently slowly that equilibrium is reached at each temperature. The crystals that form first on cooling liquid b have composition a. If these do not have time to re-equilibrate with the liquid on further cooling, they are effectively lost from the system. Each new crystal that precipitates will be slightly richer in B and the result is that crystals form which have composition ranging from a to somewhere between b and B. In practice, the crystals that precipitate during cooling are often '*cored*'. The central part that forms first may have composition a and on moving out radially from the centre, the crystal becomes increasingly rich in B.

Coring occurs in many rocks and metals. The plagioclase feldspars, solid solutions of anorthite, $CaAl_2Si_2O_8$, and albite, $NaAlSi_3O_8$, have the phase diagram shown in Fig. 7.11. Igneous rocks contain plagioclase feldspars that formed by slow cooling of liquids. Such feldspar liquids and crystals are notoriously slow to equilibrate and although the cooling of melts in Nature may have been very slow, it is common nevertheless to find rocks in which fractional crystallisation has occurred. In these, the plagioclase crystals have calcium-rich centres and sodium-rich outer regions.

Coring may occur in metals during the manufacture of bars and ingots. The molten metal is poured into moulds (or 'sand cast' in moulds of sand) and allowed to cool. If the metal composition is part of a solid

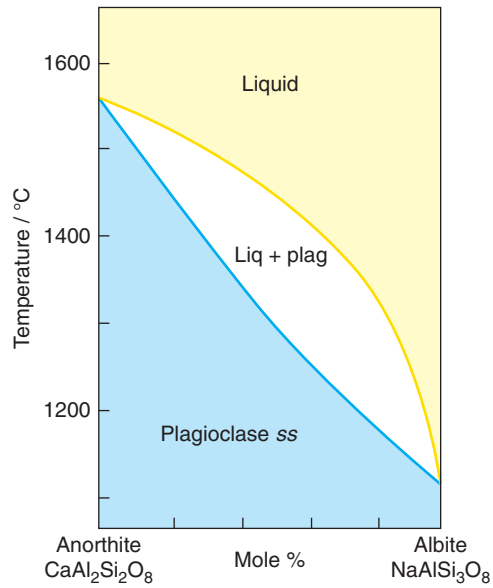


Figure 7.11 ***The plagioclase feldspar system, anorthite–albite.***

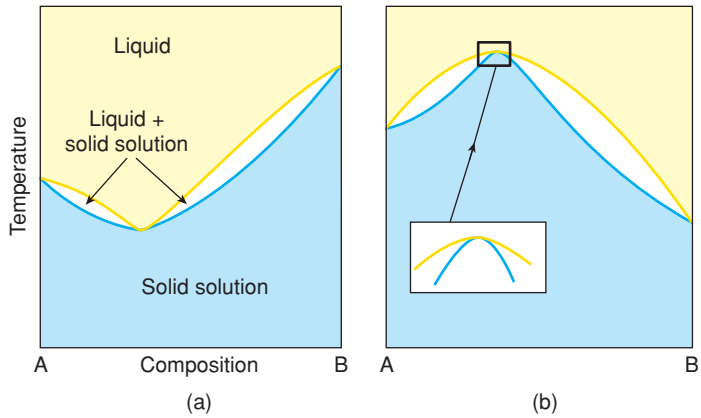


Figure 7.12 ***Binary solid solution systems with (a) thermal minima and (b) thermal maxima in liquidus and solidus curves.***

solution then coring may occur as the melt freezes. Coring is usually deleterious to the properties of the metal and has to be eliminated. This can be done by subsequently heating the bars to just below the solidus temperature at which homogenisation of the metal by solid state diffusion occurs.

7.3.3.3 Thermal maxima and minima

The simplest solid solution phase diagram is shown in Fig. 7.10; both liquidus and solidus curves decrease smoothly on changing from composition A to composition B. It is possible, however, for solid solution phase diagrams to exhibit either a thermal minimum or a thermal maximum in the liquidus and solidus curves, Fig. 7.12. These maxima and minima are called *indifferent points* because they are not true invariant points. For an invariant point, three phases exist in equilibrium ($F = 0$, $P = C + 1 = 3$), but this condition never arises in solid solution systems with indifferent points because there are never more than two phases present, i.e. solid solution and liquid solution. The liquidus and solidus are continuous through the thermal maximum or minimum [as shown in the inset, Fig. 7.12(b)] and do not show a discontinuity (such as shown by peritectics and eutectics).

7.3.3.4 Partial solid solution systems

Complete solid solubility, Figs. 7.10–7.12, occurs only when the cations or anions that replace each other are similar in size and the end-member phases have the same crystal structure. It is far more common for crystalline phases to be only partially soluble in each other. The simplest case, Fig. 7.13, is a straightforward extension of the simple eutectic system, inset (i). Crystals of B dissolve in crystals of A to form a solid solution (shaded) the maximum extent of which depends on temperature and is given by the curves *xmp* (b). The extent of solid solution is a maximum at the eutectic temperature, point *m*. Similarly, B is able to partially dissolve A (shaded region, B ss). The extent of the B ss region also varies with temperature, as given by the curve *ynq*; again, the maximum extent is at the eutectic temperature, point *n*.

In the two-phase region (A ss + B ss), the composition of the A solid solution is given by the intersection of the tie-line (dashed) at the temperature of interest, T_1, and the curve *mx*; it corresponds to point *a* on the

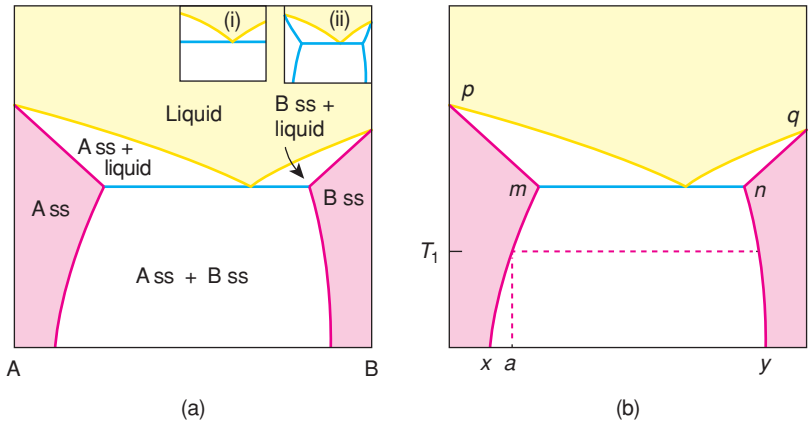


Figure 7.13 ***Simple eutectic system showing partial solid solubility of the end members.***

composition axis. The composition of the B solid solution in this two-phase region is given similarly by the intersection of the tie-line and curve *ny*.

In many phase diagrams, the solids are *line phases*, i.e. they are stoichiometric phases of fixed composition [as shown for A and B end members in inset (i)]. In practice, compositions may vary somewhat, as in inset (ii). An example of a real system similar to Fig. 7.13 is forsterite, Mg_2SiO_4–willemite, Zn_2SiO_4, Fig. 7.14; substitutional solid solutions in this system are discussed in Chapter 2.

Another type of simple binary system with partial solid solubility is shown in Fig. 7.15. This rather strange-looking diagram can be derived from a simple system showing complete solubility [inset (i)]. First, suppose that an *immiscibility dome* exists within the solid solutions which has an *upper consolute temperature* as shown in inset (ii). Above the upper consolute temperature, a single-phase solid solution exists, but below

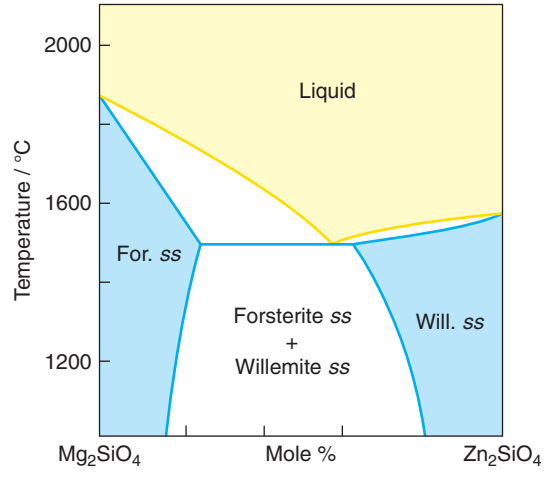


Figure 7.14 ***The system Mg_2SiO_4–Zn_2SiO_4. Reproduced with permission from E.R. Segnit and A.E. Holland, J. Amer. Ceram. Soc., 48, 412, © 1965 John Wiley & Sons.***

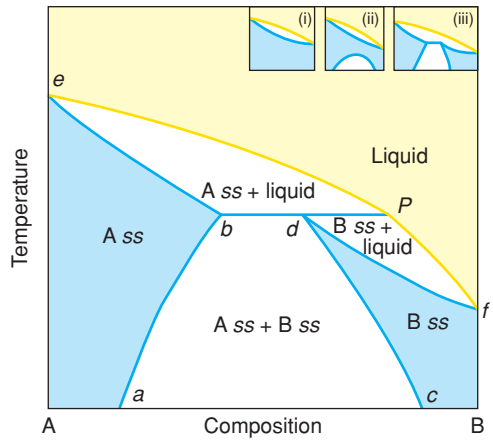


Figure 7.15 ***Binary system with partial solid solution formation.***

it a mixture of two phases exists. Second, let the dome expand to higher temperatures until it intersects the melting curves. The result is shown in inset (iii) and on an expanded scale as the main diagram in Fig. 7.15.

Let us now describe this phase diagram. A limited range of A solid solutions exists and their compositional extent increases with increase in temperature as shown by the solid solubility curve *ab*. Similarly, B forms a range of limited solid solutions whose compositional extent increases with increase in temperature as shown by solubility curve *cd*. Melting of the A solid solutions is straightforward since they enter a region of A ss + liquid and with increase in temperature the A ss becomes increasingly rich in A as shown by the solidus curve *be*. Most of the B solid solutions also melt in a similar way by entering the two-phase region B ss + liquid as given by the solidus curve *fd*. With increase in temperature the B solid solution in equilibrium with liquid becomes increasingly rich in A. The behaviour of the B solid solution of limiting composition at point *d* is completely different, however. This composition melts incongruently to give A ss + liquid with an associated peritectic point *p*.

A more complex phase diagram containing an incongruently melting phase that forms a range of limited solid solutions is shown in Fig. 7.16. This can be understood from the progressive introduction of solid solutions shown in insets (i)–(iv), again working on the principle that the maximum extent of solid solution occurs at the temperatures T_1 and T_2 which are the peritectic and eutectic temperatures. The AB ss area shown in Fig. 7.16 is, in principle, the same as that of the B ss shown in Fig. 7.15, although the actual shapes are completely different. In the AB area, the A-rich limit is shown as a vertical line and is independent of temperature, whereas the B-rich limit of the solid solution is temperature dependent. The A-rich limit melts incongruently whereas the rest of the solid solutions enter the two-phase region AB ss + liquid.

7.3.4 Binary systems with solid–solid phase transitions

Phase transitions refer to a change in crystal structure from one polymorph to another, usually as a function of temperature but also pressure. An example is tetragonal to cubic $BaTiO_3$, which occurs at the ferroelectric Curie temperature, 127 °C. Several other examples are found in the polymorphism of SiO_2, Fig. 7.5. Phase transitions that thermodynamically are first order and involve a change in some property such as volume or enthalpy, can be treated in much the same way as melting phenomena. In a one-component condensed system, e.g. pure A, two solid polymorphs may coexist in equilibrium at one fixed point ($C = 1$, $P = 2$, therefore $F = 0$). In binary systems that do not contain solid solutions, phase transitions in either end member

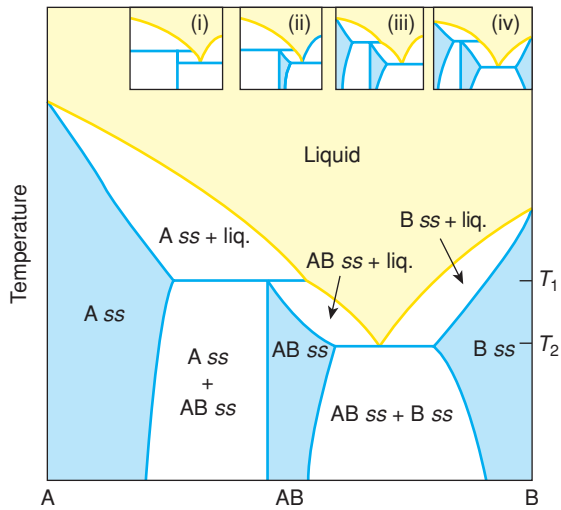


Figure 7.16 ***Binary system with incongruently melting compound and partial solid solution formation.***

are represented by horizontal (i.e. isothermal) lines. In Fig. 7.17, the low-temperature polymorphs of A and B are labelled αA and αB, respectively. Transition temperatures are $\alpha B \rightarrow \beta B$ at T_1, $\alpha A \rightarrow \beta A$ at T_2 and $\beta A \rightarrow \delta A$ at T_4 and are the same for the pure phases as for phase mixtures. This means that any mixture of phases A and B will show the various transitions in the A and B phases independently of the presence of the second phase.

7.3.5 Binary systems with phase transitions and solid solutions: eutectoids and peritectoids

In systems which have both solid solutions and phase transitions, three types of phase diagram are possible, Fig. 7.18. For the end members A and B, the transitions occur at a fixed temperature ($C = 1$, $P = 2$, $F = 0$). However, for the intermediate compositions, two phases can coexist over a range of temperatures or

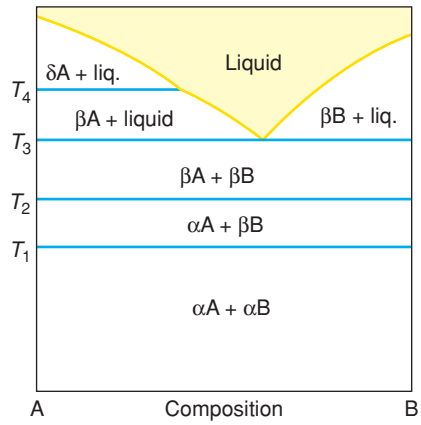


Figure 7.17 ***Simple eutectic system with solid–solid phase transitions.***

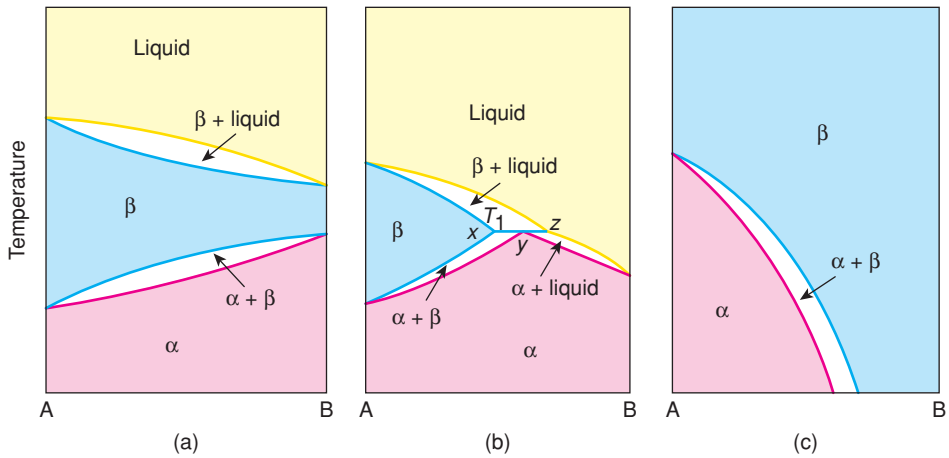


Figure 7.18 ***Binary solid solution systems with polymorphic phase transitions.***

compositions because there is now one degree of freedom ($C = 2$, $P = 2$, $F = 1$). Thus, regions containing two solid phases, e.g. ($\alpha + \beta$), are generally observed, as shown in Fig. 7.18(a). The treatment of the $\alpha \rightarrow \beta$ change in (a) is exactly the same as for the melting of β solid solutions, as discussed for Fig. 7.10. The melting relations in Fig. 7.18(a) are the same as in Fig. 7.10 but could be any of the three types given in Fig. 7.10 and Fig. 7.12. In Fig. 7.18(a), A, B and the entire range of solid solutions show both α and β polymorphs.

In Fig. 7.18(b), the $\alpha \rightarrow \beta$ transition intersects the solidus curve because the $\alpha \rightarrow \beta$ transition in B and in solid solutions rich in B would occur, hypothetically, above the melting point of these compositions. The nature of the intersection of the three one-phase fields (α, β and liquid) and the three two-phase regions at T_1 is typical of solid solution systems. Always, pairs of adjacent one-phase areas such as α and β must be separated by a two-phase area ($\alpha + \beta$), although in practice the width of the two-phase regions may be difficult to detect experimentally. Three phases coexist at one temperature, T_1, and on the horizontal line *xyz*. This line therefore represents an invariant condition and point *z* on the liquidus surface is a peritectic because α composition at point *y* melts incongruently to β solid solution of composition *x* and liquid of composition *z*.

In Fig. 7.18(c), the temperature of the $\alpha \rightarrow \beta$ transition decreases increasingly rapidly as the solid solutions become richer in B. For pure B, the α polymorph does not exist at any real temperature.

A schematic binary system that has phase transitions and partial solid solubility is shown in Fig. 7.19. Two-phase regions of (αA + βA) and (αB + βB) solid solution again exist. We have already seen analogies between melting and phase transition behaviour in solid solution systems. A further analogy exists between the eutectic, *e*, and the *eutectoid*, *b*, in Fig. 7.19. The line *abc* represents an invariant condition over which three phases coexist, βA ss (composition *a*), βB ss (composition *b*) and αB ss (composition *c*). The *eutectoid reaction* at *b* on cooling is 1 solid (βB ss) $\rightarrow$ 2 solids (βA ss + αB ss). Thus, both the eutectic and eutectoid reactions are disproportionation reactions: on cooling, a liquid at point *e* crystallises to a mixture of βA ss and βB ss, whereas at point *b*, a βB ss disproportionates into a solid mixture of βA ss and αB ss.

We can also have a solid state analogue of a peritectic known as a *peritectoid*. Point *y* is a peritectoid. On heating, αA ss at point *y* decomposes to give a mixture of two solids, βA ss and αB ss. The line *xyz,* therefore, represents an invariant condition over which three phases coexist, αA ss (composition *y*), βA ss (composition *x*) and αB ss (composition *z*). On heating, decomposition of αA ss is analogous to the melting of an incongruent compound at a peritectic temperature. Peritectic melting involves the reaction 1 solid $\rightarrow$ 1 solid + 1 liquid whereas a peritectoid transformation involves the reaction, 1 solid $\rightarrow$ 2 solids.

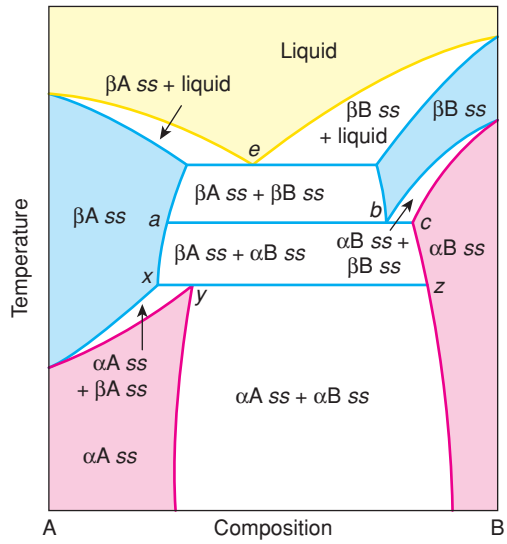


Figure 7.19 ***Binary eutectic system with polymorphic transitions and partial solid solution formation.***

7.3.6 Binary systems with liquid immiscibility: $MgO–SiO_2$

All the phase diagrams considered so far show a single liquid phase over the entire range of compositions A–B. Liquid immiscibility is fairly common, however, as shown by the familiar example of hydrocarbon oils and water, which tend not to mix. A simple schematic liquid immiscibility system is shown in Fig. 7.20(a). It is characterised by an *immiscibility dome*, inside which a mixture of two liquids occurs but outside which, the liquids are miscible. The dome maximum is the *upper consolute temperature*.

At lower temperatures, the liquid(s) crystallise and a return to more familiar solid–liquid equilibria occurs. In Fig. 7.20(b), a combination of liquid immiscibility at B-rich compositions and simple eutectic behaviour is shown. The liquidus surface in the primary phase field of B is interrupted by the presence of the immiscibility dome. At this intersection, we encounter a new kind of invariant condition since along the line *xyz*, two liquids (*x*, *y*) and one solid (B at *z*) exist, i.e. $C = 2$, $P = 3$, $F = 0$. Such a condition is known as a *monotectic*.

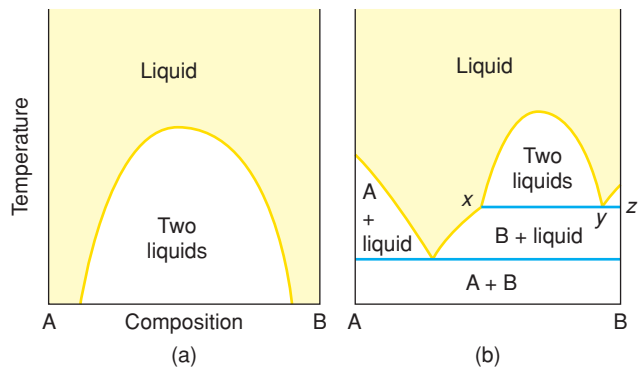


Figure 7.20 ***Liquid immiscibility domes in phase diagrams.***

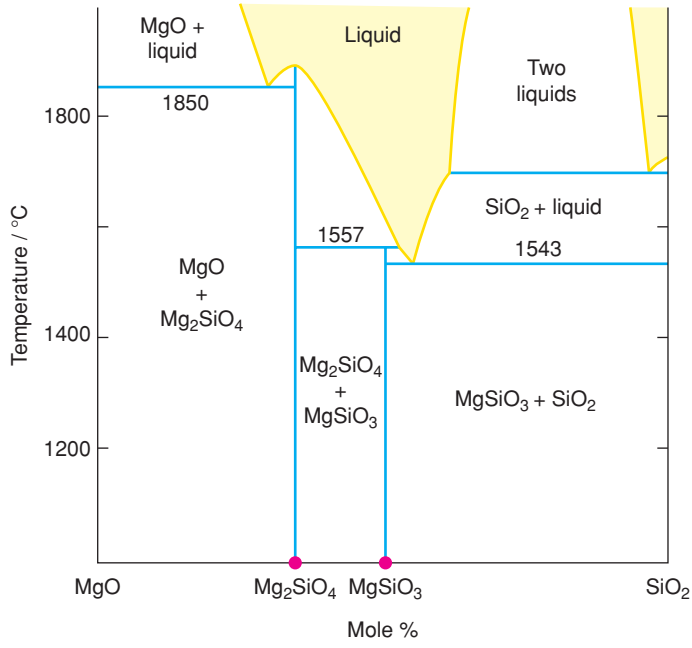


Figure 7.21 ***$MgO–SiO_2$ phase diagram.***

Liquid immiscibility is common in silicate systems at high temperatures in, e.g., $MO–SiO_2$: M = Mg, Ca, Zn, Co. It arises because the very different size, coordination and bonding requirements of, e.g., Ca^{2+} and Si^{4+} make it difficult to form homogeneous liquids over a considerable range of SiO_2-rich compositions. The phase diagram for $MgO–SiO_2$ is shown in Fig. 7.21. The liquid immiscibility region occurs above 1700 °C for SiO_2-rich compositions. The remainder of the diagram shows one congruently melting phase, Mg_2SiO_4, and one incongruently melting phase, $MgSiO_3$. $MgSiO_3$ melts incongruently to Mg_2SiO_4 and liquid at 1557 °C. The system also contains two eutectics at 1850 and 1543 °C together with a monotectic at 1700 °C.

7.3.7 Some technologically important phase diagrams

7.3.7.1 The system Fe–C: iron and steel making

A eutectoid reaction in the system Fe–C is of great importance in steel making. The Fe-rich part of this phase diagram is shown in Fig. 7.22. In addition to the eutectoid at 723 °C representing decomposition of γFe ss on cooling, the diagram shows peritectic melting of the γ solid solutions at ~1500 °C and a eutectic between γFe ss and Fe_3C at 1147 °C and 4.3 wt% C. The changes that occur during steel manufacture focus on thermal cycling of Fe–C alloys and in particular changes associated with the eutectoid decomposition of γFe ss.

Fe exists as three polymorphs: *bcc* α, stable below 910 °C; *fcc* γ, stable between 910 and 1400 °C; and *bcc* (again) δ, stable between 1400 °C and the melting point 1534 °C. γ-Fe can dissolve appreciable amounts of carbon, up to 2.06 wt%, whereas the α and δ forms dissolve very much less, up to a maximum of 0.02 and 0.1 wt%, respectively (see Fig. 2.12 and associated discussion).

Most carbon steels contain less than 1 wt% C, viz. 0.2–0.3 wt% for use as structural members. On cooling from the melt, and in the temperature range 800–1400 °C, a solid solution of C in γ-Fe called *austenite* forms. However, the austenite solid solutions are unstable on further cooling (below 723 °C) because, when the structure changes from that of γ-Fe to α-Fe, exsolution or precipitation of the carbide phase, Fe_3C, must

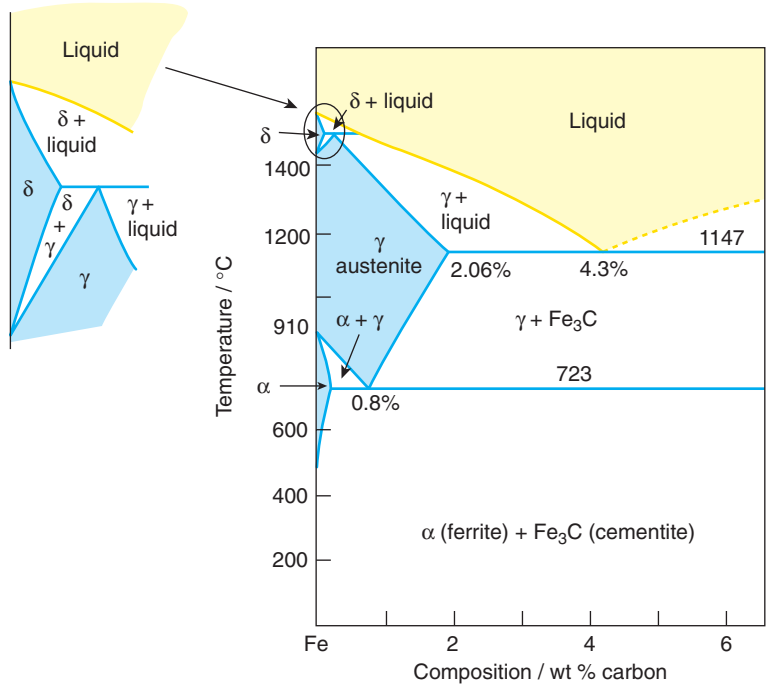


Figure 7.22 ***The Fe–C diagram.***

occur. In practice, this eutectoid decomposition starts at the boundaries of the austenite grains; the *ferrite* (α-Fe) and *cementite* (Fe_3C) crystals grow side by side to give a lamellar texture known as *pearlite*.

If the steel is cooled quickly, there is not enough time for decomposition to ferrite and cementite to occur; instead, *martensite* forms. Martensite has a deformed austenite structure in which the carbon atoms are retained in solid solution. It is possible to release these carbon atoms, as cementite, by tempering, i.e. reheating, to give a fine scale pearlite texture.

The hardness of steel depends on the cooling conditions and/or tempering treatment. Martensitic steels are hard, largely, it seems, because of the stressed state of the martensite crystals which prevents easy motion of dislocations. In steels with pearlite texture, the hardness depends on the size, amount and distribution of the very hard cementite grains; a fine texture with a large number of closely spaced grains gives hard steel. If the steel is cooled slowly or held just below 723 °C, the decomposition is slow, giving a coarse pearlite texture. A finer texture is obtained by using a faster cooling rate to produce martensite, which is then subsequently tempered at a low temperature, e.g. 200 °C.

7.3.7.2 The system $CaO–SiO_2$: cement manufacture

A phase diagram which is one of the most important in silicate technology is the system $CaO–SiO_2$, part of which is shown in Fig. 7.23. The compound Ca_2SiO_4 ($\equiv 2CaO{\cdot}SiO_2 \equiv 33.3\%SiO_2{\cdot}66.7\%CaO$) melts congruently at 2130 °C. The compound Ca_3SiO_5 melts incongruently at 2150 °C to CaO and liquid. It also has a lower limit of stability at 1250 °C, decomposing to give a mixture of CaO and Ca_2SiO_4 at lower temperatures. This part of the diagram has one peritectic, P, and one eutectic, E.

The $CaO–SiO_2$ phase diagram is critically important to cement manufacture. The key ingredient in rapid-hardening Portland cement is Ca_3SiO_5, but from the phase diagram, it is unstable below 1250 °C. In order

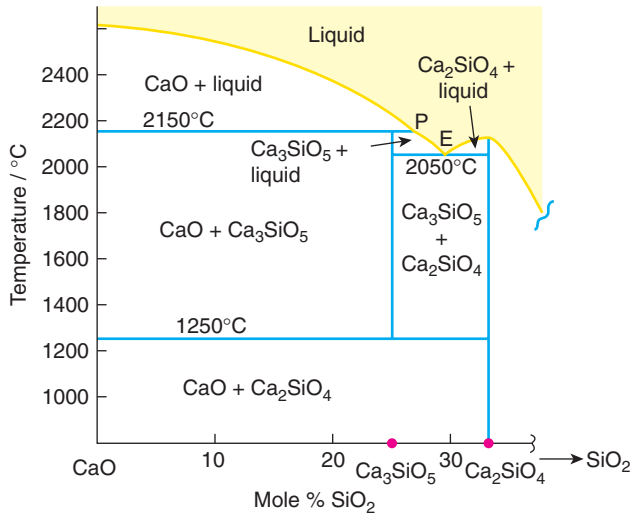


Figure 7.23 ***Partial diagram for lime-rich compositions in the system $CaO–SiO_2$.***

to obtain it, therefore, the reactants are fired in a kiln at 1400–1500 °C and cooled rapidly in a blast of air. With such rapid cooling rates, there is no time for the Ca_3SiO_5 to decompose and it is preserved to room temperature, where it is kinetically stable. On subsequent reaction with water, it forms a semicrystalline calcium silicate hydrate gel which is a key hardening component of set cement and concrete.

7.3.7.3 The system Na–S: Na/S batteries

The system Na–S is of practical importance in the operation of Na/S batteries or 'beta batteries' which are used on the megawatt scale for power station load – levelling applications. The phase diagram for S-rich compositions is shown in Fig. 7.24. It has a liquid immiscibility dome (upper consolute temperature not known) with a monotectic condition at 240 °C and several crystalline Na sulfides. Na_2S_5 and Na_2S_4 both melt congruently whereas Na_2S_2 melts incongruently to Na_2S and liquid. The liquidus surface therefore has three eutectics, one monotectic and one peritectic in the diagram shown.

Molten S forms the cathode in Na/S batteries which operate at 300–350 °C; the discharge temperature and products are represented by the shaded region in Fig. 7.24. During discharge, with a general cell reaction

$$x\text{Na} + \text{S} \rightarrow \text{Na}_x\text{S (liq)}$$

Na^+ ions diffuse through the β-alumina electrolyte (see Section 8.5.4.11) and react with the S anode to form molten sodium sulfides. It is necessary to avoid formation of crystalline Na sulfides since these would 'freeze' the cell, leading to a collapse in performance; hence discharge is permitted to continue until Na_2S_2 starts to precipitate when the composition of the molten sodium sulfides crosses the liquidus surface at, e.g., ~57% S at 320 °C.

The cell voltage as a function of degree of discharge is shown in Fig. 7.24; in the early stages of discharge, the products are a mixture of two liquids and the emf is constant, 2.08 V. With further discharge, the product is a single phase liquid and emf varies with composition over the range 2.08–1.78 V. In the later, deep discharge stage, when Na_2S_2 precipitates, the emf is again constant at 1.78 V.

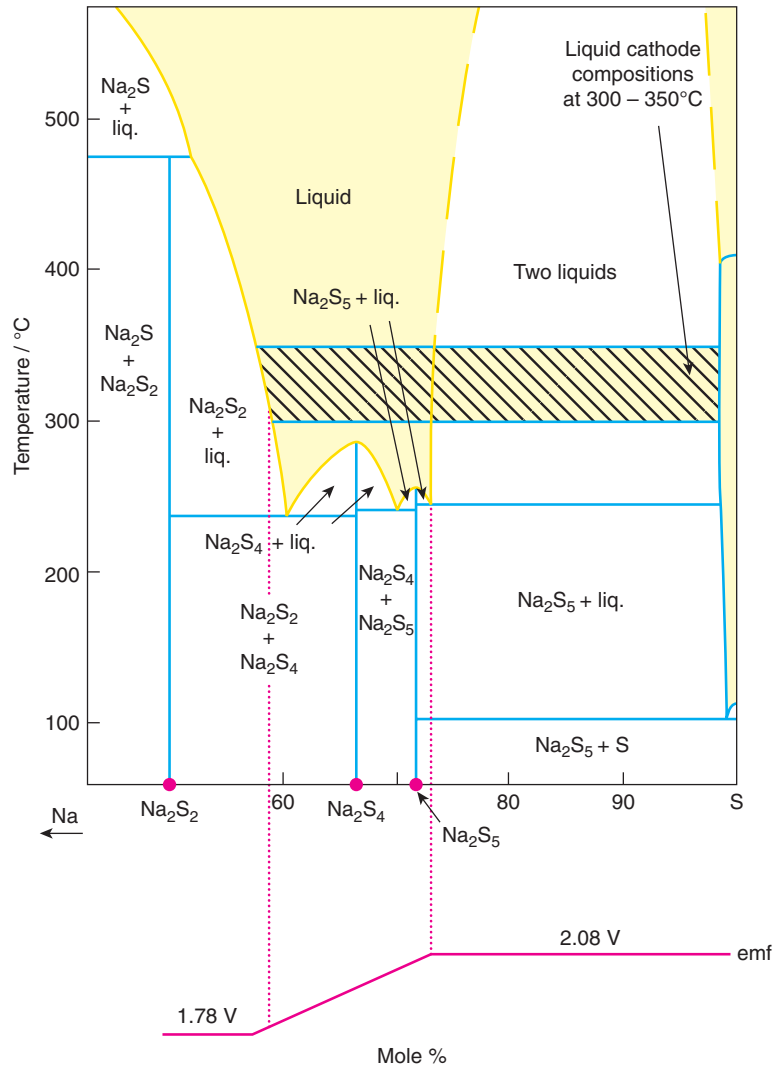


Figure 7.24 ***Na–S phase diagram and open-circuit cell voltage as a function of degree of discharge/cathode composition for the Na/S cell.***

7.3.7.4 The system Na_2O–SiO_2: glass making

Glass for windows, bottles and many other uses is based on silica, SiO_2, to which various other oxides have been added, mainly to reduce melting temperature, increase the melt fluidity and decrease the tendency to devitrification. Pure SiO_2 melts above 1700 °C to give a highly viscous, polymeric liquid which readily solidifies into a glass (an amorphous solid without 3D order in its structure) on cooling. Because of its very high melting temperature, the high viscosity and low workability of the melt, even at 2000 °C, glasses made from pure SiO_2 are expensive, notwithstanding their valuable properties such as a high softening temperature, ~1200 °C. Pure SiO_2 is a speciality, expensive glass, therefore.

The effect of adding Na_2O to SiO_2 is shown by the Na_2O–SiO_2 phase diagram, Fig. 7.25. With 30% Na_2O added, the melting temperatures have decreased from ~1700 to ~800 °C. The liquids still form a glass on

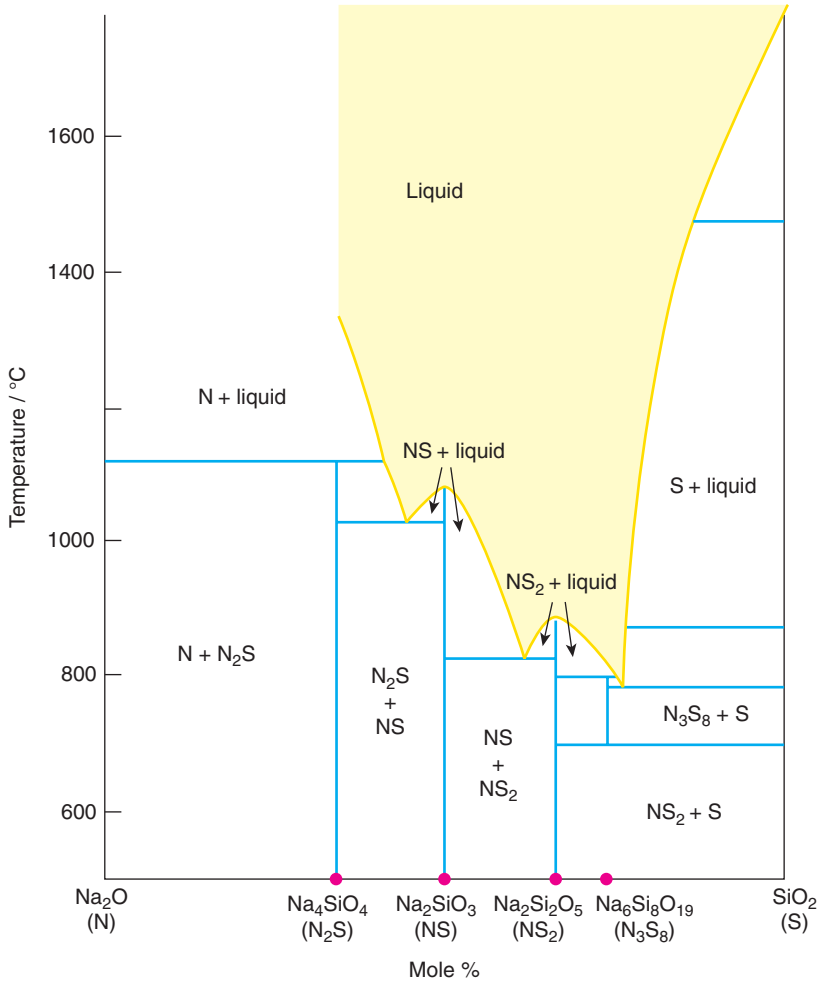


Figure 7.25 ***Na_2O–SiO_2 phase diagram. $N = Na_2O$, $S = SiO_2$, $N_2S = Na_4SiO_4$, $NS = Na_2SiO_3$, $NS_2 = Na_2Si_2O_5$ $N_3S_8 = Na_6Si_8O_{19}$.***

cooling and because they have much greater fluidity than SiO_2 liquid, the glasses can be readily shaped. There is, however, one major difficulty: Na_2O–SiO_2 glasses are water soluble! Windows made from sodium silicate glass may be perfectly suitable in the Atacama Desert but would not last very long in Sheffield. In practice, atmospheric stability is achieved by adding other oxides, especially CaO, MgO, Al_2O_3 and B_2O_3; the resulting glasses are not water soluble and do not readily devitrify (crystallise). Thus, although commercial glasses are compositionally complex, the effect of modifying the composition so as to reduce melting temperatures can be appreciated from Fig. 7.25. In order to understand the variation in melt fluidity and its structure in terms of the degree of polymerisation of the silicate component, see Section 1.17.16.

7.3.7.5 The system Li_2O–SiO_2: metastable phase separation and synthetic opals

The system Li_2O–SiO_2 is a glass-forming system at high silica compositions, similar to Na_2O–SiO_2, but glasses with compositions around 20 mole % Li_2O have beautiful opalescent colours (orange in transmitted

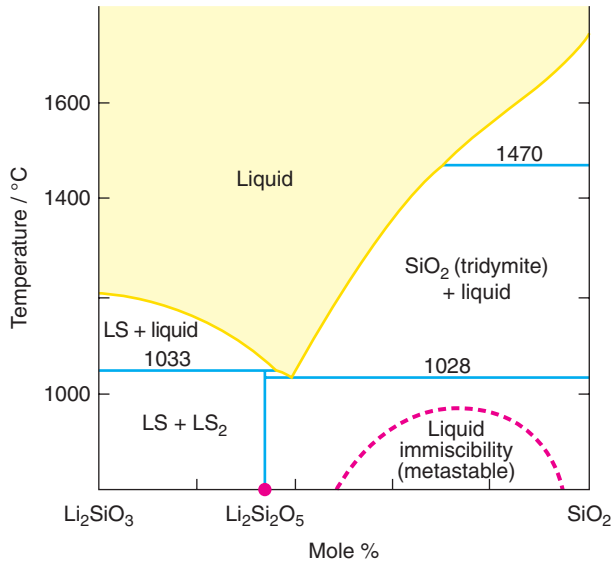


Figure 7.26 ***Li_2SiO_3–SiO_2 phase diagram. LS = Li_2SiO_3, LS_2 = $Li_2Si_2O_5$. The existence of a metastable immiscibility dome in rapidly cooled liquids that have avoided crystallisation is shown schematically, dashed.***

light, blue in reflected light). The Li_2O–SiO_2 phase diagram, Fig. 7.26, shows an unusual-shaped liquidus at high silica compositions. This is because although there is no stable liquid immiscibility dome in this region, there is a tendency, at the local structure level, for lithia-rich and silica-rich clusters to form in the liquid. On cooling such liquids, this tendency to cluster formation increases (associated with a reduction in $T\Delta S$ and hence reduced driving force for mixing of the liquids). On cooling rapidly, crystallisation is suppressed, and the undercooled liquids enter an immiscibility dome with an upper consolute temperature of ~950 °C. Once inside the dome, they spontaneously break up into two liquids with a very fine texture (the *spinodal decomposition* mechanism). The resulting liquids, of compositions ~95% SiO_2 and ~25% Li_2O·75% SiO_2, have different refractive indices and this is the origin of their opalescence.

The immiscibility dome in the Li_2O–SiO_2 glasses is entirely metastable and is observed only in undercooled liquids that have departed from the expected phase diagram behaviour as a consequence of being cooled rapidly to avoid crystallisation. Although it is not an equilibrium property, we can nevertheless understand it in terms of typical phase diagram behaviour: a range of single phase liquids enters an immiscibility dome on cooling. It also illustrates that the world would be a less colourful place if everything behaved within the constraints of thermodynamic equilibrium!

7.3.7.6 Purification of semiconducting Si by zone refining

Ultrahigh-purity Si is essential for semiconductor applications. It is obtained in the zone refining method by a process of repeated melting and crystallisation. On crystallisation, impurities tend to concentrate in the remaining liquid; by carrying out the crystallisation in a temperature gradient, impurities are effectively 'swept out of the Si'. This is illustrated by the schematic phase diagram in Fig. 7.27. Si forms a solid solution with an impurity (unspecified); solidus and liquidus temperatures decrease smoothly. On partially melting composition a, at temperature T, there is a greater concentration of impurities in the liquid than in the overall composition a and, provided that the liquid, c, can be separated from the solid at this temperature, the remaining solid, b, is of higher purity than the starting solid, a. Further comments on the application of the zone refining method are given in Section 4.6.3.

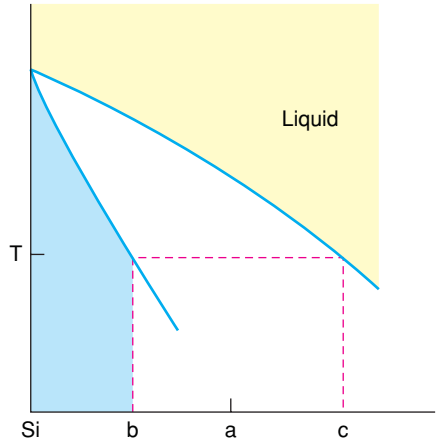


Figure 7.27 ***Purification of Si by zone refining: impurities concentrate in the melt.***

7.3.7.7 The system ZrO_2–Y_2O_3: yttria-stabilised zirconia, YSZ, solid electrolyte

Zirconia, ZrO_2, is potentially a very useful ceramic material with a high melting point of ~2700 °C but on cooling it undergoes a series of phase transitions:

$$\text{cubic (fluorite)} \xrightarrow{2400\,^\circ\text{C}} \text{tetragonal} \xrightarrow{1050\,^\circ\text{C}} \text{monoclinic (baddeleyite)}$$

The tetragonal to monoclinic transition is associated with an increase in unit cell volume, and therefore sample volume, by ~9%; consequently, ceramic bodies fabricated at high temperature shatter on cooling. Fortunately, these transitions can be avoided by partly replacing Zr with aliovalent cations such as Mg, Ca and Y to give solid solutions such as $Zr_{1-x}Mg_xO_{2-x}$ and $Zr_{1-y}Y_yO_{2-y/2}$.

The ZrO_2–Y_2O_3 (given as $YO_{1.5}$) phase diagram is shown in Fig. 7.28. Both phase transition temperatures decrease rapidly with increasing $YO_{1.5}$ content. With ≥~15% $YO_{1.5}$, the cubic polymorph is stabilised to room temperature, to give the material known as YSZ, or yttria-stabilised zirconia. With smaller yttria contents, the tetragonal polymorph is partially stabilised down to a eutectoid at ~500 °C, below which it decomposes to a mixture of monoclinic and cubic polymorphs under conditions of thermodynamic equilibrium. In practice, both partially stabilised tetragonal zirconia and fully stabilised cubic YSZ are used for various applications.

YSZ is a particularly important material since not only is it a stable ceramic over an enormous temperature range, it is also an excellent oxide ion conductor by virtue of the solid solution formation mechanism which leads to creation of oxide ion vacancies. These vacancies are mobile at high temperatures, leading to applications of YSZ as an oxide ion conductor in fuel cells and sensors (see Chapter 8).

7.3.7.8 The system Bi_2O_3–Fe_2O_3: multiferroic $BiFeO_3$

The system Bi_2O_3–Fe_2O_3 contains the phase $BiFeO_3$, which is currently of great interest as a multiferroic material which is both magnetic and ferroelectric. The phase diagram, Fig. 7.29, shows that $BiFeO_3$ melts incongruently at 933 °C to give $Bi_2Fe_4O_9$ + liquid. $Bi_2Fe_4O_9$ itself then melts incongruently to Fe_2O_3 + liquid at 961 °C. Phases that melt incongruently, such as $BiFeO_3$, are often difficult to synthesise phase-pure by solid state reaction and this is indeed the case with $BiFeO_3$. It is very common that synthesis products contain $BiFeO_3$ as the majority phase but with small amounts of other phases which are extremely difficult

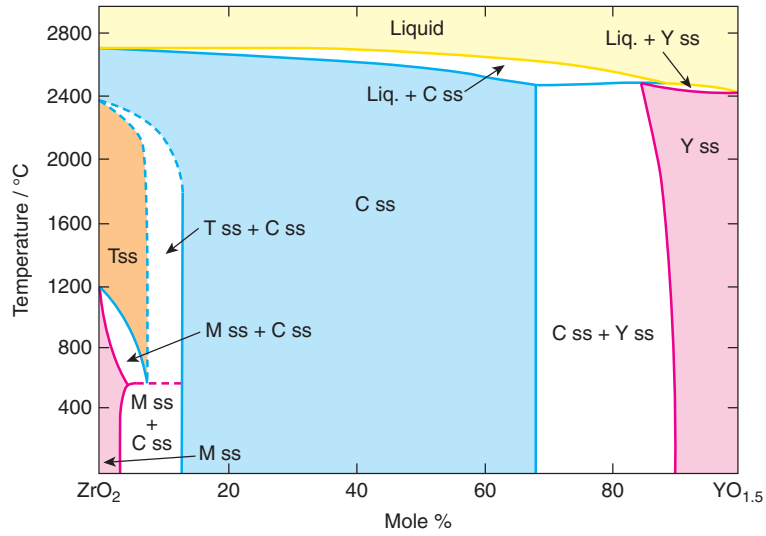


Figure 7.28 ***ZrO_2–Y_2O_3 phase diagram. M, T and C refer to the monoclinic, tetragonal and cubic polymorphs of zirconia, and their solid solutions, ss. Y = yttria, Y_2O_3.***

to eliminate by further heat treatment. Synthesis of phase-pure $BiFeO_3$ has tested the ingenuity of scientists and a survey of many of the methods that have been used are described in Section 4.3.8.

7.4 Some Tips and Guidelines for Constructing Binary Phase Diagrams

This section is directed towards the information that is required to construct a phase diagram, and how it may be obtained. First, we consider the nature of reaction products and whether they are pure phases or phase

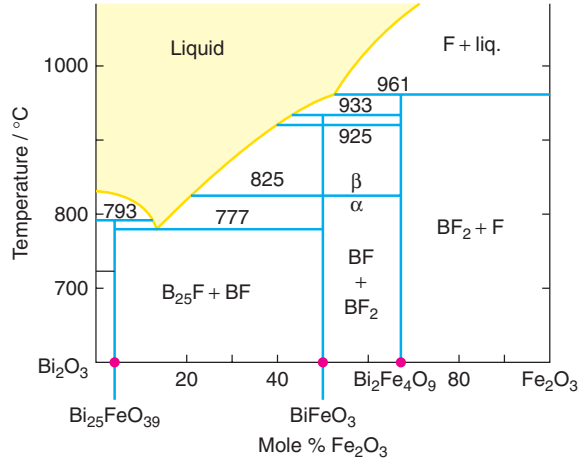


Figure 7.29 ***Bi_2O_3–Fe_2O_3 phase diagram. Adapted with permission from E. I. Speranskaya et al.,*** **Bull. Acad. Sci. USSR,** 5 ***873–4 © 1965.***

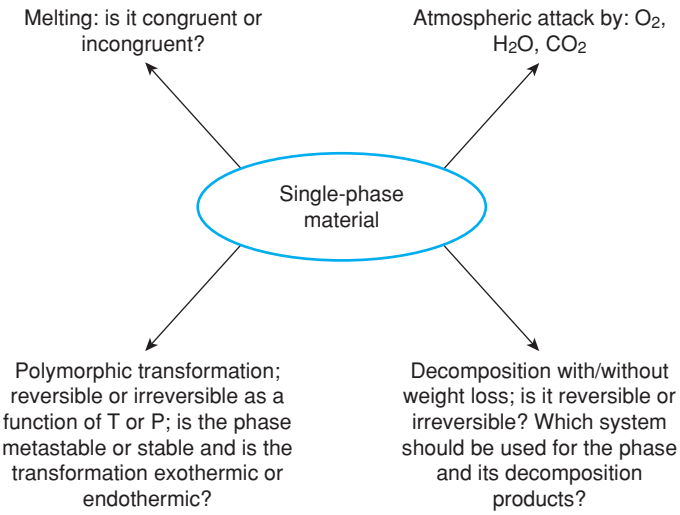


Figure 7.30 ***Summary of the changes that can occur to a single-phase material as a consequence of either heating or atmospheric attack.***

mixtures. The main question, as always, is: Does the phase or phase assemblage that I have obtained represent thermodynamic equilibrium under the conditions used for reaction?

The experimental tests that can be applied, and the information that can be obtained from the results, revolve around whether or not changes occur with changing experimental conditions. Thus, if a mixture of phases is found to be present at the end of a reaction, is this phase assemblage independent of increased reaction time or raising the reaction temperature? If changes occur with time, it is unlikely that equilibrium has been reached; if changes occur with increasing temperature, it is necessary to determine whether these are reversible on returning to the original temperature or irreversible. If the changes are irreversible, the next question to ask is whether this is because the original sample was not in equilibrium or because the reverse process takes place too slowly to be observed on the timescale of the experiment?

Starting with a single-phase material, there are four main things that can happen to the phase as a function of temperature, time and atmosphere. These are summarised in Fig. 7.30. To study these various possibilities, a furnace is needed, preferably with a controlled atmosphere and access to powder XRD for phase analysis; it can also be very useful to have access to TG to study weight changes, usually on heating and DTA/DSC, to study exothermic and endothermic events on heating or cooling. A variety of changes may occur to our phase, Fig. 7.30; if these are reversible on changing the experimental conditions, then they may represent thermodynamic equilibrium or at least a metastable equilibrium (Section 6.4).

Many new materials that are prepared by chimie douce methods are either stable thermodynamically only at low temperatures or are stable kinetically at low temperatures but thermodynamically, are only metastable. If our phase does not undergo atmospheric attack, decomposition or polymorphic transformations then the only question is: What is its melting temperature, does it melt congruently or incongruently or does it decompose prior to melting? The type of melting can be determined by DTA/DSC linked to XRD of the products of crystallisation on cooling.

There are four main processes that can occur on cooling a melt. First, the melt may not crystallise but instead may solidify to form a glass; in this case, the product will be X-ray amorphous and can be confirmed as such using optical microscopy. Second, if melting is congruent then it is likely that recrystallisation of the

original phase will occur and can be confirmed by XRD. Third, if melting is incongruent, recrystallisation of the original phase is unlikely to occur during the timescale of a rapid cool and a mixture of phases may result. A fourth possibility is that the phase may decompose, especially if it is an oxysalt; this may occur before or during melting, and again, the product after quenching is likely to be different from that of the starting phase. If decomposition is suspected, then TG can be a useful technique (provided that the TG instrument is not attacked by the sample during/after decomposition). The combination of TG with DSC/DTA is very useful for separating transformations and reactions that involve weight change from those which do not (Section 6.4).

A useful test to help decide whether or not a sample has changed on heating is to quench samples rapidly to room temperature at the end of a heat treatment and compare the results with those obtained on samples that have been cooled more slowly. Quenching can be carried out by dropping samples, wrapped in Pt foil envelopes, into liquid N_2 (liquid Hg is a better quench medium because of its high thermal conductivity but health and safety concerns over possible vapourisation of Hg do not permit its use these days).

Quenching experiments, followed by analysis of the products at room temperature, may permit a direct answer to the question: is the sample at room temperature the same as the one at the temperature of heat treatment or have changes occurred during cooling? Examples of processes or reactions that occur on cooling and therefore may depend on cooling rate, are as follows:

- Samples may lose a small amount of oxygen at high temperatures which is recovered during a slow cool but not on rapid quenching, and this can have a dramatic effect on the electrical properties.
- Many polymorphic transformations may occur during slow cooling but can be prevented by quenching (Section 6.4).
- Some solid solution compositions become supersaturated during cooling and whether or not precipitation of a second phase occurs may be controlled by the cooling rate.
- In addition to possible reoxidation during slow cooling, many materials are hygroscopic or sensitive to CO_2 and the products at room temperature may depend on both cooling rate and atmospheric conditions during cooling.

Once various tests have been conducted to determine the stability and melting behaviour of a phase, attention can be paid to the melting of phase mixtures. For this, DTA/DSC is again very useful, particularly for establishing solidus temperatures since these should represent the first endothermic peak on heating. It is then relatively easy to obtain a good estimate of the behaviour of the various two-phase regions and thereby construct the rest of the phase diagram.

In summary, the key point in determining phase diagrams is first to ascertain the behaviour of pure compounds or phases and determine the conditions under which they are thermodynamically stable. Thus: if changes subsequently occur to a material that is already single phase then you can be fairly sure that they represent changes in thermodynamic stability; if, however, changes occur to a mixture of phases, it may not be clear whether these represent a change in thermodynamic stability or kinetic changes to a sample that has not, yet, finished reacting. Second, once the behaviour of single-phase materials has been established, it should be relatively straightforward to determine the behaviour, in particular the melting behaviour, of phase mixtures. In cases where the vapour phase is important, such as if the system contains transition elements with variable oxidation states, it is clearly essential to establish to which system the phase diagram refers since what may, at first sight, be a binary system may in fact be part of a ternary system if oxygen is an important compositional variable.

8

Electrical Properties

8.1 Survey of Electrical Properties and Electrical Materials

The electrical properties of materials depend on whether they are conductors or dielectrics (insulators) and, if the former, whether the current carriers are electrons or ions. A wealth of electrical phenomena occur and find a huge variety of applications. Here, we are concerned with the chemical aspects of electrical properties, which primarily involve the relation between crystal structure and properties.

Electronic conductivity occurs to varying extents and by different mechanisms in a wide variety of materials. It is responsible for the characteristic electrical properties of metals, superconductors and semiconductors. *Metallic conductivity* has certain characteristics:

(a) A significant proportion of the outer shell or valence electrons are free to move throughout the structure and are completely delocalised.
(b) Collisions between these electrons and *phonons* (lattice vibrations) are responsible for the residual resistance to current flow and associated heat losses.
(c) Metallic bonding and conductivity are described using band theory.
(d) Metallic conductivity occurs in metals, alloys and a variety of inorganic solids, such as certain oxides and sulfides. In the latter, the metal atoms form a sublattice with overlapping valence orbitals. Metallic conductivity also occurs in conjugated organic systems such as polyacetylene and polyaniline.

The closely related phenomenon of *superconductivity* has the following characteristics:

(a) Valence shell electrons again appear to be delocalised.
(b) The electrons move cooperatively, perhaps as loosely coupled pairs.
(c) No electron–phonon collisions occur, there is no heat loss and there is no resistance to current flow.
(d) Until 1986, superconductivity was confined to very low temperatures, <23 K, but is now known in 'high-T_c ceramic superconductors' at temperatures up to 138 K (and 166 K under very high pressures).

Semiconductivity has the following characteristics:

(a) It is associated with a limited degree of electronic conduction.
(b) It is intermediate between metallic conduction, in which a significant number of outer shell electrons are free to move, and insulating behaviour, in which all the valence electrons are either tightly bound to individual atoms or localised in bonds between atoms.

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

(c) It is common in transition metal compounds (apart from metallic ones), Si, Ge and related materials and some organic solids, e.g. anthracene.
(d) It either may be regarded as an electron hopping process, or may be described using band theory, depending on the particular material.
(e) The number of electrons involved depends on temperature and impurity level, in contrast to metallic conductivity.
(f) The ability to dope semiconductors and modify their properties gives rise to many applications and is the basis of the microelectronics industry.

Ionic conductivity occurs in materials known variously as solid electrolytes, superionic conductors or fast ion conductors:

(a) These have a rigid framework structure within which one set of ions forms a mobile sublattice.
(b) Conductivities as high as 1 Ω^{-1} cm^{-1}, similar to those of strong liquid electrolytes, occur.
(c) Solid electrolytes are intermediate between typical ionic solids, in which all ions are fixed on their lattice sites, and liquid electrolytes, in which all ions are mobile.
(d) The structural requirements for high ionic conductivity are (i) there must be available empty sites for ions to hop into and (ii) the energy barrier for ions to hop between sites is small.
(e) High levels of ionic conduction are rare but modest amounts are common, especially in non-stoichiometric or doped materials.

Dielectric materials, or insulators, are characterised by the complete absence of electrical conductivity, both electronic and ionic. Their bonding may be strongly ionic, as in MgO and Al_2O_3, strongly covalent, as in diamond, C, or strongly polar covalent, as in SiO_2.

Typical values for ionic and electronic conductivities are given in Table 8.1. Conductivities increase with increase in temperature for all materials except metals and superconductors: in superconductors, the conductivity is independent of temperature since the resistance is zero; in metals, the conductivity decreases gradually with increase in temperature due to electron–phonon collisions.

Table 8.1 ***Typical values of electrical conductivity***

Property	Material	σ/Ω^{-1} cm^{-1} [a]
Ionic conduction	Ionic crystals	$<10^{-18}$–10^{-4}
	Solid electrolytes	10^{-3}–10^{1}
	Strong (liquid) electrolytes	10^{-3}–10^{1}
Electronic conduction	Metals	10^{-1}–10^{5}
	Semiconductors	10^{-5}–10^{2}
	Insulators	$<10^{-12}$

[a]Units of *conductivity* are defined as follows. The *resistivity* of a solid is obtained by first measuring its *resistance* and then correcting for its *geometric factor*:

$$\text{resistivity}\rho\ (\Omega\ \text{cm}) = \text{resistance}\ (\Omega) \times \frac{\text{area}\ (\text{cm}^2)}{\text{thickness}\ (\text{cm})}$$

The conductivity σ, with units Ω^{-1} cm^{-1} or siemens, S cm^{-1}, is the reciprocal of the resistivity.

Ferroelectricity has the following characteristics:

(a) It is intermediate between solid electrolyte behaviour, in which ions migrate over long distances, and dielectric behaviour, in which there is no net displacement of ions from their regular lattice sites; in ferroelectrics, certain ions can be displaced by ~0.1 Å.
(b) These limited atomic displacements lead to a net polarisation and creation of a dipole moment.
(c) Ferroelectricity occurs in a small group of materials: e.g. in perovskite, $BaTiO_3$, at room temperature, Ti atoms are displaced towards one corner of the TiO_6 octahedra because Ti is slightly too small for a regular octahedral site.
(d) It is related to *pyroelectricity* and *piezoelectricity*. In pyroelectricity, ionic displacements occur spontaneously and vary with temperature; in piezoelectricity, the displacements also occur spontaneously but vary under an applied pressure; in ferroelectricity, the spontaneous displacements may be reversed in an applied electric field, giving rise to *hysteresis*.

8.2 Metallic Conductivity

The description of bonding in metals in terms of band theory is outlined in Chapter 3. The two theoretical approaches to model metallic conductivity, the *nearly free electron model* and the *tight binding model*, both give rise to a qualitatively similar, pictorial description of metallic bonding. This can be extended to both semiconductors and insulators with the proviso that not all energies are available to electrons; instead, in the plots of energy versus density of states, forbidden gaps or band gaps exist and are responsible for a fundamental difference between metals on the one hand and semiconductors/insulators on the other. Specifically, for conduction in semiconductors and insulators to occur, it is necessary first to promote electrons across the forbidden gap, which is an activated process and therefore, very temperature dependent. In metals there is no band gap and no necessity to promote electrons across a band gap.

Although metals are excellent conductors of electricity, they have a residual resistivity associated with two sources: first, electrostatic interactions between mobile electrons and lattice vibrations, ρ_{vib}, known as *electron–phonon collisions*; second, collisions between mobile electrons and defects or impurities in the metal crystal lattice, ρ_{def}. These two effects are combined in Matthiesen's rule:

$$\rho = \rho_{vib} + \rho_{def} \tag{8.1}$$

The resistivity of metals increases gradually with increase in temperature due to the increase in electron–phonon collisions, Fig. 8.1; nevertheless, the *mean free path* between collisions is typically several hundred atomic spacings at room temperature and can extend to centimetre lengths at very low temperatures. The number of electrons that contribute to metallic conductivity is relatively small and is limited to those with energies close to the Fermi energy, E_F; experimental evidence for this comes from measurements of the specific heat of metals whose values can be attributed almost entirely to lattice vibrations with only a limited electronic contribution. Conductivity data at room temperature for a selection of metals are summarised in Table 8.2.

Metallic conductivity is not confined to metals but is also exhibited by inorganic materials such as certain oxides and organic materials such as certain polymers. It is shown in Chapter 3 how, in transition metal-containing oxides, *d*-orbital overlap gives rise to bands of energy levels for the *d* electrons; provided that certain conditions are met regarding the strength of overlap and the number of *d* electrons available, metallic conduction may occur. Good examples are titanium monoxide, TiO, and tungsten bronzes such as Na_xWO_3.

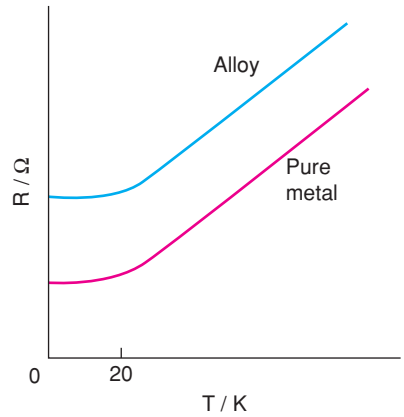


Figure 8.1 ***Resistivity of metals, which typically is constant below ~20 K and increases linearly with temperature above 20 K.***

The possibility of preparing electrically conducting polymers or 'organic metals' is now a reality. Such materials combine the mechanical properties of polymers – flexibility, ease of fabrication as thin films – with the high conductivity normally reserved for metals. There are two main categories of 'organic metal': conjugated systems and charge-transfer complexes. There is also much current interest in two new forms of carbon, carbon nanotubes and graphene, both for their electrical and mechanical properties.

8.2.1 Organic metals: conjugated systems

8.2.1.1 Polyacetylene

Organic solids are usually insulators. Electrons cannot move freely either within molecules or from one molecule to another. Exceptions are conjugated systems with a skeleton of alternate double and single C–C bonds, as in graphite. Polymers such as polyethylene are insulators because although the precursor, ethylene, contains a double bond, polyethylene itself is saturated and contains only single bonds, Fig. 8.2(a).

A conjugated long-chain polymer with the potential for electrical conductivity is polyacetylene. The acetylene precursor contains a –C≡C– triple bond whereas polyacetylene contains alternate single and

Table 8.2 ***Conductivities of some metals at 25 °C***

Metal	$10^{-5}\sigma/\Omega^{-1}$ cm^{-1}
Cu	5.9
Ag	6.2
Na	2.1
Al	3.7
Fe	1.0
Zn	1.7
Cd	1.4
Pb	0.5

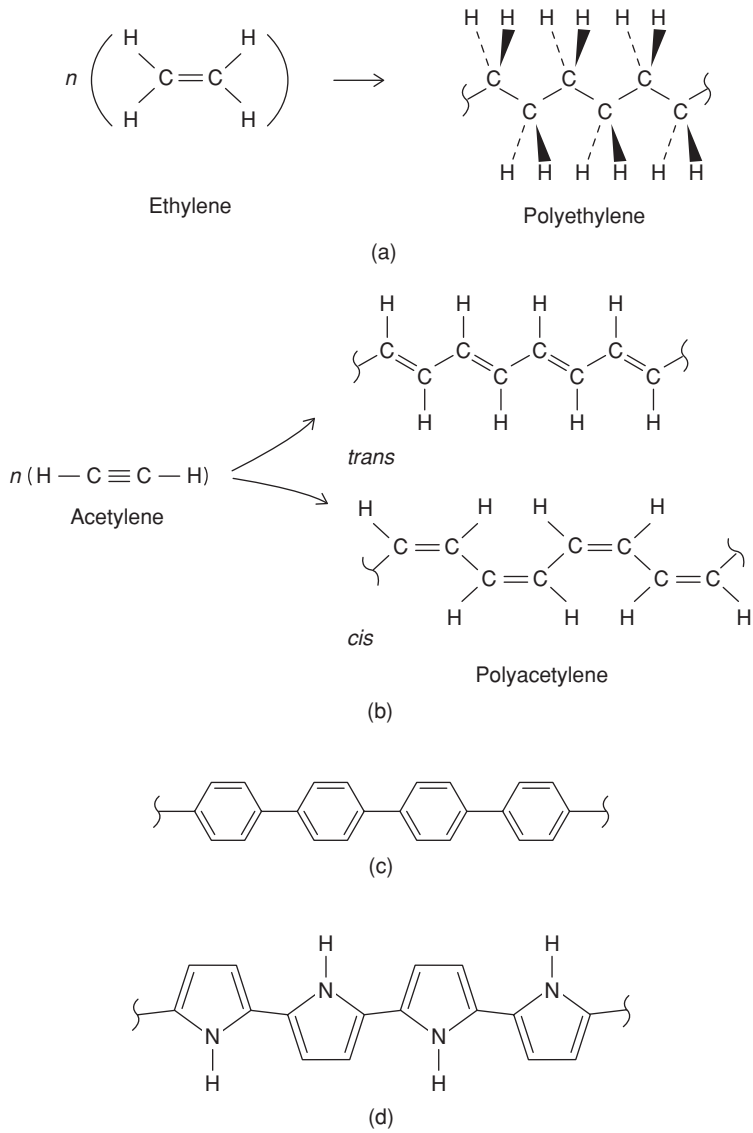


Figure 8.2 ***(a) Polyethylene, (b) polyacetylene, (c) poly-p-phenylene and (d) polypyrrole.***

double bonds, Fig. 8.2(b). In fact, pure polyacetylene has a modest electrical conductivity, 10^{-9} Ω^{-1} cm^{-1} (*cis* form) to 10^{-5} Ω^{-1} cm^{-1} (*trans* form), comparable to those of semiconductors such as pure Si. These conductivity values are low because the π electron system is not completely delocalised in polyacetylene. It has a band gap of 1.9 eV.

A significant discovery, by MacDiarmid, Heeger and Shirakawa, Nobel Prize in Chemistry, 2000, was that on doping polyacetylene with suitable inorganic compounds, its conductivity increases dramatically. With dopants such as (a) Br_2, SbF_5, WF_6 and H_2SO_4, all of which act as electron acceptors to give, e.g., $(CH)_n^{\delta+}Br^{\delta-}$, and (b) alkali metals, which act as electron donors, conductivities as high as 10^3 Ω^{-1} cm^{-1}

in *trans*-polyacetylene are obtained. This is similar to the conductivity of many metals and, indeed, such materials have been termed 'synthetic metals'. The conductivity increases rapidly as dopant is added and a semiconductor-to-metal transition occurs with 1–5 mole % of added dopant.

Polyacetylene is prepared by catalytic polymerisation of acetylene in the absence of oxygen. A Ziegler–Natta catalyst may be used, which is a mixture of $Al(CH_2CH_3)_3$ and $Ti(OC_4H_9)_4$. In one method, acetylene is bubbled through a solution of the catalyst and polyacetylene precipitates. In another method, acetylene gas is introduced into a glass tube whose inner surface is coated with a thin layer of catalyst; a layer of polyacetylene forms on the surface of the catalyst. It is usually desired to prepare the *trans* form because it has higher conductivity. It may be prepared directly, at 100 °C, or by heating the *cis* form, which changes rapidly to the *trans* form on heating to ~150 °C. Doping may be achieved by exposing polyacetylene to gaseous or liquid dopant.

The electronic structure of polyacetylene films is unclear, especially the mechanism of electron transfer between polyacetylene chains. The films have a complex morphology in which polyacetylene chains appear to fold up to form plates and the plates overlap to form fibres. More work is needed to characterise and control the film texture better and to relate specific textural features to the desirable, high electronic conductivity. Conducting polyacetylene has a variety of potential applications. Heavily doped material could be used in place of metals for certain electrical applications. Undoped or lightly doped materials could be used in place of semiconductors. Thus, *p–n diode junctions* can be fabricated in which two films of polyacetylene are brought into contact, one doped with an electron acceptor and p-type, the other doped with an electron donor and n-type (see the next section). Such devices are easy to make and, with their large surface areas, have potential application in solar energy conversion. One problem that remains to be solved is the sensitivity of polyacetylene to oxygen. Perhaps substituted or modified polyacetylenes will be made which retain the high conductivity but are not liable to atmospheric attack.

The above discussion is concerned with electronic conductivity. Ionic conductivity is also possible and certain doped polyacetylenes may be used as reversible electrodes in new types of battery. The doping is carried out electrochemically. In one arrangement, a polyacetylene film is dipped into a liquid electrolyte composed of $LiClO_4$ dissolved in propylene carbonate. A lithium metal electrode also dips into the electrolyte. On charging the cell to 1.0 V, at room temperature, perchlorate ions from the electrolyte enter the polyacetylene electrode to give $(CH)_y^+(ClO_4)_y^-$, where y is in the range 0–0.06. At the same time, electrons leave the polyacetylene electrode via the external circuit and Li^+ ions are deposited at the lithium electrode to preserve overall charge balance in electrodes and electrolyte. The perchlorate ions enter the polyacetylene structure reversibly and subsequently diffuse back into the electrolyte on discharge. The polyacetylene therefore behaves as a mixed ionic–electronic conductor.

The possibility of using polymeric electrodes in solid state batteries is very attractive because (a) the polymers are very light (compared with lead in lead/acid batteries) and (b) polymers have a flexible structure, which should reduce contact resistance problems at the electrode–electrolyte interfaces and allow the fabrication of wafer-thin, flexible batteries.

*8.2.1.2 Poly-*p*-phenylene and polypyrrole*

A long-chain polymer with considerable potential, poly-*p*-phenylene, Fig. 8.2(c), consists of a long chain of benzene rings. It has been studied less extensively than polyacetylene but has been doped successfully to increase its electronic conductivity, e.g. on doping with $FeCl_3$, a product, $[C_6H_4(FeCl_3)_{0.16}]_x$, with conductivity 0.3 S cm^{-1} at 25 °C was obtained.

Pyrrole, C_4H_5N, can be polymerised to give a long-chain structure with alternate double and single bonds, and a delocalised π-electron system (d). Polypyrrole has low conductivity but can be oxidised by perchlorate to give p-type conductivity as high as $10^2\ \Omega^{-1}\ cm^{-1}$. It is stable in air and at temperatures up to 250 °C.

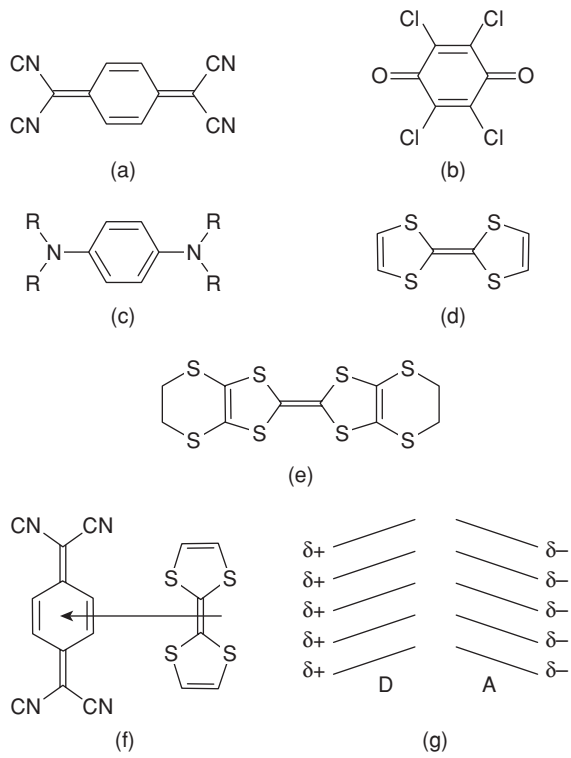


Figure 8.3 ***(a) Tetracyanoquinodimethane (TCNQ), (b) chloroanil, (c) p-phenylenediamines, (d) tetrathiafulvalene (TTF), (e) bisethylenedithiotetrathiafulvalene (BEDT-TTF, also known as ET), (f) a TTF–TCNQ donor–acceptor interaction and (g) adjacent donor and acceptor stacks in ID organic metal.***

8.2.2 Organic metals: charge-transfer complexes

Some two-component organic systems in which one component is a π-electron donor and the other an electron acceptor behave as highly conducting, synthetic metals and a few are even superconducting at very low temperatures. Although many examples of organic charge-transfer complexes are known, in order to achieve long-range conduction, it appears to be essential that the donors and acceptors form separate, alternating stacks. Hence the donor stacks transfer some charge to the acceptor stacks, giving a partially occupied, π electron band.

Examples of good π-electron acceptors, Fig. 8.3, are (a) TCNQ, tetracyanoquinodimethane, and (b) chloranil. TCNQ is one of the best acceptors found so far as it is a planar, conjugated molecule with strong, electron-withdrawing CN groups. Examples of good π-electron donors are (c) *p*-phenylenediamine, with R = H; TMPD, tetramethyl-*p*-phenylenediamine, with R = CH_3, (d) TTF, tetrathiafulvalene, and (e) BEDT-TTF, bisethylenedithiotetrathiafulvalene. A single donor–acceptor interaction between TTF and TCNQ molecules is shown schematically in (f) and an array with two stacks of tilted donor, D, and acceptor, A, molecules in (g). The TTF–TCNQ complex, discovered in 1972, is metallic from room temperature down to 54 K, and with this material, the possibility of preparing organic metals was demonstrated for the first time.

In order to achieve high conductivity in the stack direction, it appears to be essential that donors and acceptors form separate stacks; with mixed stacks, electron transfer still occurs locally but there is no

delocalised π-electron system to support long-range conduction. It is also necessary that electron transfer between donors and acceptors should be only partial as this then gives rise to partially occupied delocalised π-electron systems.

Electronically conducting systems also form in which only one of the stacks has delocalised π electrons, e.g. TCNQ complexes with alkali metals, such as $(TCNQ)_3Cs_2$, and TTF complexes with halogens, such as $(TTF)_2Br$.

A further dramatic advance in the field of organic metals was the discovery of superconductivity in 1981 by Bechgard and colleagues in TMTSF, tetramethyltetraselenafulvalene salts, first in the hexafluorophosphate salt, $T_c = 0.9$ K at 12 kbar, and then in the perchlorate salt, $T_c = 1.4$ K at ambient pressure. The current record for T_c in a superconducting charge-transfer complex is 11.4 K in $(BEDT\text{-}TTF)_2Cu(NCS)_2$.

8.3 Superconductivity

One of the most exciting scientific breakthroughs in recent years has been the discovery in 1986–1987 of superconductivity at relatively high temperatures in the ceramic oxides $La_{2-x}Ba_xCuO_{4-\delta}$ and $YBa_2Cu_3O_7$. Prior to this discovery, superconductivity was confined to very low temperatures, within a few degrees of absolute zero. Now, with $YBa_2Cu_3O_7$ in particular, superconductivity occurs at temperatures ranging from absolute zero to above the boiling point of liquid nitrogen, 77 K. In the space of a few months, superconductivity changed from being what, for most people, was a low-temperature curiosity, out of reach of normal laboratory conditions, to a readily accessible phenomenon with many exciting possibilities for new developments and applications.

Superconductivity was discovered by the Dutch scientists Holst and Onnes at Leiden University in 1911, in mercury at liquid helium temperatures; this happened just 3 years after He was liquefied for the first time, in the same Dutch laboratory. Since then, superconductivity has been found in a large variety of metals, alloys, oxides and organic compounds, but only at low temperatures. Prior to 1986, the highest critical temperature, T_c, for the superconductor-to-metal transition was just over 20 K, in the alloys Nb_3Ge and Nb_3Sn. An enormous leap forward was the discovery, by IBM Zurich scientists Bednorz and Müller, of superconductivity in a La–Sr–Cu–O phase with a T_c of about 36 K; for this, they were awarded the 1987 Nobel Prize in Physics. This discovery was followed rapidly by the synthesis of $YBa_2Cu_3O_7$ with a T_c of 91 K and other complex cuprates of Bi, Tl and Hg, with T_c values up to the current record at ambient pressure of 138 K. As a consequence of these discoveries, and since it is possible to fabricate materials such as $YBa_2Cu_3O_7$ into wires and tapes, which remain superconducting over large distances, a revolution in many areas of the electrical, magnetic and electronics industries is widely anticipated.

8.3.1 The property of zero resistance

Superconductors have several characteristic features, one of which is zero resistance to the flow of electrical current, implying that a 'supercurrent' could flow indefinitely. The electrical resistance of $YBa_2Cu_3O_7$ as a function of temperature is shown in Fig. 8.4. Below 90 K in the superconducting state, the resistance is zero. Above about 92 K, the material is metallic and the resistance gradually rises with increase in temperature. This increase, characteristic of so-called *normal state* metallic behaviour, is associated with *electron–phonon* collisions, equation (8.1). Although we speak of the conduction electrons of metals as being delocalised or free, they do collide with vibrating atoms as well as with crystal defects and other imperfections and this gives rise to the residual, but small, resistance of metals, which increases with temperature.

Prior to the discovery of the *high-T_c ceramic superconductors*, as materials such as $YBa_2Cu_3O_7$ are known, the generally accepted *BCS theory of superconductivity* (due to Bardeen, Cooper and Schrieffer) involved a

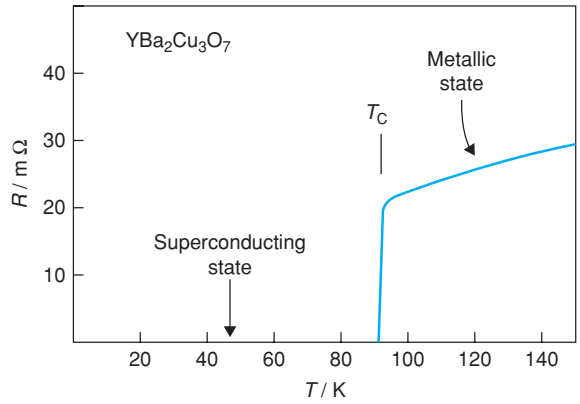


Figure 8.4 ***Electrical resistance of $YBa_2Cu_3O_7$ as a function of temperature.***

loose association of electrons into pairs, known as *Cooper pairs*. These pairs move cooperatively through the lattice in such a way that electron–phonon collisions are avoided. In support of BCS theory, the free energy of the superconducting state is known to be lower than that of the corresponding metallic state in the same material; a low-energy band gap is present and absorption of radiation in the microwave region causes promotion of electrons from the superconducting to the metallic state.

Heat capacity measurements show the superconducting state to have lower entropy and hence be more ordered than the metallic state. The transition from the superconducting to the metallic state is thermodynamically second order, which means that there is no discontinuity in ΔH or ΔS at T_c but there is a discontinuity in the first derivative of ΔH, $\Delta H/\Delta T$, i.e. heat capacity. On cooling, the low-temperature, superconducting state becomes gradually more ordered and a small band gap, typically 1 meV, opens up (note that this is about three orders of magnitude smaller than the band gap in semiconductors). In a second-order phase transition, all of the changes occur over a wide temperature range below T_c and these changes are complete (or non-existent) at the so-called transition temperature; this contrasts with a first-order transition where, ideally, all of the changes occur at T_c. The magnitude of the superconducting band gap, $E_g(T)$ is given by

$$E_g(T) = E_g(0)\left(\frac{T_c - T}{T_c}\right)^{\frac{1}{2}} \tag{8.2}$$

where $E_g(0)$ is the energy gap at 0 K. Thus, $E_g(T) = 0$ at $T = T_c$.

In the band theory of metals, we are familiar with the idea that the conduction band contains an enormous number of discrete levels which are very close, but nevertheless distinct, in energy. The band conduction model cannot be applied to the superconducting state since energy would be lost, or dissipated, as electrons moved between occupied and empty states, whereas once a superconducting current is set up, it flows indefinitely without loss of energy. The supercurrent in a superconductor belongs to a single quantum state, therefore, and is referred to as a quantum phenomenon in which a piece of superconducting material behaves as though it were a giant single atom.

Evidence that electron–phonon interactions are somehow related to Cooper pair formation and superconductivity comes from the *isotope effect*. Comparing the critical temperatures, T_c, of superconducting elements that have different isotopes, the following correlation is observed:

$$T_c \propto M^{-\beta} \tag{8.3}$$

where M is the mass of the isotope and the exponent β has the approximate value 0.5. Fundamental vibrational frequencies, ν, of lattice atoms depend on their reduced mass, M_0, and the force constant, k, between vibrating atoms according to

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{M_0}} \tag{8.4}$$

Equations (8.3) and (8.4) therefore provide a link between lattice phonon frequencies and superconductivity.

In traditional superconductors, the *correlation length* of Cooper pairs is large, hundreds or thousands of ångströms. This is completely different from chemical bond lengths, which involve electron pairs on a scale of 2–3 Å. In the high-T_c ceramic superconductors, BCS theory would require correlation lengths of the order of a few ångströms. The tantalising possibility exists that the Cooper pairs in the superconducting state may be associated with the occurrence of more closely coupled electron–electron interactions and in some way, with bond formation in solids, rather than electron–phonon interactions which are responsible for superconductivity in low-T_c materials. It should be emphasised, however, that current understanding of superconductivity in high-T_c ceramic materials is limited and that much more theoretical work remains to be done.

8.3.2 Perfect diamagnetism; the Meissner effect

Superconducting materials exhibit 'perfect diamagnetism' and expel a magnetic field, provided that the field is below the critical field strength, H_c. This is demonstrated convincingly by the *Meissner effect*, Fig. 8.5(a), discovered by Meissner and Ochsenfeld in 1933. When in the superconducting state, prepared by cooling in, e.g., liquid N_2, a sample, M, of $YBa_2Cu_3O_7$ suspended by a thread swings out from between the jaws of a magnet. As the sample warms, it loses its superconductivity and swings back into the magnetic field. The Meissner effect shows that external magnetic fields are expelled from the interior of a superconducting material; nevertheless, it is possible for the magnetic field to penetrate the surface of the sample to a small extent known as the *London penetration depth*, given by

$$\text{penetration depth, } \lambda = \left(\frac{m}{\mu_0 n e^2}\right)^{\frac{1}{2}} \tag{8.5}$$

where m, n and e are the mass, concentration and charge, respectively, of the conducting species (electrons) and μ_0 is the magnetic moment of the material. The magnetic field, and the associated current density, decrease exponentially with distance from the surface and λ is the distance at which the magnetic field has decreased to ~37% (1/e) of its external strength, typically $\lambda \approx 100$ nm.

For most chemists, diamagnetism is a rather uninteresting weak repulsion effect shown by all materials towards magnetic fields and is a necessary correction factor when making accurate measurements of, for instance, paramagnetic susceptibility. The perfect diamagnetism of superconductors is, by contrast, an enormous effect which could lead to applications in areas such as levitation, Fig. 8.5(b).

8.3.3 Critical temperature T_c, critical field H_c and critical current J_c

The superconductivity in a particular material is lost on either (a) heating above a certain *critical temperature* T_c, (b) increasing the strength of an applied magnetic field beyond a *critical field* H_c or (c) increasing the electrical current passing through it beyond a *critical current density* J_c. These parameters are interlinked; thus, at or just below T_c, the value of H_c is vanishingly small but at lower temperatures, H_c increases. A

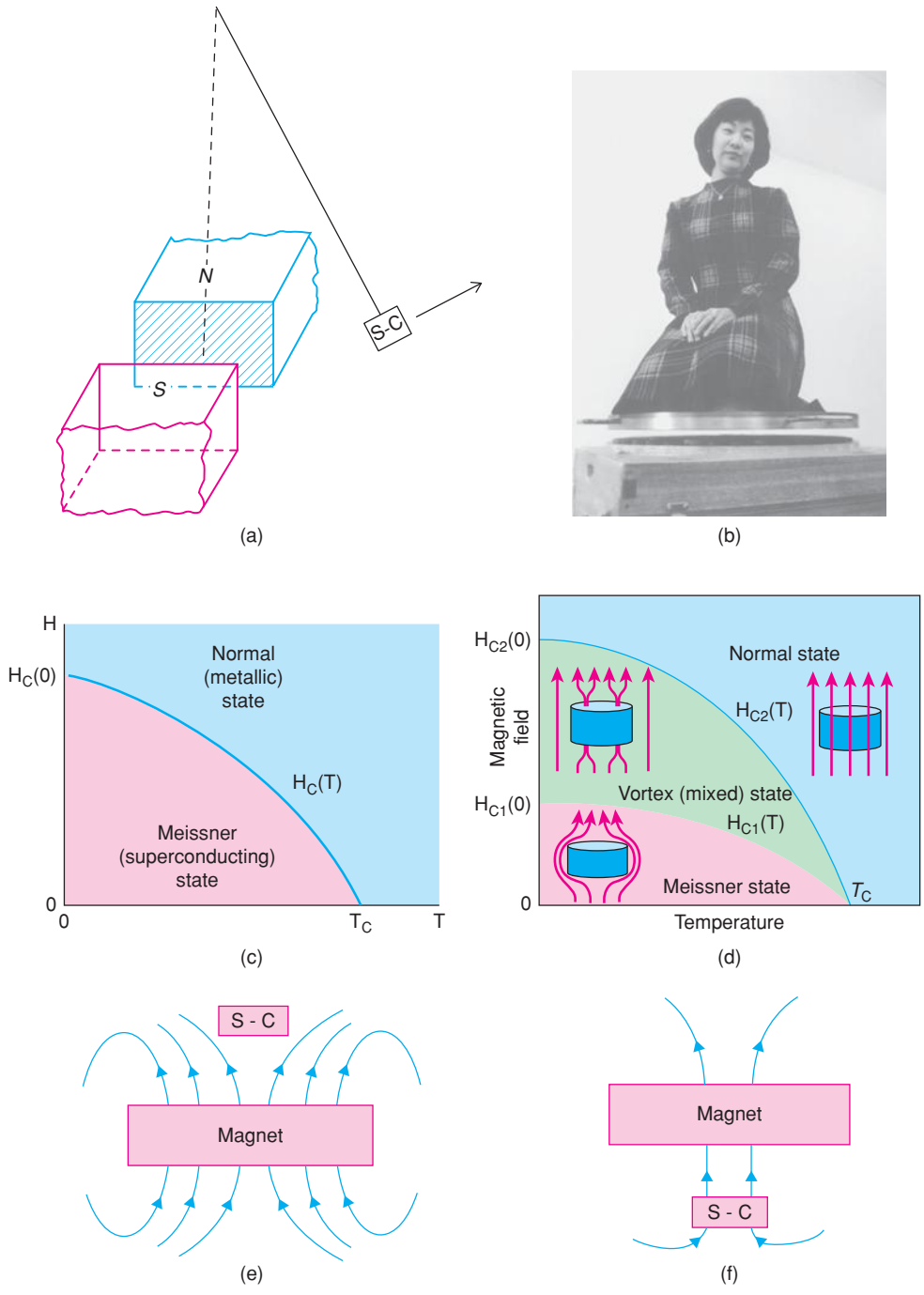


Figure 8.5 *(a) The Meissner effect showing repulsion of a superconductor, S-C, by a magnetic field. (b) Levitation in action: the Japanese lady is kneeling on a magnet that is floating above a bed of superconducting melt-processed $YBa_2Cu_3O_7$. Photograph courtesy of Dr M. Murakami, ISTEC, Japan. (c) A type I and (d) a type II superconductor. Distribution of flux lines in a demonstration of (e) levitation and (f) suspension by vortex pinning.*

schematic H–T phase diagram is shown in Fig. 8.5(c): $H_c(T)$ gradually increases from 0 at T_c to a maximum, $H_c(0)$, at absolute zero and is given by the empirical relationship

$$H_c(T) = H_c(0)\left[1 - \left(\frac{T}{T_c}\right)^2\right] \tag{8.6}$$

The value of T_c is of critical importance for applications that depend on the property of zero resistance, whereas H_c governs whether or not superconductors can withstand applied magnetic fields. J_c is a critical parameter for power applications in which it is desired to pass a high current through superconducting wires or devices. Since the high-T_c superconductors are ceramic materials, rather than conventional metals or alloys, one of the major challenges facing researchers is to produce these in the form of wires or tapes in which there is a continuous current pathway over large distances and in which J_c is large: target current density values are around 10^6 A cm^{-2}.

8.3.4 Type I and type II superconductors: the vortex (mixed) state

The so-called *type I* superconductors have an H–T phase diagram of the kind shown in Fig. 8.5(c) in which, with increasing H or T, an abrupt change from a superconducting to a non-superconducting state occurs. Type I superconductors are modelled well by BSC theory in which the Cooper electron pairs are coupled by lattice vibrations. This correlates with the observation that the best metallic conductors at room temperature, Cu, Ag and Au, are not superconductors, even at low temperatures: they have the smallest lattice vibrations, consistent with their high metallic conductivity, which correlates with the weakness of the electron–phonon interactions required for Cooper pair formation. Most elemental superconductors are type I and are characterised by small H_c values, which means that they lose their superconductivity in modest magnetic fields. In type I superconductors, the Cooper pairs have opposed spins, i.e. they couple antiferromagnetically, with zero angular momentum and spin: they are said to be *s-wave singlet superconductors*; an applied magnetic field decouples the Cooper pairs easily so that the spins align preferentially with the field.

In *type II* superconductors, there is a transition state, the so-called *vortex state*, or *mixed state*, between superconducting and metallic (or 'normal' state) regimes, (d). In this state, magnetic lines of force are able to pass through the material but only in narrow regions called vortices and the lines of force or *flux lines* bunch together through these vortices. A key requirement for certain applications is to pin the vortices so that they cannot undergo lateral displacement; they are then trapped inside the material. You may have seen demonstrations in which a superconducting material is suspended *beneath*, but not touching, a magnet or, alternatively, in which a superconducting pellet is hanging, in mid-air, alongside a magnet. These are possible when the superconductor is type II and the local T, H conditions are such that the superconductor is in the vortex state. A schematic illustration of the principle of levitation with a superconductor in its Meissner state is shown in Fig. 8.5(e) together with the real-life example of levitation in (b). The suspension of a type II superconductor in its vortex state beneath a magnet, by means of trapped lines of force, is shown schematically in (f).

The successful application of vortex state levitation is associated with the mechanisms of *flux pinning*; it appears that crystal defects, compositional inhomogeneities or impurity phases often act as pinning centres. In cases where the vortices are unpinned, they repel each other and may adopt a regular hexagonal array (like a single layer of *cp* spheres) known as a *vortex lattice*. If the vortices are pinned but form a random array, this is known as a *vortex glass*.

Type II superconductors, typified by the alloys Nb_3Sn and Nb_3Ge, have much higher H_c values than type I superconductors; in these, the Cooper pairs of electrons have parallel spins, forming a *triplet state*, and cannot be decoupled readily on application of a magnetic field. Because of their ability to pass high currents without loss of their superconductivity, they have given rise to a new generation of permanent magnets with critical H_c values as high as 16 T, leading to a range of applications in magnetic resonance imaging, NMR and mass spectrometers, beam-steering magnets in particle accelerators and magnetic separation techniques.

8.3.5 Survey of superconducting materials

A selection of some important superconducting materials is given in Table 8.3. The property of zero resistance and superconductivity was discovered in Hg. Many metals are now known to become superconducting just above absolute zero. There is no simple explanation as to which metals become superconductors and which do not, but it is noticeable that several early transition metals do, e.g. La, Ti, V, Mo and Nb, as do a group of late transition metals and heavy p-block metals, e.g. Zn, Hg, Sn and Pb. The pure metal with the highest T_c is Nb, 9.5 K.

A wide range of alloys or intermetallics superconduct; the highest T_c is found in Nb_3Ge, 23.2 K, and for many years it was regarded by theoreticians as having T_c close to the theoretically attainable limit. Alloys such as Nb_3Sn and NbTi are well developed technologically and form the basis of the superconducting magnets used in magnetic resonance imaging (MRI) devices, such as the body scanners that are now used widely in hospitals.

Table 8.3 ***Some superconducting materials***

Material	Substance	T_c/K	Substance	T_c/K
Metals	Hg	4	Nb	9.5
	V	5.4	Tc	7.8
	Pb	7.2		
Alloys	Nb_3Ge	23.2	V_3Si	17
	Nb_3Sn	18	V_3Ga	16.5
	Nb_3Al	17.5	NbN	16
Non-cuprate oxides	$LiTi_2O_4$	12	$BaPb_{0.75}Bi_{0.25}O_3$	13
			$Ba_{0.6}K_{0.4}BiO_3$	30
Cuprates	$La_{1.85}Sr_{0.15}CuO_4$	40	$Tl_2Ba_2CuO_6$	80
	$YBa_2Cu_3O_7$	93	$Tl_2Ba_2CaCu_2O_8$	105
	$Bi_2Sr_2Ca_2Cu_3O_{10}$ (BiSCCO)	110	$Tl_2Ba_2Ca_2Cu_3O_{10}$	125
	$HgBa_2Ca_2Cu_3O_{10}$	138[a]	$Tl_2Ba_2Ca_3Cu_4O_{12}$	115
	Sr_2CuO_4	1.5		
Organics	$(ET)_2Cu(NCS_2)$	11.4		
Others	YNi_2B_2C	15.5	MgB_2	39
	Fullerenes, e.g. Cs_2RbC_{60}	33	$LaO_{1-x}F_xFeAs$	28
	Chevrel phases, e.g. $HoMo_6S_8$	0.7–1.8	$SmO_{1-x}F_xFeAs$	55
	Sulfur: at 93 GPa	10	$ErRh_4B_4$	1.0–8.7
	at 160 GPa	17[b]		

[a] Under pressure, T_c of $HgBa_2Ca_2Cu_3O_{10}$ increases to 166 K, only 20 K below the lowest temperature ever recorded in Nature on Earth. 166 K is still (in 2013), the current record for T_c.

[b] Sulfur is normally non-metallic but at very high pressures becomes metallic and then superconducting with the highest T_c for any pure element to date; its T_c increases further under pressure, in contrast to that of Nb, which decreases.

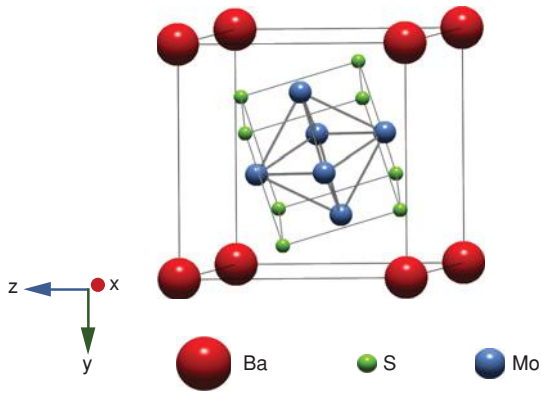
Prior to the high T_c revolution of 1986, superconductivity was known in certain oxides such as reduced $SrTiO_3$, $T_c < 1$ K, but particularly in $LiTi_2O_4$ spinel containing mixed-valence Ti- and Bi-doped $BaPbO_3$ perovskite containing mixed-valence Pb.

A large family of cuprates is now known to superconduct. Almost all are layered structures with mixed-valence Cu (between 2+ and 3+). The main ones of interest for applications are the 123 phases, $YBa_2Cu_3O_7$ and $NdBa_2Cu_3O_7$, the latter because it has an extremely high H_{c2} value of ~7 T which could lead to applications in high-field magnets, and $Bi_2Sr_2Ca_2Cu_3O_{10+\delta}$ (BiSCCO 2223) for fabrication as wires, tapes and eventually, cables, with high J_c values. It is a curious fact that cuprate materials discovered to have still higher T_c values become increasingly exotic, unstable and downright poisonous. Thus, following on from the discovery of the BiSCCO phases, Pb- and Tl-based cuprates were discovered, which pushed T_c up to 126 K and then, most recently, the Hg-based cuprate $HgBa_2Ca_2Cu_3O_{8+\delta}$, which after fluorination treatment has a T_c of 138 K (which increases to 166 K at 23 GPa pressure). The crystallochemical principles of perovskite-related cuprates are discussed in the next section.

Numerous other materials and structure types have been found to superconduct. Particularly interesting ones include the following:

- Sulfide-based Chevrel phases of general formula $REMo_6X_8$:X = S, Se, whose crystal structure is shown in Fig. 8.6(a). These are low-temperature superconductors but analogues such as $HoMo_6S_8$ are *re-entrant superconductors* that superconduct over a limited range of temperatures that does not continue to 0 K. Thus, this material is a normal state metallic conductor above T_{c1}, 1.8 K, is superconducting down to 0.7 K, but below this temperature, T_{c2}, it becomes a normal state ferromagnet. Complex rhodium boride phases of formula XRh_4B_4: X = Y, Th, RE, are also low-temperature superconductors and the X = Er analogue is a re-entrant superconductor with $T_{c1} \approx 8.7$ K and $T_{c2} \approx 1$ K.
- MgB_2, an old material that was discovered by Akimitsu in 2001 to be superconducting with $T_c = 39$ K. It has a very simple crystal structure shown in Fig. 1.51 in which three-coordinate B atoms form graphite-like layers which alternate with layers of Mg atoms. MgB_2 is metallic in its normal state because the B–B bonds are not complete electron pair, covalent σ bonds but are electron-deficient; in the formula $Mg^{\delta+}B_2{}^{\delta-}$ electron transfer to give an ionic structure with $\delta = 2$ is incomplete. MgB_2 shows a strong isotope effect on replacing ^{11}B by ^{10}B, showing the importance of B lattice vibrations, consistent with BCS theory. MgB_2 is of considerable interest for commercial applications since it can be prepared readily as wires and is very resistant to magnetic fields with H_{c2} approaching 200 T at 4.2 K.
- A family of *heavy fermion* compounds typified by $CeCu_2Si_2$ are superconducting; examples include $CeCoIn_5$ ($T_c = 23$ K) and radioactive $PuCoGa_5$ ($T_c = 18$ K). The characteristic feature of these compounds is that their 4*f*(Ce) and 5*f*(Pu) electrons whose energies are close to the Fermi level, E_F, have an *effective mass*, m^*, that is higher than normal by a factor of 10^3. The high density of states, $N(E)$ at E_F causes electrons to move more slowly due to their high concentration, giving rise to the high m^* values. The high $N(E)$ also leads to magnetic instability, causing the energy band to split into spin-up and spin-down sub-bands.
- Single-walled carbon nanotubes are 1D superconductors ($T_c = 15$ K); fullerides are superconductors (see later) and, as indicated above, there is much interest in organic charge-transfer salts.
- Oxypnictide phases such as $REO_{1-x}F_xFeAs$ ($T_c = 28$ K for RE = La, rising to 55 K for RE = Sm, discovered by Hosono in 2008) are of much current interest as they have the second highest T_c values after the cuprates. Many are Fe based, which demonstrates the possibility of links between magnetic phenomena, often associated with Fe-containing materials and superconductivity: it was thought previously that magnetism and superconductivity are incompatible phenomena. Many of these phases were first synthesised and characterised by Jeitschko; one example is shown in Fig. 8.6(b). These structures are derived from that of

Chevrel Phase ($BaMo_6S_8$)



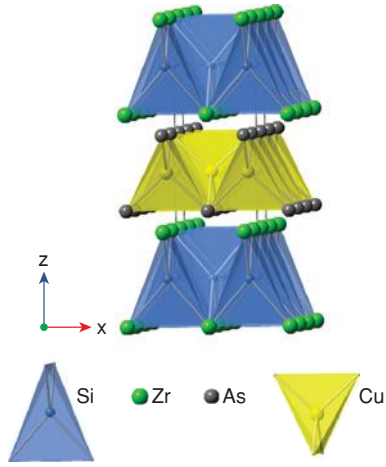
Atom	Wyckoff positions	x	y	z
Ba	1a	0	0	0
Mo	6f	0.23479	0.56589	0.41741
S1	6f	0.1233	0.3907	0.7365
S2	2c	0.2512	0.2512	0.2512

Trigonal (Rhombohedral) : a = 6.6441 Å, $\alpha = 88.562$

Space Group : R $\bar{3}$ (No. 148)

(a)

ZrCuSiAs



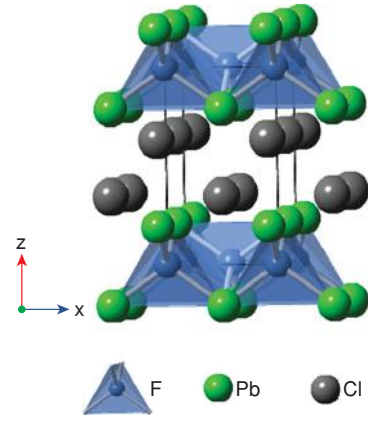
Atom	Wyckoff positions	x	y	z
Zr	2c	0	0.5	0.2046(6)
Cu	2b	0	0.5	0.5
Si	2a	0	0	0
As	2c	0	0	0.6793(8)

Tetragonal : a = 3.6736, c = 9.5712 Å

Space Group : P4 / nmm (No. 129)

(b)

Matlockite (PbFCl)



Atom	Wyckoff positions	x	y	z
Pb	2c	0	0.5	0.2055(1)
Cl	2c	0	0.5	0.6485(1)
F	2a	0	0	0

Tetragonal : a = 4.1062, c = 7.2264 Å

Space Group : P4 / nmm (No. 129)

(c)

Figure 8.6 ***Crystal structures of (a) Chevrel phase, (b) ZrCuSiAs and (c) PbFCl.***

PbFCl, the mineral matlockite, Fig. 8.6(c). These phases typically contain 2D layers of $FeAs_4$ or $CuAs_2$ tetrahedra separated by oxide or metal layers.

Although all of these are fascinating new materials, they are at the early stages of development into practical applications; most interest for applications is focused on the high-T_c cuprates together with high-field magnet applications of more traditional superconductors such as NbTi and Nb_3Sn.

8.3.6 Crystal chemistry of cuprate perovskites

Simple cuprates with the ideal, cubic perovskite structure appear not to exist, probably because Cu, which would occupy the smaller, octahedral B sites, does not usually adopt regular octahedral coordination; instead, depending on its valence state, it is found in a variety of linear, square planar, pyramidal and distorted octahedral coordinations. Much of the crystallochemical complexity, and beauty, of superconducting cuprate structures occurs because the structures are oxygen deficient to varying degrees and this permits Cu to have a range of coordinations and valence states.

We may see how these arise, beginning with the ideal cubic perovskite structure, Fig. 8.7 (a), using the unit cell setting with the small B cation at the corners and the larger 12-coordinate A cation at the body centre, with the oxygens in edge-centre positions. The regular octahedral coordination of one B cation is shown; as discussed in Section 1.17.7, these octahedra link up by corner sharing to form a 3D framework with cavities for the A cations. These cavities are empty in the ReO_3 structure, partially occupied in tungsten bronzes and fully occupied in perovskites and in the perovskite-related cuprate (and bismuthate) superconductors. The B cation sites are usually fully occupied in perovskites (but an example of B site vacancies is found in La-doped $BaTiO_3$: La^{3+} substitutes for Ba^{2+} on the A sites and the observed charge compensation mechanism involves creation of Ti vacancies: $4Ba^{2+} + Ti^{4+} = 4La^{3+}$, with general formula $Ba_{1-4x}La_{4x}Ti_{1-x}O_3$).

Consider now the consequences of removing some of the oxygens from the perovskite structure in Fig. 8.7(a). In (b), all oxygens at positions $(00\frac{1}{2})$ are absent. As a result, the coordination of A is reduced to eightfold; the coordination of B is square planar and the square planar units link at their corners (oxygens) to form infinite 2D sheets. In some cuprate perovskite-related phases, square planar coordination is found for Cu; to chemists, this indicates that the Cu is probably a d^8 ion (square planar coordination is common with d^8 ions, Ni^{2+}, Pd^{2+}, Pt^{2+}), with the d electron configuration shown in (e) and, therefore, Cu is probably in the 3+ valence state.

Next, let us remove the $(00\frac{1}{2})$ oxygens, but only from alternate layers, Fig. 8.7(c). To represent this, a double-perovskite unit cell is required, of ideal dimensions, a, a, $2a$, i.e. with the same, a, b axes, but with c doubled. In (c) the new unit cell is shown, with oxygens at $(00\frac{3}{4})$ present and those at $(00\frac{1}{4})$absent. The coordination of A is now either 12-coordinate (A at $\frac{1}{2}\frac{1}{2}\frac{3}{4}$) or eight-coordinate (A′ at $\frac{1}{2}\frac{1}{2}\frac{1}{4}$). The coordination of B is square pyramidal (i.e. five-coordinate). The (BO_5) pyramids link up to form infinite sheets, and the sheets link up through common apices to form pairs or sandwiches in which one layer of pyramids points 'up' while the other layer points 'down'. The A atoms lie inside these sandwiches and the A′ atoms form layers between adjacent sandwiches. Such pyramidal coordination for Cu is common in cuprates and is taken to be a characteristic of Cu in its common d^9, Cu^{2+} valence state [d^9 ions often show pyramidal or distorted octahedral coordination associated with the Jahn–Teller effect; the d configuration is similar to that in (e) but with an additional electron in a singly occupied $d_{x^2-y^2}$ orbital (f)].

Finally, let us remove all oxygens from the $(\frac{1}{2}00)$, $(0\frac{1}{2}0)$ positions and instead leave those at $(00\frac{1}{2})$. The original unit cell is retained, Fig. 8.7(d); the coordination of A is reduced to four, square planar. The coordination of B is linear; if all unit cells were the same, then infinite (BO) chains would occur, parallel to c. Linear (CuO) segments such as this occur in some cuprates (but infinite chains are not common); such two-coordination for Cu is evidence for a d^{10} state and therefore for Cu^+ species. The rationalisation usually

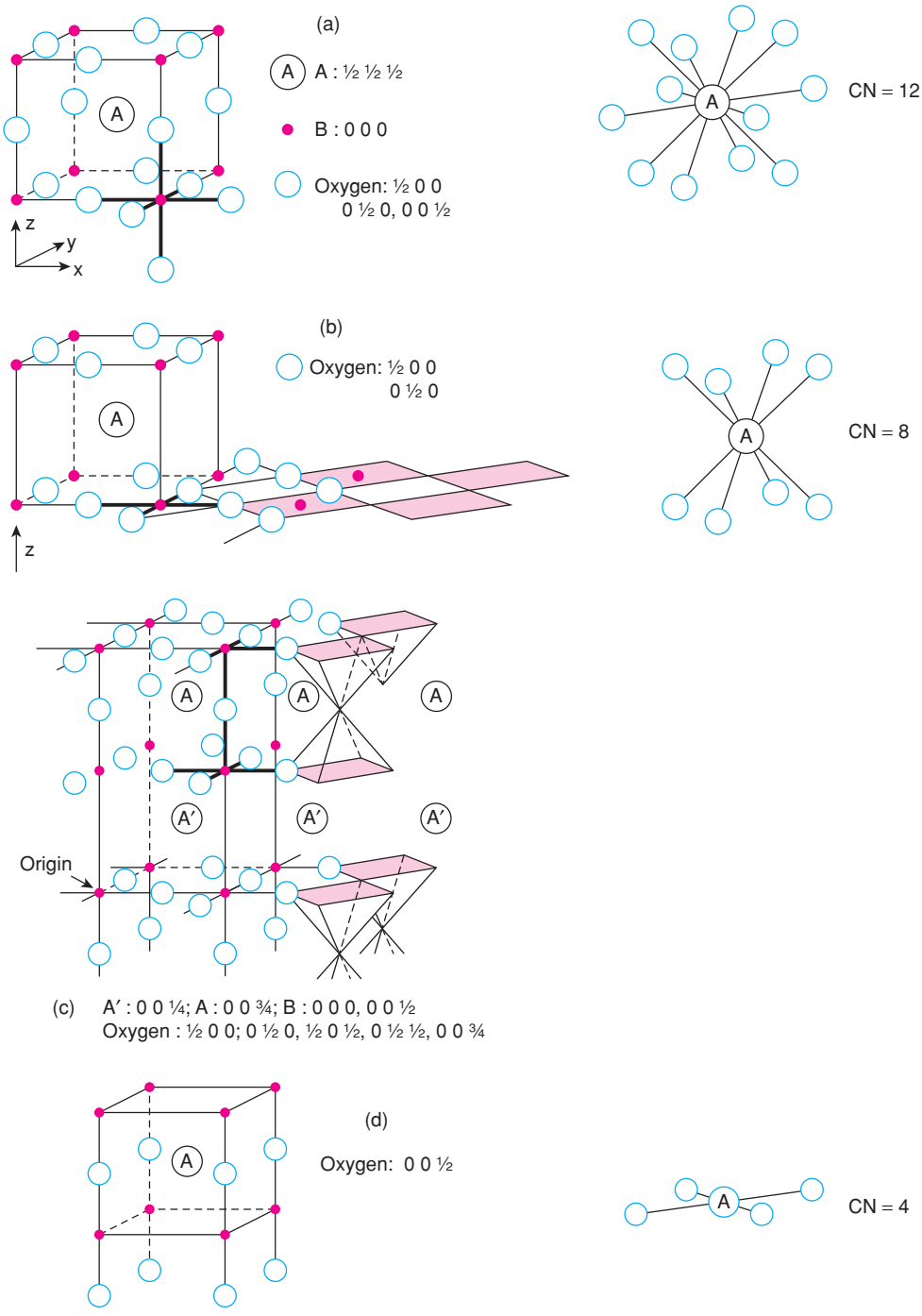


Figure 8.7 ***Perovskite-related cuprate structures showing (a) octahedral, (b) square planar, (c) pyramidal and (d) linear coordination for Cu and the rationalisation of (b)–(d) in terms of (e) Cu^{3+}, d^8, (f) Cu^{2+}, d^9, and (g) Cu(I), $d^{10}s^1$, respectively. In (g), the 4s electron is not ionised (to give Cu^+) but is assumed to take part in covalent bond formation.*** (***Continued***)

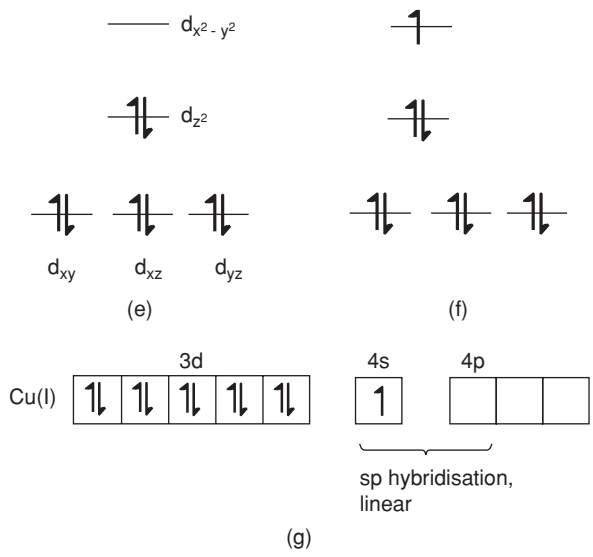


Figure 8.7 **(*Continued*)**

given is that the 3*d* shell on Cu is full (d^{10}) and bonding electrons are associated with the 4*s* and one 4*p* orbital of Cu, giving a linear *sp* hybridised state (g).

Using the crystallochemical principles set out here, the structures of many complex cuprates can be broken down into simple, perovskite-related components, as shown next for $YBa_2Cu_3O_7$. Others are more complex and involve intergrowths of perovskite-related segments of various kinds, together with rock salt-like or fluorite-like segments.

8.3.7 $YBa_2Cu_3O_{7-\delta}$, YBCO

$YBa_2Cu_3O_{7-\delta}$ is a solid solution phase characterised by an oxygen content that varies over the range $0 \leq \delta \leq 1$. As the oxygen content, δ, varies, the average Cu oxidation state varies and this is responsible for the dramatic variation in electrical properties with changing δ. When $\delta = 0$, the solid solution end member $YBa_2Cu_3O_7$ is a superconductor with $T_c = 90$ K. With increasing δ, T_c decreases increasingly rapidly and composition $\delta = 1$ is a semiconductor with no evidence of superconductivity at any temperature. Control of oxygen content is critical to optimising the superconducting properties, therefore, and features prominently in the methods used to synthesise and fabricate YBCO ceramics, films and wires, and also the development of experimental techniques to measure δ in a given sample.

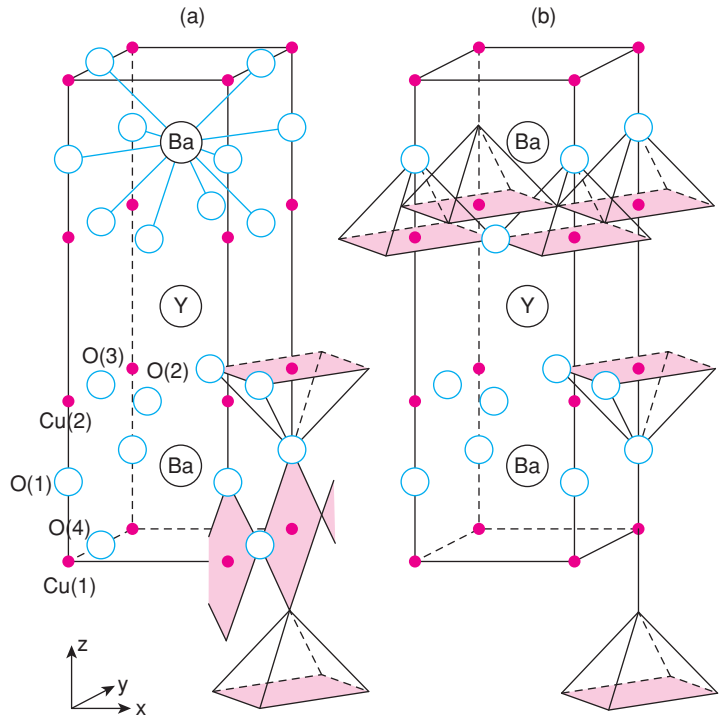
8.3.7.1 Crystal structure

The crystal structure of $YBa_2Cu_3O_7$ is fascinating. Superficially, it is a perovskite with two-ninths of the oxygens missing (compare the stoichiometry of $BaTiO_3$ perovskite with that of $YBa_2Cu_3O_7$). The unit cell is shown in Fig. 8.8 (a) and has three perovskite-related units stacked in the *c* direction. Two of these, comprising the top and bottom thirds of the unit cell, have the net stoichiometry '$BaCuO_{2.5}$' with Ba in A-site positions and have two of the twelve edge centre O sites vacant. This gives Ba a coordination number of ten, as shown

for the upper Ba in (a). The middle unit has the effective stoichiometry '$YCuO_2$' and is characterised by four missing edge centre oxygens. Consequently, Y has a coordination number of eight.

Two kinds of Cu site are present. In perovskite, these sites would be octahedral but, because of the oxygen deficiency, one is five-coordinate [Cu(2), square pyramidal arrangement] and the other is four-coordinate [Cu(1), square planar]. The square planar units link at their corners to form chains parallel to *y*. The square pyramidal units link up to form sheets in the *xy* plane. Pairs of adjacent sheets are linked, via the square planar chains, to form symmetrical complex triple layers. The repeat unit of this complex layer is indicated in the bottom right corner, (a).

The unit cell of $YBa_2Cu_3O_7$ is orthorhombic with $a \approx b \approx c/3$. From the atomic coordinates listed in Fig. 8.8(a), bond distances can be calculated. Thus, Cu(1)–O(1) $= 0.16c = 1.87$ Å, whereas Cu(2)–O(1) $= 0.19c = 2.22$ Å. Cu(1)– O(1) is shorter than Cu(2)–O(1), consistent with the higher formal charge on Cu(1) and hence stronger, shorter bonds to oxygen (see the next section).



a = 3.817, b = 3.882, c = 11.671 Å
Ba : ½ ½ 0.18; ½ ½ 0.82
Y : ½ ½ ½
Cu(1) : 0 0 0
Cu(2) : 0 0 0.35; 0 0 0.65
O(1) : 0 0 0.16; 0 0 0.84
O(2) : ½ 0 0.38; ½ 0 0.62
O(3) : 0 ½ 0.38; 0 ½ 0.62
O(4) : 0 ½ 0

a = b = 3.8, c = 11.8 Å
fractional coordinates :
as for (a), but with
O(4) unoccupied

Figure 8.8 ***Crystal structure of (a) $YBa_2Cu_3O_7$ and (b) $YBa_2Cu_3O_6$.***

8.3.7.2 Atom valences and the superconducting mechanism

The oxidation state of Cu in $YBa_2Cu_3O_7$ is unusual. If we assume that Y, Ba and O have their usual oxidation (valence) states of +3, +2 and −2, respectively, then for charge balance, Cu must be an average of +2.33. This may be rationalised in terms of a +2 state for square pyramidal Cu(2), of which there are two in the unit cell, and a +3 state for square planar Cu(1), of which there is one per cell; these nominal charge states are consistent with the observed coordination environments, as discussed in Section 8.3.6.

There is, however, considerable debate as to the most appropriate way to describe the valence states because some believe that instead of having Cu^{3+} ions as well as Cu^{2+} ions, some of the oxide ions are present as O^- ions. Either way, one of the species, either Cu or O, is partially oxidised compared with the more usual states of Cu^{2+} and O^{2-}. An alternative view is to regard the basal plane of the square pyramids as units, i.e. as $(CuO_2)_\infty$, which are partially oxidised, but leaving open the question as to whether Cu, O or the band structure comprising overlapping Cu 3*d* and O 2*p* orbitals is oxidised. Consistent with this view, $YBa_2Cu_3O_7$ is metallic in its normal state, with delocalised electrons in the *xy* plane: the conductivity is anisotropic and is much higher in the *xy* plane than parallel to *z*, consistent with the 2D structure of the cuprate layers.

It is likely that the more oxidised part of the $YBa_2Cu_3O_7$ structure involves the Cu–O band structure but also that electron charge transfer can occur between the Cu(2) pyramids and the Cu(1) square planar units. This is because, although the superconducting mechanism appears to require positive holes and, therefore, partial oxidation of Cu to give mixed 2+, 3+ states, those Cu atoms which appear to be most oxidised are Cu(1), in the chains, whereas the superconductivity is thought to involve primarily Cu(2) in the square pyramids, which is nominally Cu^{2+}. For this reason, Cu(1), which is nominally Cu^{3+}, is regarded as a 'sink' which can accept electrons from the $Cu(2)O_2$ basal planes in an internal redox or charge-transfer process.

The key to the superconducting behaviour in $YBa_2Cu_3O_7$ certainly appears to be the Cu atoms, their oxidation states and the manner in which they link up with O to form the complex structure shown in Fig. 8.8(a). The superconductivity may be destroyed by partial reduction of the structure. Oxygens are removed, specifically from the O(4) sites (a), and consequently, the average oxidation state of Cu is reduced. The dependence of T_c on δ is shown in Fig. 8.9 (a). Two types of behaviour are seen, depending on whether samples are cooled slowly or rapidly; these differences are not fully understood but may be related to variable amounts of internal charge transfer of electrons between Cu(1) chains and Cu(2) planes.

The structural consequences of removing oxygen are seen in Fig. 8.8(b). At $\delta = 1$, i.e. $YBa_2Cu_3O_6$, all O(4) oxygen sites are empty and the coordination of Cu(1) is reduced to 2 (linear). As discussed in Section 8.3.6, we take this to be a characteristic coordination for Cu^+, which would be consistent with the oxidation states for $\delta = 1$ of $Y^{3+}Ba_2^{2+}Cu(1)^+Cu(2)_2^{2+}O_6$. Hence, on removal of oxygen, the Cu(2) planes are largely unaffected and Cu(2) remains primarily as Cu^{2+} whereas Cu(1) is reduced from Cu^{3+} to Cu^+.

The unit cell of $YBa_2Cu_3O_6$ is tetragonal, in contrast to the orthorhombic unit cell for $\delta = 7$; thus, parallel to *z*, the unit cell (b) has a fourfold rotation axis passing through the A-site cations and the top/bottom face centres. In $YBa_2Cu_3O_7$ (a), this axis is only twofold, however, because of the presence of O(4) oxygens at $0\frac{1}{2}0$ but the absence of similar oxygens at $\frac{1}{2}00$.

8.3.7.3 Oxygen content of $YBa_2Cu_3O_{7-\delta}$

Most oxides have accurately fixed oxygen contents; $YBa_2Cu_3O_{7-\delta}$ is a remarkable exception since δ may be varied easily over the range $0 \leq \delta \leq 1$ simply by heating the material at different temperatures and at different oxygen partial pressures. This is shown by thermogravimetry in Fig. 8.9(b), where the mass of $YBa_2Cu_3O_{7-\delta}$ is given as a function of temperature in air. Starting with fully oxidised $YBa_2Cu_3O_7$, oxygen loss commences above ∼400 °C and continues smoothly to ∼1000 °C, which is close to the melting temperature (which

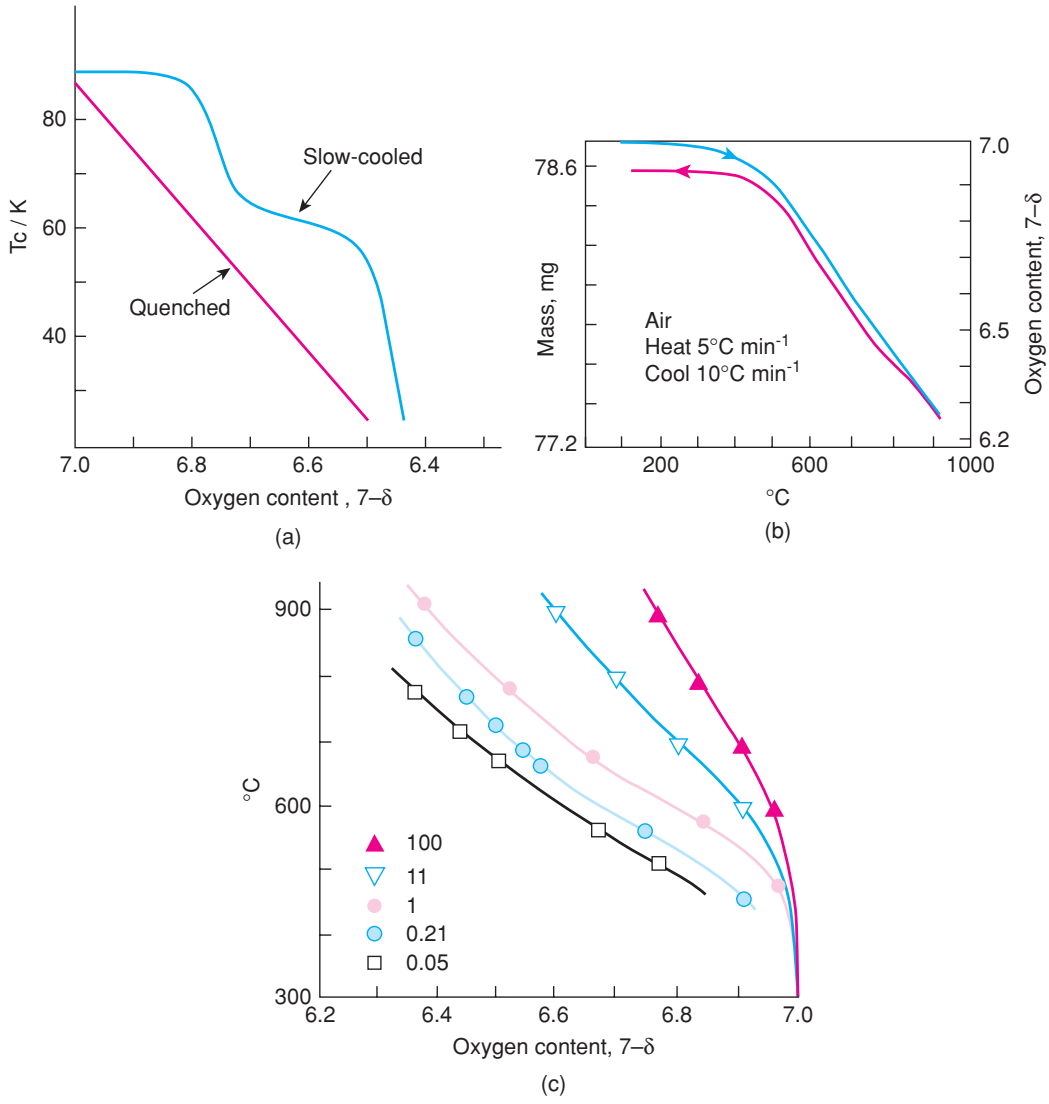


Figure 8.9 *(a)* T_c *versus oxygen content* δ *for* $YBa_2Cu_3O_\delta$ *showing the importance, also, of sample history (i.e. cooling rate after synthesis). (b) Thermogravimetric analysis on* $YBa_2Cu_3O_7$ *showing oxygen loss on heating which is largely reversible on cooling; cooling rates of much less than 10 °C* min^{-1} *would be required to regain fully an oxygen content of 7.0. (c) Phase diagram for* $YBa_2Cu_3O_\delta$ *showing* δ *as a function of temperature for different oxygen pressures (bar* O_2*). (d)* T_c *as a function of cubic unit cell parameter,* a*, for the fullerides: (▲)* K_3C_{60} *measured under different pressures, (▽)* Rb_3C_{60} *at different pressures and (□)* $(M_xM'_{3-x})C_{60}$*: M, M′ = K, Rb, Cs, with different compositions, x. Reproduced with permission from Gunnarsson,* Rev. Mod. Phys.*, 69, 575–606, © 1997 American Physical Society. (e) Schematic decrease in conduction band width and increase in* $N(E)$ *at* E_F *as the* C_{60}^{3-} *ions move further apart.* (*Continued*)

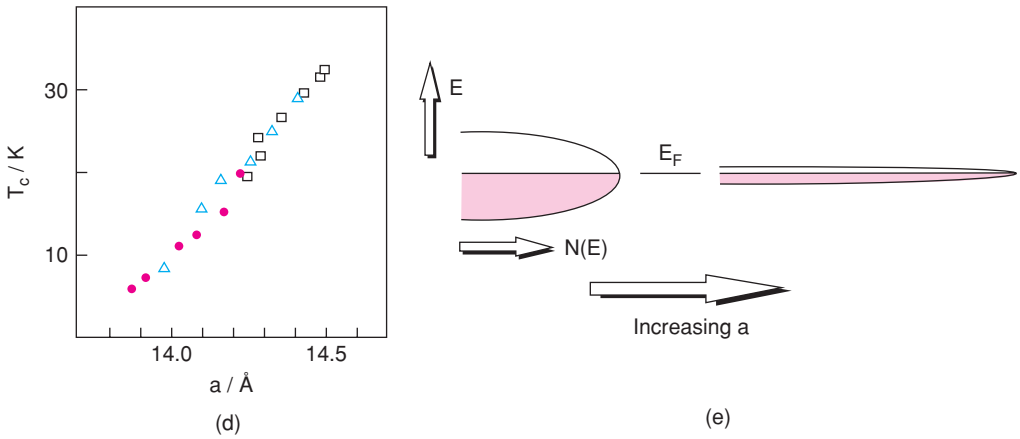


Figure 8.9 **(*Continued*)**

also varies with δ!); on cooling, oxygen uptake occurs rapidly and equilibration is achieved reversibly at the relatively rapid heat/cool rates used, 10 °C min^{-1}. To prepare samples, at room temperature, with intermediate δ values, samples may be quenched after equilibration at a certain temperature; with rapid cooling, there is no time for oxygen uptake to occur.

The phase diagram for $YBa_2Cu_3O_{7-\delta}$ showing temperature versus δ for different atmospheres is given in Fig. 8.9(c); at a given temperature, δ increases as P_{O_2} decreases; alternatively, for a given P_{O_2}, δ increases with increase in temperature.

8.3.7.4 Determination of oxygen content, 7–δ

Determination of the oxygen content of a solid is not easy. Chemical methods involving dissolution followed by subsequent elemental analysis using, e.g., atomic absorption methods are inappropriate since the dissolution media usually have high oxygen contents themselves. Direct elemental analysis of a solid is possible using electron microscopy (EPMA, EDAX) and X-ray fluorescence but, since oxygen is a light element, the X-rays emitted are 'soft' (i.e. of relatively low energy), and are readily absorbed by the sample from which they are emitted. Quantitative analysis of oxygen is possible with certain instruments but the methods are not routine.

A generally-applicable, indirect method is to heat samples in a reducing atmosphere and, provided the products are known, determine the oxygen content from the weight change on reduction. In *hydrogen reduction TG*, samples are heated in an atmosphere of, e.g., 5% H_2–95% N_2 on a programmed heating cycle until a constant mass is obtained. For $YBa_2Cu_3O_{7-\delta}$, the reaction is

$$YBa_2Cu_3O_{7-\delta} + (3.5 - \delta)H_2 \rightarrow \tfrac{1}{2}Y_2O_3(s) + 2BaO(s) + 3Cu(s) + (3.5 - \delta)H_2O(g)$$

Hence the Cu component is reduced to Cu metal but, under these conditions, Y and Ba stay as their simple oxides.

A second method, applicable to substances containing mixed-valence elements, is iodometric titrations or *iodometry*. The substance is dissolved in, e.g., dilute HCl, to which an excess of an acidified solution of KI is added. The I^- ions play two roles, as a reducing agent (Cu^{3+}, Cu^{2+} to Cu^+; O^- to O^{2-}) and as a precipitating agent (CuI is insoluble). The amount of I^- used in these various reactions is obtained from the

amount of unused I^-, which is obtained by back-titration against thiosulfate. An independent measurement of the total Cu content is required; this may be achieved by dissolving the sample in dilute HCl containing Fe^{2+}. Any Cu^{3+} present reduces to Cu^{2+} by oxidising Fe^{2+} to Fe^{3+}. An iodometric analysis is then performed. Comparison of results obtained with/without prior reduction using Fe^{2+} leads to a determination of the Cu^{3+} content, and therefore of the oxygen content δ.

8.3.8 Fullerides

Many of the alkali metal fullerides, such as K_3C_{60}, are superconducting. They are 3D superconductors, in contrast to the cuprates, which are 2D superconductors. The structures and bonding in C_{60} and the fullerides are discussed in Sections 1.15.6 and 3.4.4.3. On reaction of C_{60} with alkali metals, electrons enter the conduction band of C_{60} and convert it from an insulator to a metal. The conductivity is optimised at the composition A_3C_{60} at which the conduction band is half full, for the C_{60}^{3-} anion, and then decreases again to a minimum at C_{60}^{6-} at which the conduction band is full.

Many A_3C_{60} compounds are superconducting; for some, the pressure dependence of T_c has been determined and a good correlation is found between T_c and the cubic unit cell parameter, a, Fig. 8.9(d). This gives us important information into the conditions under which superconductivity occurs. First it shows that the alkali metal cations, M^+, are merely 'packing' ions and have little influence on the electrical properties once they have donated their electrons to the C_{60} network. Second, T_c rises as the C_{60}^{3-} ions are pushed apart, leading to a reduction in orbital overlap between adjacent C_{60}^{3-} ions and hence to a reduction in 'metallicity' or conduction band width. A curious fact about superconductors, especially the high-T_c cuprates, is that superconductivity is found not in very good metals, but in very bad ones that are on the edge of a metal–insulator transition. This phenomenon is illustrated very well with the fullerides, Fig. 8.9(d). As the C_{60}^{3-} ions are pushed apart, the conduction band must become increasingly narrow; however, the number of energy levels is unchanged (three for each C_{60}) and hence the density of states, $N(E)$, becomes very large near the Fermi level, i.e. at E_F there is a very large number of mobile electrons with very similar energies, Fig. 8.9(e). At this point, theories of the origin of superconductivity diverge somewhat and there is, as yet, no widely-accepted explanation of the origin of superconductivity in the cuprates or in the fullerides.

8.3.9 Applications of superconductors

The main commercial application of superconductors at present is in superconducting magnets for magnetic resonance imaging, MRI, and NMR spectrometers with a $4bn global market. These use conventional superconductors, such as NbTi ($T_c = 9.2$ K), and are made from coils of NbTi wire, which produce higher magnetic fields than are obtained readily using standard electromagnets; they are not expensive to operate since there is no heat loss in the windings during current flow but they do need to be cooled in liquid He. NbTi is the material of choice at present as it is well studied and easy to wind; there is interest in the new MgB_2 superconductor ($T_c = 39$ K) as it is cheaper, whereas YBCO, which would be entirely satisfactory, is costly to produce as wires.

Ever since the discovery of YBCO and BiSSCO 2223, which superconduct at liquid N_2 temperatures, there has been much interest in the fabrication of wires and tapes of these materials for possible use as current-carrying applications and as electrical circuit components and interconnects in microelectronic devices. However, YBCO and BiSSCO are ceramic oxide phases, and cannot readily be prepared as wires, which is obviously straightforward for metals and alloys. There are two main routes to circumvent this problem. One is to pack Ag tubes with powder of superconducting material and then draw the tubes into wires; in this way, BiSSCO wires with a current density, J, of 10^4 A cm^{-2} at 77 K have been produced and formed the first generation of commercial wires that required cooling to only 77 K to become superconducting. The second

method is to deposit the powders on a metal substrate such as Ni or stainless steel. This has been achieved satisfactorily for YBCO, giving wires that are hundreds of metres long and with current densities as high as 10^6 A cm^{-2} at 77 K. This current density is approximately 100 times that of Cu metal, as used in electrical circuits, but such YBCO wires are expensive to produce. They do, nevertheless, possess all the attributes required, including the ability to resist any magnetic field strength (H_{c_2} is too high to measure at 4 K), but it seems that, at present at least, we shall not see the installation of YBCO current-carrying cables in national electricity grids, because of the high costs of fabrication and operation.

In the early days of high-T_c discovery and materials development, it was anticipated that microelectronics applications would be forthcoming, but as yet these have not materialised, mainly because of the development of MOSFETS, which allow increased miniaturisation of microelectronics technology without the need for cooling to liquid N_2 temperatures.

The demonstrations of levitation using the new type II high-T_c superconductors are very impressive but as yet have not seen commercialisation in, for instance, maglev trains. Apart from the Yamanashi test line in Japan, most maglev systems use conventional, Fe-core electromagnets since there is little advantage in using superconducting magnets.

A novel magnetic property-based area which is a direct application of superconductivity is in the development of SQUIDs (Superconducting QUantum Interference Devices) for the detection of minute magnetic fields. SQUIDs were developed following the discovery of the *Josephson effect* in which a current can tunnel between two superconducting materials separated by a thin insulating barrier. SQUIDs use Josephson junctions in their construction and are able to detect very small magnetic fields. They have medical applications in magnetocardiography, to detect cardiac anomalies, and in magnetoencephalography, to study anomalies in the human brain. SQUIDs also have scientific applications in research laboratories for characterising the magnetic behaviour of materials.

Large superconducting magnets are vital to the construction of synchrotron storage rings and particle accelerators in which electrons are accelerated to velocities approaching the speed of light around a circular path. The electrons are confined to a curved trajectory by superconducting bending magnets. For instance, the Large Hadron Collider at CERN, Switzerland, uses bending magnets that require a current of 8000 A, yielding a magnetic field of 8 T at 2 K. One application area of such particle accelerators is to increase the energy of collisions between electrons and other particles; to date, impact energies of up to 7 TeV at CERN have been achieved and have provided scientific results that underpin the recent announcement of the discovery of the Higgs Boson. A second application area arising from electrons constrained to curved pathways is that they emit radiation tangentially. Such synchrotron radiation has many applications in basic scientific research, for spectroscopic studies and diffraction measurements; some of these techniques are described in Chapter 6.

8.4 Semiconductivity

Let us first clarify some important differences between metals and semiconductors. In general, conductivity, σ, is given by

$$\sigma = ne\mu \tag{8.7}$$

where n is the number of current carriers, e their charge and μ their mobility. In metals, the number of mobile electrons is essentially constant, but their mobility gradually decreases with increase in temperature due to electron–phonon collisions. Consequently, conductivity gradually decreases with rising temperature, as shown schematically in Fig. 8.10, which is an Arrhenius plot of log conductivity against reciprocal temperature.

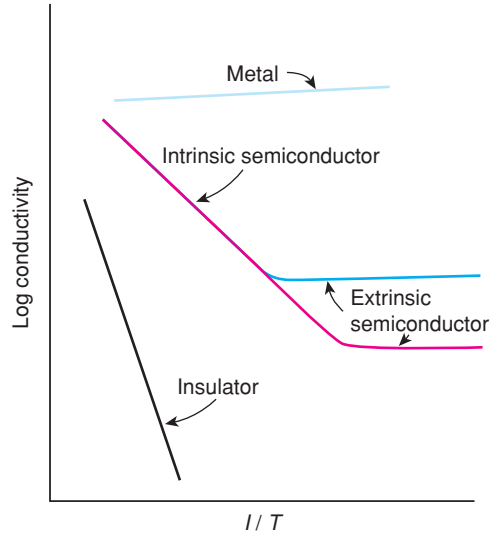


Figure 8.10 ***Conductivity of metals, semiconductors and insulators.***

In semiconductors, the number of mobile electrons is small. This number may be increased, either by raising the temperature to promote more electrons from the valence band to the conduction band, or by doping with impurities that provide either electrons or holes. In the first of these, n is given by

$$n = n_0 \exp(-E/kT) \tag{8.8}$$

where n_0 is a constant (the total number of electrons), E is the activation energy for promotion of electrons to the conduction band, k is Boltzmann's constant (chemists may prefer to use the gas constant, R, if the units of E are kJ mol^{-1}; if k is used, the units of E are in eV) and T is temperature (K). The number of mobile electrons n and, therefore, σ rises exponentially with increase in temperature, as shown for the *intrinsic* region in Fig. 8.10. The mobility, μ, changes little with temperature, as is the case for metals, and therefore the temperature dependence of the conductivity is essentially determined by the temperature dependence, if any, of n.

In the second case, extra mobile carriers are generated by the addition of dopants; at low temperatures in the *extrinsic* region, Fig. 8.10, the concentration of extra carriers is much greater than the thermally generated intrinsic concentration. Consequently, in the extrinsic region, carrier concentration is independent of temperature and σ shows a slight decrease with increase in temperature due to the mobility effect mentioned above.

Insulators differ from semiconductors only in the magnitude of their conductivity, which is also sensitive to both temperature and dopants. Since the band gap of insulators is much larger than that of semiconductors, the activation energy for carrier creation is much greater and, therefore, σ is very small, even at higher temperatures.

The application of band theory to describe the electronic properties of metals, semiconductors and insulators is outlined in Chapter 3. The key parameter is the band gap, E, Fig. 8.11, and Tables 3.18 and 3.19. In order to promote electrons across the band gap, absorption of energy is required. For small band gaps, <1 eV, some promotion due to thermal excitation occurs, especially at high temperature; materials with $E < 0.01$ eV are essentially metallic. For larger band gaps, radiation of appropriate wavelength can cause promotion and *photoconductivity* results. Thus CdS, $E = 2.45$ eV, absorbs visible light and is used in photocells to convert sunlight into other forms of energy.

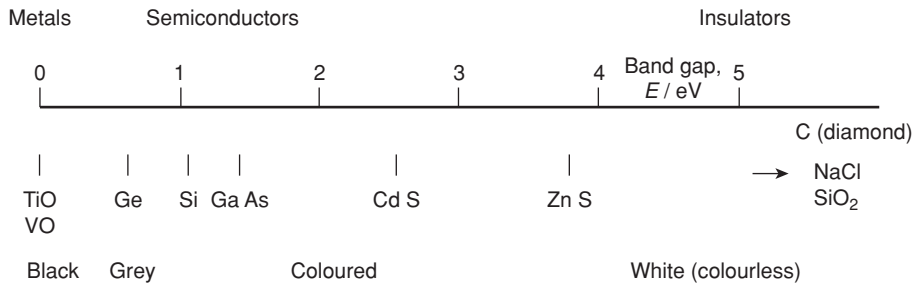


Figure 8.11 ***Relation between electronic properties and magnitude of the band gap with typical examples.***

8.4.1 Elemental and compound semiconductors with diamond and zinc blende structures

The most important semiconductors, technologically, by far are the Group IV elements, C, Si and Ge, with the diamond structure and isostructural compounds consisting of elements from Groups III and V of the Periodic Table such as GaAs or Groups II and VI such as CdTe. These compounds have the same basic atomic arrangement as in diamond but the III–V and II–VI elements are ordered onto alternate tetrahedral sites. This ordered structure, known as the zinc blende or sphalerite structure, is described in Section 1.17.1.2 and shown in Fig. 1.33. Unit cell data for the Group IV elements with the diamond structure are listed in Table 1.9 and some band gap data for elemental and compound semiconductors in Tables 3.18 and 3.19. Of these various semiconducting materials, silicon is by far the most important technologically and most of our modern microelectronics industry is Si based.

Many of the technological applications of semiconductors are associated with doped or extrinsic materials. Si is an extrinsic semiconductor when doped with a Group III or V element. First, let us see the effect of doping with a small amount, say 0.02 at.%, of a trivalent element, e.g. Ga. The Ga atoms replace Si in the tetrahedral sites of the diamond structure to form a substitutional solid solution. In pure Si, using covalent bond theory, all the Si–Si bonds are electron pair single bonds since Si has four valence electrons and is bonded to four other Si atoms, Fig. 8.12(a). Ga has only three valence electrons, however, and in Ga-doped Si, one of the Ga–Si bonds must be deficient by one electron. Using band theory calculations and data from experimental results, it is found that the energy level associated with each single-electron Ga–Si bond does not form part of the valence band of Si. Instead, it occupies a discrete energy level or atomic orbital just above the top of the valence band. This level is known as an *acceptor level*.

The gap between the acceptor levels and the top of the valence band is small, $\leq$0.1 eV. Consequently, electrons from the valence band have sufficient thermal energy to be promoted readily into the acceptor levels. The acceptor levels are discrete if the Ga concentration is low and it is not possible for electrons in the acceptor levels to contribute directly to conduction. When electrons are promoted (thermally) from the valence band to the acceptor levels, the positive holes that are left behind in the valence band are able to move. Ga-doped Si is known as a positive hole or *p-type semiconductor*.

The concept of positive hole conduction is very useful although somewhat misleading; in reality, a positive hole can move only when an electron hops into it, leaving its original site vacant, as a fresh positive hole. In a sequential hopping mechanism, positive holes move in one direction by virtue of electrons hopping in the opposite direction!

At normal temperatures, the number of positive holes created by the presence of Ga dopants far exceeds the number created by thermal promotion of electrons into the conduction band, i.e. the extrinsic, positive hole concentration far exceeds the intrinsic concentration. Hence the conductivity is controlled by the Ga concentration. With increasing temperature, the concentration of intrinsic carriers increases exponentially until, at sufficiently high temperatures, it exceeds the extrinsic value and a changeover to intrinsic behaviour occurs, Fig. 8.10.

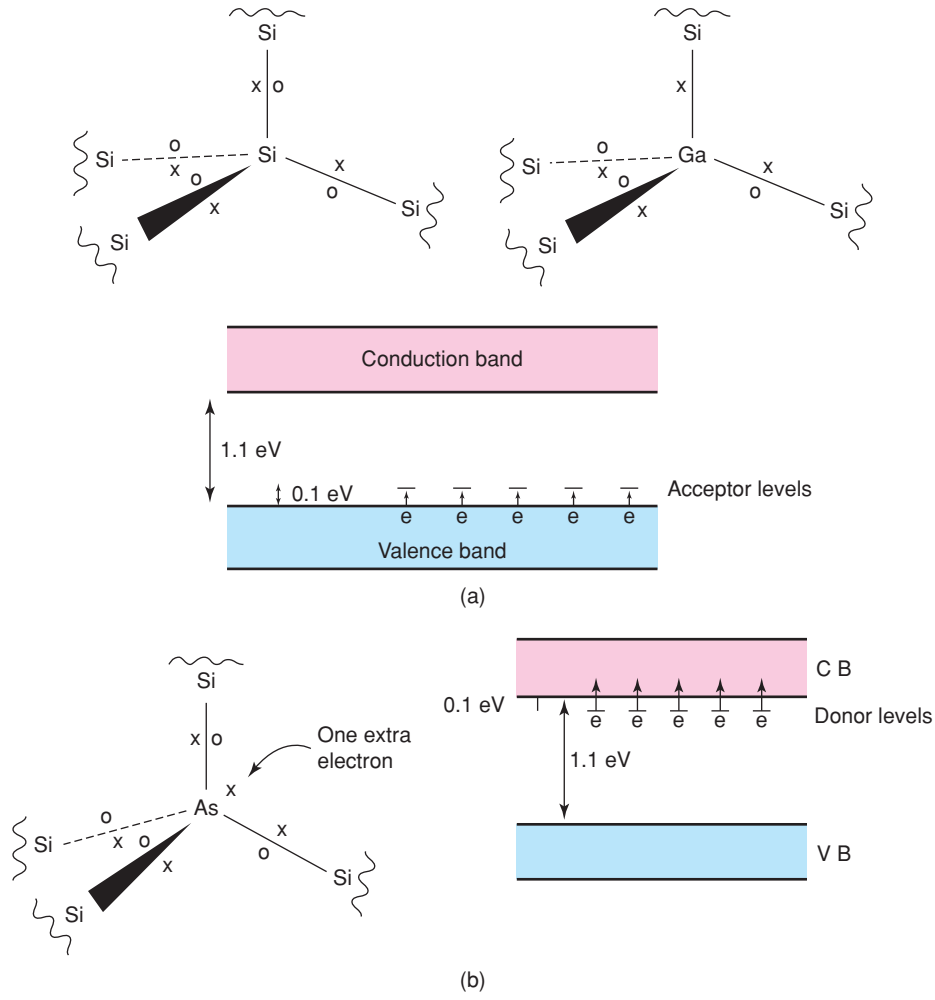


Figure 8.12 ***(a) p-Type semiconductivity in gallium-doped silicon; (b) n-type semiconductivity in arsenic-doped silicon.***

Consider now the effect of doping Si with a pentavalent element such as As. The As atoms again substitute for Si in the diamond-like structure, but for each As there is one electron more than needed to form four Si–As covalent bonds, Fig. 8.12(b). On the band description, this extra electron occupies a discrete level that is found to lie about 0.1 eV below the bottom of the conduction band. Again, electrons in these levels cannot move directly as there are insufficient of them to overlap and form a continuous band. However, they do act as *donor levels* because electrons in them have sufficient thermal energy to get up into the conduction band where they are free to move. Such a material is known as an *n-type semiconductor*.

To summarise, the main differences between extrinsic and intrinsic semiconductors are as follows:

(i) Extrinsic semiconductors have higher conductivities than similar intrinsic semiconductors at low temperatures. Thus, at 25 °C, pure Si has an intrinsic conductivity of $\sim 10^{-2}\ \Omega^{-1}\ \text{cm}^{-1}$. By appropriate doping, this increases by several orders of magnitude.

(ii) The conductivity of extrinsic semiconductors is accurately controlled by the concentration of dopant. Materials with desired values of conductivity may therefore be designed and manufactured. With intrinsic semiconductors, conductivity is strongly dependent on temperature and the presence of stray impurities.
(iii) The extrinsic region can extend over a wide temperature range by choosing (a) a material with a large intrinsic band gap and (b) dopants whose associated energy level is close to the relevant valence or conduction band.

8.4.2 Electrical properties of semiconductors

The electrical conductivity of semiconductors is given by equation (8.7); the magnitude of the conductivity and its temperature dependence are governed essentially by the carrier concentration, n, which is independent of temperature in the extrinsic region but temperature dependent in the intrinsic region, as given by equation (8.8). In the intrinsic region, for each electron that is promoted into the conduction band, a hole is left behind in the valence band and therefore

$$[e] = [h] = [n] \tag{8.9}$$

where $[n]$ is the defect concentration. Creation of intrinsic defects and carrier concentrations may be regarded as an equilibrium process:

$$null \leftrightarrow e' + h^{\bullet} \tag{8.10}$$

for which the equilibrium constant, K, may be written as:

$$K = [e][h] = [n]^2 \tag{8.11}$$

For extrinsic semiconductors, in which either the electron or hole concentration is increased as a consequence of adding either donor or acceptor dopants, respectively, it can be assumed that K is unaffected by small dopant concentrations and, therefore, any increase in $[e]$ is compensated by a decrease in $[h]$ and vice versa. This then leads on to the notion that extrinsic semiconductors contain both *majority* and *minority carriers.*

In many semiconductor device applications, the movement of charge carriers controls the speed at which devices operate. Three factors contribute to this movement. First, the *drift velocity*, v_{d}, depends directly on the applied electric field strength, E, according to

$$v_d = \mu_{\mathrm{d}} E \tag{8.12}$$

where μ_d is the drift mobility; μ_{d} is controlled by two terms, electron–phonon collisions, which increase with increase in temperature, and electron–impurity interactions, which increase with increase in impurity concentration, equation (8.1); one reason why high purity of semiconductors is essential is so that the deleterious effects of unwanted impurities, and their effect on mobility, are reduced. The purification of semiconductors using the zone refining technique is described in Chapters 4 and 7.

Second, carriers have an intrinsic *thermal velocity*, v_{t}, which is temperature dependent. Carriers may be treated as gas-like particles, which is an over-simplification but nevertheless useful. From the kinetic theory of gases, their energy is given by

$$KE = {}^1\!/_2 m^* v_t^2 = 3kT/2 \tag{8.13}$$

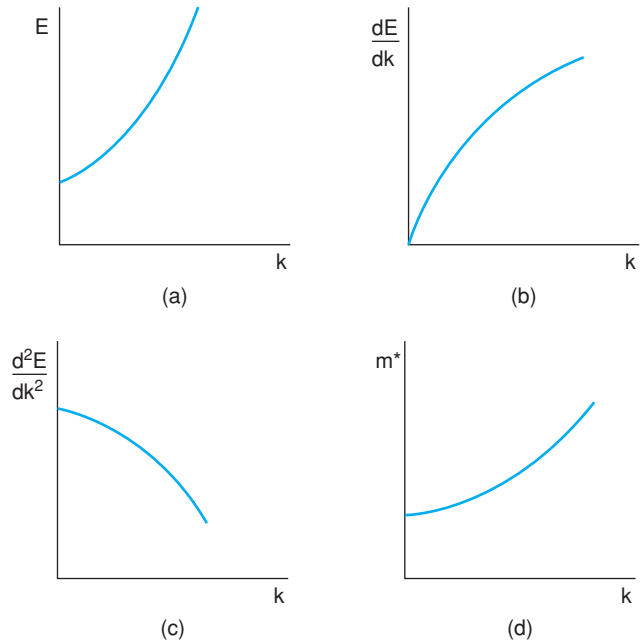


Figure 8.13 ***Variation of effective mass, m*, with wave vector, k (d) and its derivation from E–k space (a, b, c).***

where m^* is the *effective mass* of the carriers. Rearranging:

$$v_t = (3kT/m^*)^{1/2} \tag{8.14}$$

The concept of effective mass was introduced to allow Newton's laws of motion to be applied to electrons treated as particles. In the nearly free electron theory, the velocity of electrons is modified by electron–phonon collisions and, schematically, this reduction in velocity is taken into account by increasing the effective mass of the electrons. Electron momentum, p, is related to the wave number, k, by equations (3.41)–(3.44) and the idealised E–k parabolic relationship shown in Fig. 3.27(c). The effective mass is then given by

$$m^* = h^2/(\mathrm{d}^2E/\mathrm{d}k^2) \tag{8.15}$$

The E–k curve is shown in Fig. 3.27(c) and in part in Fig. 8.13(a), from which $\mathrm{d}E/\mathrm{d}k$–k is shown schematically in (b) and $\mathrm{d}^2E/\mathrm{d}k^2$–$k$ in (c); from equation (8.15), m^* is inversely proportional to $\mathrm{d}^2E/\mathrm{d}k^2$ and therefore the m^*–k dependence is as shown in (d).

The third factor that contributes to charge carrier movement arises when there is a concentration gradient of charge carriers through the sample leading to *diffusion* of carriers away from the region of higher concentration. The carrier motion that gives rise to thermal velocity is the same motion that is responsible for diffusion; in the presence of a concentration gradient, there is a net migration of carriers from regions of high concentration to those of low concentration. When a uniform carrier distribution is achieved, the diffusional processes still occur but there is no longer any net motion in a particular direction.

The combined effects of drift velocity, thermal velocity and diffusion give the overall motion of charge carriers, which is usually different for holes and electrons. In both Si and GaAs, electron velocity is higher than hole velocity and this needs to be taken into consideration in the fabrication of high-speed devices.

Table 8.4 ***Typical values of v_d and v_t for n-type Si at 25 °C***

Assume:	$m^* = 1.18m_0 = 1.18 \times 9.11 \times 10^{31}$ kg where m_0 is the mass of the electron $\mu_d = 0.15$ m^2 V^{-1} s^{-1} $V_{appl} = 2$ kV m^{-1}
Then, from equation (8.12):	$v_d = 300$ m s^{-1}
and from equation (8.14):	$v_t = 1.075 \times 10^5$ m s^{-1}

Source: Reproduced with permission from S. A. Holgate, *Understanding Solid State Physics*, © 2010 Taylor & Francis Group LLC.

Although the value of the current flowing in a particular direction is given by v_d, with a contribution from any diffusion effects, it is interesting to compare the relative magnitudes of v_d and v_t as summarised for a typical case in Table 8.4 . Typical values for the effective mass and electron mobility are given and assuming a typical small applied voltage of 2 V mm^{-1}, we see that the calculated thermal velocity is several orders of magnitude greater than the drift velocity. This shows that the applied voltage merely gives a small bias to the random motion of electrons, leading to a gradual drift in a particular direction; it is certainly incorrect to regard the effect of a small applied voltage as forcing electrons to move in a particular direction.

From the drift velocity, it is possible to estimate the mean free path or 'time of flight', τ, between successive electron–phonon collisions, since velocity is the product of acceleration and time: acceleration is given by force/mass; in this case, force is the product of the voltage, V, applied to a charge, e; therefore,

$$v_d = (Ve)(m^*)^{-1}\tau \tag{8.16}$$

Combining equation (8.16) with equation (8.12) gives

$$\mu = e(m^*)^{-1}\tau \tag{8.17}$$

8.4.3 Oxide semiconductors

Band theory is able to account satisfactorily for the properties of semiconductors such as Si and Ge. In certain transition metal compounds, however, it appears not to be applicable, mainly because the outer valence orbitals on adjacent atoms do not overlap sufficiently strongly to form bands of delocalised energy levels. For instance, in MnO, which has the rock salt structure, the t_{2g} orbitals on Mn^{2+} (a d^5 ion) contain three d electrons only since it is in the high-spin configuration (the other two electrons are in the e_g orbitals). If overlap of the t_{2g} orbitals on adjacent atoms occurred, then a half-filled band would be expected, with consequent metallic conductivity. This does not occur and MnO is instead a semiconductor.

A similar situation exists with FeO and CoO, which are also semiconductors. In TiO and VO, however, significant overlap of the t_{2g} orbitals must occur since these materials are metallic. The degree of orbital overlap required for band formation is difficult to quantify, in part because of the ephemeral nature of the orbitals themselves. Nevertheless, the contrast in properties of VO and MnO indicates that it is critically important. The explanation of this variation in properties is that, on progressing across the transition metal series, electrons in the d orbitals experience an increased nuclear charge with increasing atomic number. Consequently, they are more tightly held to the individual atoms and become localised rather than itinerant (i.e. in orbitals or bonds rather than in bands).

Semiconductors for which band theory is not applicable are termed *hopping semiconductors*. The conduction electrons are localised on individual atoms but can hop to adjacent atoms provided that they gain sufficient energy. There is an activation energy to be overcome, therefore, similar to the band gap of intrinsic semiconductors, but the freed electrons get no further than adjacent atoms where they become trapped again.

The conductivity of semiconducting transition metal compounds often increases when the transition element is present in more than one oxidation state. A good example of such *mixed valency semiconductors* is provided by the oxides of nickel. Stoichiometric nickel(II) oxide, NiO, is a pale-green solid of very low conductivity. Its colour is due to internal *d–d* transitions within the octahedrally coordinated Ni^{2+} ions and, hence, the *d* electrons must be localised on the individual cations. On heating NiO in air at, e.g., 800 °C, it oxidises to a material of empirical stoichiometry, $Ni_{1-x}O$, where $x \approx 0.1$, which is black and semiconducting. It has the same basic rock salt structure as NiO, but the cation sites contain a mixture of Ni^{2+} and Ni^{3+} ions and cation vacancies, corresponding to the formula

$$Ni^{2+}_{1-3x}Ni^{3+}_{2x}V_x O$$

The electronic conductivity arises because electrons can hop readily from Ni^{2+} to Ni^{3+} ions. Although the nickel ions themselves do not move, the net result of this electron transfer is the same as if the Ni^{3+} ions were moving, but in the opposite direction to the electron transfer. The Ni^{3+} ions are effectively positive holes and, hence, black nickel oxide is a p-type semiconductor. In contrast to Ga-doped Si, which is also p-type, the Ni^{2+}/Ni^{3+} exchange is thermally activated and therefore, highly temperature dependent.

The disadvantage of using oxidised NiO as a semiconductor is that its conductivity is difficult to control; it depends on the degree of oxidation, which is affected by both temperature and the partial pressure of oxygen. In *controlled valency semiconductors*, the concentration of, say, Ni^{3+} ions is controlled by the addition of an aliovalent dopant. For instance, lithium oxide reacts with nickel oxide and oxygen to form solid solutions of formula:

$$Li_xNi^{2+}_{1-2x}Ni^{3+}_{x}O$$

In these, the concentration of Ni^{3+} ions, and hence the conductivity, depend directly on the concentration of Li^+ ions. The magnitude of the conductivity varies enormously with x, from $\sim 10^{-10}\ \Omega^{-1}\ cm^{-1}$ for $x = 0$ to $\sim 10^{-1}\ \Omega^{-1}\ cm^{-1}$ for $x = 0.1$ at 25 °C.

Other examples of controlled valency materials associated with solid solution mechanisms and aliovalent doping are given in Chapter 2. It is a widespread phenomenon and arises whenever the charge balance mechanism on aliovalent substitution is electronic rather than ionic. The electrical properties of cuprate high-T_c superconductors, Section 8.3, depend critically on this same mechanism of valency control. These materials also show a transition from semiconducting to metallic behaviour depending on dopant concentration. Thus $YBa_2Cu_3O_6$, which contains a mixture of Cu^{2+} and Cu^+ ions, is a semiconductor; on increasing the oxygen content to convert this into $YBa_2Cu_3O_7$, which contains a mixture of Cu^{2+} and Cu^{3+} ions, the material becomes metallic (and superconducting below 90 K); this is attributed to formation of a band structure consisting of overlapping Cu 3*d* and O 2*p* orbitals.

8.4.4 Applications of semiconductors

The main use of semiconductors is in solid state devices such as transistors, silicon chips and photocells. These devices are possible because the energies of electrons in the highest occupied levels are different depending on whether the materials are doped to be either n-type or p-type. This is shown in Fig. 8.14(a) for a material such as Si that is doped with a trivalent element such as B to make it p-type or with a pentavalent element such as P to make it n-type. In the p-type material, the acceptor levels are just above the top of the valence

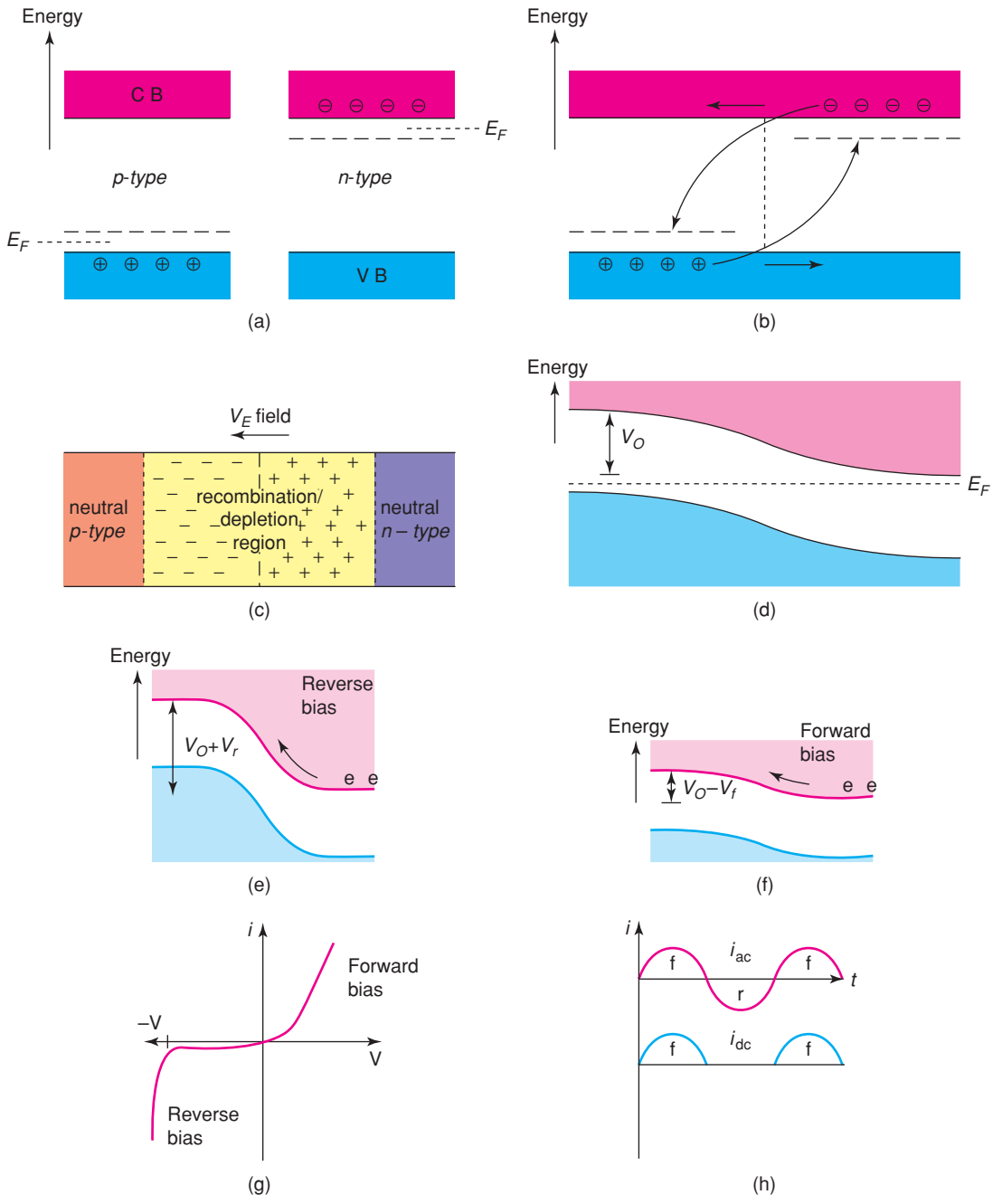


Figure 8.14 ***A p–n junction. (a) Energy levels in p-type and n-type semiconductor; (b) p- and n-type regions in contact leading to (c) creation of a charge depletion region by spontaneous diffusion and (d) equalisation of*** E_F ***across the p–n junction. The resulting band bending is (e) increased in reverse bias and (f) decreased in forward bias. (g) Voltage–current characteristics as a function of bias leading to (h) current rectification.***

band; many of these are occupied by promotion of electrons from the valence band since their thermal energy is sufficient to make the transition and, consequently, holes are present in the top of the valence band. In the n-type region, the donor levels are just below the bottom of the conduction band and some of the electrons in the donor levels have sufficient thermal energy to be promoted into the conduction band. Hence the Fermi level, E_F, is higher in the n-type region than in the p-type region; this difference in E_F at the interface between differently doped materials is the vital property that has allowed the entire microelectronics industry to evolve.

In a *p–n junction*, Fig. 8.14(b), n- and p-type regions are brought into contact, either by creating a compositional gradient of donor/acceptor impurities within a single crystal of, e.g., Si to create a *homojunction,* or by epitaxial growth of a second material, doped differently, on top of a substrate, to create a *heterojunction*. Before contact, the n- and p-type regions are electrically neutral but, because of the difference in Fermi levels at the junction between the differently-doped regions, an effective concentration gradient of charge carriers of both types occurs across the junction; spontaneous diffusion processes cause electrons from the n-type region to spill over into the p-type region, where they may enter energy levels of the conduction band or more probably, drop down to combine with holes in the valence band. Conversely, holes in the valence band of the p-type region may diffuse into the valence band of the n-type region (of course, holes diffusing in one direction are, in reality, electrons diffusing in the opposite direction) or may annihilate electrons in the conduction band by hole–electron combination. The Fermi energy of electrons is similar to their electrochemical potential; provided that a difference in potential exists, electrons flow spontaneously from a region of high potential to one of lower potential.

The net effect of these processes is to create a region, either side of the physical junction between the dissimilar materials, which is depleted of charge carriers of either kind. This region is no longer neutral due to the extra negative charge at the p-type side of the interface and the extra positive charge at the n-type side. This charged *depletion region*, also known as a *space charge layer,* acts to prevent further homogenisation of charge carriers across the entire material. We may also think of the holes in the p-type region as electron traps: once they are occupied, they become negatively charged.

The net effect of the diffusion processes is that a reverse electric field, V_E, is set up as shown in Fig. 8.14(c), which effectively causes drift of charge carriers in the opposite direction to that of the spontaneous diffusion. An equilibrium situation is reached in which diffusion and drift currents cancel; since there is no net driving force for charge transfer in either direction, the Fermi levels across the interface must equalise, (d).

Creation of a charged depletion region leads to *band bending*, which is a modification, locally, of valence band and conduction band energies, as shown schematically in Fig. 8.14(d). Thus, removal of negative charge from the n-type side leads to greater attraction of valence band and remaining conduction band electrons for the atomic nuclei and hence lowering of energy. Conversely on the p-type side, energy levels of valence and conduction bands rise in the vicinity of the *p–n junction* due to the presence of the extra negative charge. The band bending hypothesis, (d), helps to explain why no net charge migration occurs once the depletion region is set up. Thus, flow of conduction band electrons from n- to p-type regions becomes much reduced because this is an uphill process. Hence few of the majority carriers are able to cross the junction in the n–p direction. Conversely, a small number of minority carriers in the conduction band in the p-type region are able to cross the junction in the p–n direction as this is a downhill process.

The effect of applying a dc bias to a p–n junction is shown schematically in Fig. 8.14(e) and (f). Under *reverse bias*, (e), a negative potential is applied to the p-type end, which has the effect of increasing still further the potential barrier for migration of electrons from n- to p-type regions to a value $V_0 + V_r$, where V_0 is the barrier height in the p–n junction at equilibrium without a dc bias, (d), and V_r is the applied reverse potential. Consequently, fewer electrons are able to move uphill from the n- to p-type regions whereas the drift current of minority carriers in the opposite direction is retained and a very small current therefore persists. In *forward bias*, by contrast, a positive potential, V_f, is applied to the p-type end and the energy barrier for electron flow is reduced to $V_0 - V_f$. Flow of majority carriers from n- to p-type regions is therefore able to

recommence and a continuous current flows. Electrons enter from the right-hand electrode at the n-type end, flow through the conduction band, drop into the valence band of the p-type region where they recombine with holes, but fresh holes are generated at the left-hand electrode where electrons leave the valence band to complete the dc current flow.

Consequently, current flow is greatly reduced under reverse bias and enhanced under forward bias. This is shown in Fig. 8.14(g) and forms the basis of *rectifying action* in which an ac current is converted into a dc current (h). The p–n junction is therefore the solid state electronic equivalent of the diode rectifier valve. Si-based p–n junctions have completely replaced diode valves in electronic circuits. A more complex arrangement is the *p–n–p* or *n–p–n junction*. This acts as a current or voltage amplifier and forms the basis of the *transistor* which has completely taken the place of triode valves.

Controlled valency, hopping semiconductors are used in *thermistors*, thermally sensitive resistors. In these, the conductivity has a large temperature dependence with large E. For example, $Li_{0.05}Ni_{0.95}O$ shows Arrhenius-type conductivity behaviour over a wide temperature range up to ~200 °C with $E \approx 0.15$ eV. Since the conductivity is reproducible, lithium nickel oxide can be used in devices to control and measure temperature. In order to achieve reproducibility, materials that are insensitive to impurities must be used, e.g. Fe_3O_4, Mn_2O_3, Co_2O_3-doped NiO and certain spinels.

Some semiconductors are *photoconductive*; their conductivity increases greatly on irradiation with light. Amorphous selenium is an excellent photoconductor and forms an essential component of the photocopying process (see Chapter 10). Conventional band theory cannot be used to explain the properties of amorphous materials such as Se since they lack long-range periodicity.

8.5 Ionic Conductivity

In most ionic solids, ions are trapped on their lattice sites. They vibrate continuously at infrared frequencies but rarely have enough thermal energy to escape from their lattice sites. If they are able to escape and move into adjacent lattice sites we have the makings of *ionic conduction*, also referred to as *ionic migration*, *hopping* or *diffusion*. Ionic conduction is easier at higher temperatures and especially if crystal defects are involved. In defect-free solids, there are no atom vacancies and interstitial sites are completely empty. For ionic conduction to occur, a minimum requirement is that either some sites are vacant and adjacent ions can hop into the vacancies, leaving their own sites vacant, or there are some ions in interstitial sites which can hop into adjacent interstitial sites. At higher temperatures, ions have greater thermal energy and vibrate more vigorously; also, defect concentrations are higher. For example, the conductivity of NaCl at ~800 °C, just below its melting point, is $\sim 10^{-3}\ \Omega^{-1}\ cm^{-1}$, whereas at room temperature, pure NaCl is an insulator with a conductivity much less than $10^{-12}\ \Omega^{-1}\ cm^{-1}$.

In contrast to most ionic solids, there exists a small group of solids called, variously, *solid electrolytes*, *fast ion conductors* and *superionic conductors*, in which one set of ions can move quite easily. Such materials often have special crystal structures with open tunnels or layers through which the mobile ions may move. The conductivity values, e.g. $10^{-3}\ \Omega^{-1}\ cm^{-1}$ for Na^+ ion migration in β-alumina at 25 °C, are comparable to those observed in strong liquid electrolytes. There continues to be great interest in studying the properties of solid electrolytes, developing new ones and extending their range of applications in solid state electrochemical devices. Let us first consider the behaviour of typical ionic solids and how their electrical properties are controlled by crystal defects.

Ionic conductivity in pure, and doped, alkali halides and Ag halides is a well-researched and well-understood topic. The magnitude of the ionic conductivity in these materials is generally too small for practical applications but the detailed measurements made on single crystals several decades ago have provided us with much information on the thermodynamics of defect formation, migration and aggregation

Figure 8.15 ***Migration of cation vacancies, i.e. Na^+ ions, in NaCl.***

which underpins much of our understanding of ionic solids and their behaviour. We therefore consider this topic in some detail here.

8.5.1 Alkali halides: vacancy conduction

In alkali metal halide crystals, cations are usually more mobile than anions. A section through the NaCl structure is shown in Fig. 8.15; a Na^+ ion is moving into an adjacent vacant cation site, leaving its own site vacant. The Na^+ ion that has moved can travel no further because (a) there are no other vacant sites in the vicinity for it to move into and (b) migration of Na^+ into tetrahedral interstitial sites in NaCl appears not to occur to any appreciable extent. The *cation vacancy* may continue to move, however, because it is always surrounded by 12 Na^+ ions, any one of which can jump into the vacancy. It is convenient, therefore, to regard cation vacancies as the main current carriers in NaCl. Anion vacancies are also present, but are less mobile than the cation vacancies.

Ionic conductivity in NaCl depends on the presence of cation vacancies, which in turn depends on purity and thermal history. Vacancies are created by one of two methods. One is simply to heat the material. The number of vacancies present in thermodynamic equilibrium increases exponentially with increase in temperature, equation (2.12), and is the number that is *intrinsic* to pure NaCl. The other method involves addition of aliovalent impurities: vacancies are created so as to preserve charge balance. For example, addition of a small amount of $MnCl_2$ yields a solid solution of formula

$$Na_{1-2x}Mn_xV_xCl$$

where, for each Mn^{2+} ion, there is an associated cation vacancy, V. Such vacancies are *extrinsic* and are not present in pure NaCl. At low temperatures (e.g. 25 °C), the number of thermally generated intrinsic vacancies is very small and, unless the crystal is very pure, is much less than the concentration of extrinsic vacancies. With increase in temperature, a changeover from extrinsic to intrinsic behaviour occurs when the thermally generated intrinsic vacancy concentration exceeds the impurity-controlled extrinsic vacancy concentration.

The temperature dependence of ionic conductivity is usually given by the Arrhenius equation:

$$\sigma = A\exp\left(-\frac{E}{RT}\right) \tag{8.18}$$

The pre-exponential factor, A, contains several terms, including the vibrational frequency of the potentially mobile ions. A graph of $\log_e\sigma$ against T^{-1} should give a straight line of slope $-E/R$; more usually, conductivity is plotted as $\log_{10}\sigma$, in which case, the slope is $-E/R\log_{10}e$. In some treatments, a reciprocal temperature term is included in the prefactor, A, and then $\log_{10}\sigma T$ is plotted against T^{-1}. This may make a small difference to the value of the activation energy that is calculated from the slope.

A schematic Arrhenius plot for NaCl is shown in Fig. 8.16. At low temperatures, the number of vacancies is dominated by the impurity level and is constant for a given impurity concentration. A set of parallel lines

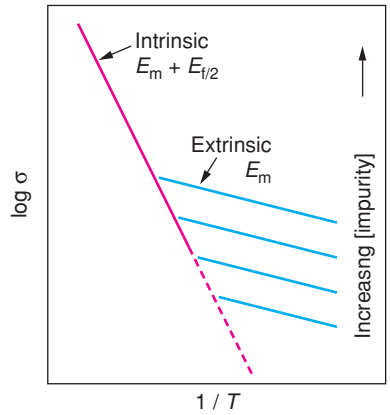


Figure 8.16 ***Schematic ionic conductivity of doped NaCl crystals. Parallel lines in the extrinsic region correspond to different dopant concentrations.***

is shown, each of which corresponds to a different amount of dopant. In this extrinsic region, the temperature dependence of σ depends only on the cation mobility, μ, equation (8.7), whose temperature dependence is also given by an Arrhenius expression:

$$\mu = \mu_0 \exp\left(-\frac{E_m}{RT}\right) \tag{8.19}$$

where E_m is the activation energy for cation vacancy migration.

8.5.1.1 Activation energy for ion hopping: geometric considerations

In order to understand the origin of E_m, consider the possible paths taken by an Na^+ ion in jumping from its lattice site into an adjacent vacancy. A small part of the NaCl structure is shown in Fig. 8.17(a). It is a simple cube with Na^+, Cl^- ions at alternate corners and corresponds to one-eighth of the unit cell of NaCl.

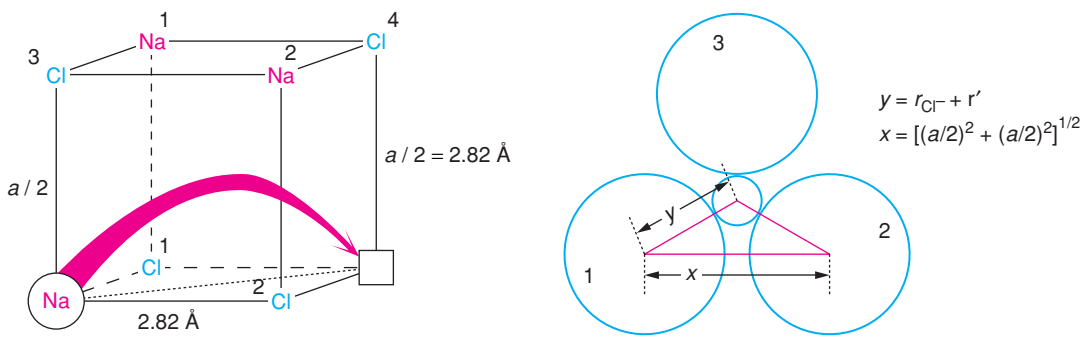


Figure 8.17 ***(a) Pathway for Na^+ migration in NaCl. (b)Triangular interstice of radius r′ through which a moving Na^+ ion must pass in NaCl. Circles 1–3 are Cl^- ions.***

One corner Na^+ site is shown empty and an Na^+ ion from one of the other three corners moves to occupy it. The direct jump (dotted line) across the cube face is not possible because Cl^- ions 1 and 2 are very close, if not in actual contact, and an Na^+ ion would not be able to squeeze between them. Instead, Na^+ must take an indirect route (curved arrow) that passes through the middle of the cube. At the cube centre is an interstitial site that is equidistant from all eight corners. Four of the corners are occupied by Cl^- ions which are arranged tetrahedrally about the cube centre. Before the moving Na^+ ion arrives at this central interstitial site, it has to pass through a triangular window formed by Cl^- ions 1, 2 and 3. Let us calculate the size of this window in order to appreciate how difficult it is for Na^+ to squeeze through.

For NaCl, the unit cell parameter, a, equals 5.64 Å. The Na–Cl bond length, Fig. 8.17(a), is $a/2 = 2.82$ Å. This equals $(r_{Na^+} + r_{Cl^-})$, assuming anions and cations to be in contact. Tabulated ionic radii (which vary somewhat, depending on the table that is consulted) of Na^+ and Cl^- are ~0.95 and ~1.85 Å; as expected, the Na–Cl bond length calculated from the sum of these radii is ~2.80 Å, close to the measured value.

In close packed structures, such as NaCl, anions are either in contact or in close proximity. The Cl^- ions, 1, 2 and 3, form part of a *cp* layer and the distance Cl(1)–Cl(2) is given by $[(a/2)^2 + (a/2)^2]^{1/2} = 3.99$ Å. This is about 0.3 Å larger than given by $2r_{Cl^-}$ and so adjacent Cl^- ions in NaCl are not quite in contact.

The radius r' of the triangular window or bottleneck formed by Cl^- ions, 1, 2 and 3, may be calculated from Fig. 8.17(b) as follows:

$$\cos 30^\circ = \frac{x/2}{y} = \frac{1.995}{(r_{Cl^-} + r')} \tag{8.20}$$

Therefore,

$$(r_{Cl^-} + r') = \frac{1.995}{\cos 30^\circ} = 2.30 \text{ Å} \tag{8.21}$$

If

$$r_{Cl^-} = 1.85 \text{ Å}, \text{ then } r' = 0.45 \text{ Å} \tag{8.22}$$

The radius r' of the interstitial site at the cube centre, Fig. 8.17(a), may be calculated similarly. The body diagonal of the cube is $(2r_{Cl^-} + 2r')$. Therefore,

$$\begin{aligned} 2(r_{Cl^-} + r'') &= \left[(a/2)^2 + (a/2)^2 + (a/2)^2\right]^{\frac{1}{2}} \\ &= 4.88 \text{ Å} \end{aligned}$$

and

$$r'' = 0.59 \text{ Å} \tag{8.23}$$

We can see, therefore, that the jumping of a Na^+ ion in NaCl is difficult and complicated. First, it has to squeeze through a narrow triangular gap of radius 0.45 Å, whereupon it finds itself in a small tetrahedral interstitial site of radius 0.59 Å. The residence time here is short since this site has a hostile environment with two Na^+ ions, 1 and 2, at a distance of 2.44 Å and also four Cl^- ions at the same distance. The Na^+ ion leaves by squeezing through another gap of radius 0.45 Å (formed by Cl^- ions 1, 2 and 4) to occupy the vacant octahedral

site on the other side. These calculations are inevitably idealised since relaxation or distortion of the structure must occur in the vicinity of the defects, thereby modifying the distances involved. They do nevertheless show that the migration of Na^+ ions is difficult and is associated with a considerable activation energy barrier.

In the extrinsic region of Fig. 8.16, conductivity depends on both vacancy concentration and mobility. It is given by combining equations (8.7) and (8.19) to give

$$\sigma = ne\mu_0 \exp\left(-\frac{E_m}{RT}\right) \tag{8.24}$$

In the high-temperature intrinsic region, the thermally induced vacancy concentration is greater than that associated with the dopant. Hence n is temperature dependent and is also given by an Arrhenius equation:

$$n = N \exp\left(-\frac{E_f}{2RT}\right) \tag{8.25}$$

This equation is the same as equation (2.12), in which $E_f/2$ is the activation energy for formation of 1 mol of cation vacancies, i.e. half the energy required to form 1 mol of Schottky defects. The vacancy mobility is again expressed by equation (8.19) and the overall conductivity in the intrinsic region is given by

$$\sigma = Ne\mu_0 \exp\left(-\frac{E_m}{RT}\right) \exp\left(-\frac{E_f}{2RT}\right) \tag{8.26}$$

i.e.

$$\sigma = A \exp\left(-\frac{E_m + E_f/2}{RT}\right) \tag{8.27}$$

In Fig. 8.16, schematic Arrhenius plots for NaCl crystals with varying degrees of purity are shown. The parallel lines in the extrinsic region correspond to different impurity levels; the single line in the intrinsic region shows the conductivity to be unaffected by the impurities. The gradient of the intrinsic line is greater than that of the extrinsic lines; if both gradients can be measured, then E_m and E_f may be determined.

8.5.1.2 Ionic conductivity of NaCl crystals

Experimental data for NaCl single crystals, Fig. 8.18, indicate the simple behaviour in Fig. 8.16 to be somewhat idealised since several complicating factors are present. Stages I and II correspond to the intrinsic and extrinsic regions of Fig. 8.16. The dashed lines represent extrapolations of regions I and II and show where deviations occur. Stage I′ occurs close to melting (802 °C) and has been attributed to two possible causes. One is that anion vacancies become increasingly mobile and make a significant contribution to σ. The other is that long-range Debye–Hückel interactions between cation and anion vacancies become appreciable at high vacancy concentration at high temperatures. An attractive force results (similar to the Debye–Hückel interaction between ions in solutions) which has the effect of reducing E_F. Thus, vacancy formation is easier, the vacancy concentration increases and σ increases. It is not known which explanation is correct for NaCl, although the Debye–Hückel effect is probably the cause of a similar deviation in σ at high temperatures in AgCl single crystals (see the discussion of Fig. 2.4).

Below ~390 °C (in this particular NaCl crystal), σ deviates downwards from the ideal extrinsic line, stage III. This is attributed to the formation of defect complexes, e.g. cation vacancy/anion vacancy pairs or cation vacancy/aliovalent cation impurity pairs. These complexes form by short-range attraction between defects of

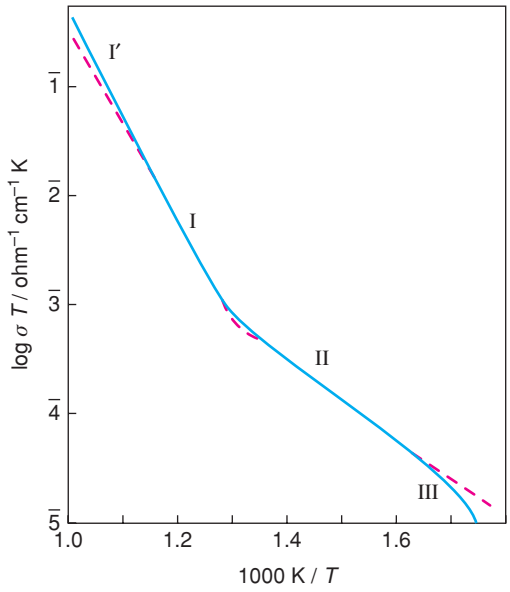


Figure 8.18 ***Ionic conductivity of 'pure' NaCl as a function of reciprocal temperature.***

opposite charge and are distinct from the long-range Debye–Hückel interactions mentioned above. In stage III, cation vacancies must first acquire the additional energy needed to dissociate themselves from the defect complexes before they can move. As a result, the activation energy of stage III is greater than E_m (stage II).

Data for the various activation energies associated with ionic conduction in NaCl are given in Table 8.5. The probable reason for the variation in E_m values is that other defects, especially dislocations, are inevitably present and have a considerable influence on ion migration: the regions of the crystal structure close to a dislocation are stressed and distorted and migration of cations along 'dislocation pipes' may be easier than through regions of undistorted crystal. If so, conductivity and the magnitude of E_m may depend on the number and distribution of dislocations and hence on thermal history of the crystals. It is difficult to make quantitative measurements of the effect of dislocations on σ because their number cannot be controlled and measured accurately.

8.5.1.3 Extrinsic conductivity in NaCl: control by aliovalent doping

The effect of impurities on σ is amenable to study since their number may be controlled and measured and they have a dramatic effect on σ in the extrinsic region. The concentration of Schottky defects at equilibrium is given by the law of mass action, equation (2.5):

$$K = \frac{[V'_{Na}][V^{\bullet}_{Cl}]}{[Na^{x}_{Na}][Cl^{x}_{Cl}]}$$

It is assumed that K is unaffected by the presence of small amounts of aliovalent impurity and that the denominator, the number of occupied sites, is essentially constant and equal to 1. Therefore,

$$[V'_{Na}][V^{\bullet}_{Cl}] = \text{constant} = x_0^2 \tag{8.28}$$

Table 8.5 **Conductivity in NaCl crystals, other ionic solids and characteristics of solid electrolytes**

Property	Value
(a) Conductivity in NaCl crystals	
Process	Activation energy/eV
Migration of Na^+	0.65–0.85
Migration of Cl^-	0.90–1.10
Formation of Schottky pair	2.18–2.38
Dissociation of vacancy pair	~1.3
Dissociation of vacancy–impurity (e.g. Mn^{2+}) pair	0.27–0.50
(b) Conductivity in other ionic solids	
Solid	*E/eV*
MgO	3–5 for Mg^{2+}, O^{2-}
Na β-alumina	0.16 for Na_i
Na β^{11}-alumina	0.08–0.35 for Na_i
$RbAg_4I_5$	0.05 for Ag^+
(c) Characteristics of solid electrolytes	
Only one species mobile (unlike liquid electrolytes)	
Large number of carriers (unlike most ionic solids)	
Easy hops, low E (~0.1–0.9 eV)	
A relatively rare phenomenon	
Occurs in glasses, ceramics, crystals, gels	
Occurs mainly with M^+, F^-, O^{2-} ions	
Structures have partial site occupancies of mobile species	
Cooperative conduction mechanisms are common	

where, in the intrinsic region, $x_0 = [V'_{Na}] = [V^{\bullet}_{Cl}]$. If the number of cation vacancies in the extrinsic region is increased by adding aliovalent cation impurity, then the number of anion vacancies must decrease in accord with equation (8.28). Let x_a and x_c be the concentration of anion and cation vacancies under extrinsic conditions (and so $x_a \neq x_c$) and c be the concentration of divalent impurity cations. Then,

$$x_c = x_a + c \tag{8.29}$$

This condition arises because both anion vacancies and divalent cation impurities carry a net positive charge of $+1$, whereas the cation vacancy has a charge of -1. Overall, charge balance must be retained. On combining equations (8.28) and (8.29) and solving the resulting quadratic, the following positive values of x_c and x_a are obtained:

$$\begin{aligned} x_c &= \frac{c}{2}\left[1 + \left(\frac{4x_0^2}{c^2}\right)^{\frac{1}{2}}\right] \\ x_a &= \frac{c}{2}\left[\left(\frac{1+4x_0^2}{c^2}\right)^{\frac{1}{2}} - 1\right] \end{aligned} \tag{8.30}$$

If $x_0 \ll c$, then $x_c \rightarrow c$ and $x_a \rightarrow 0$; this result applies to the extrinsic region and is the same as that deduced above. If $x_0 \gg c$, then $x_c = x_0 = x_a$; this applies to the intrinsic region of conductivity.

Not all impurities lead to an increase in σ in the extrinsic region. Impurities of the same valence, e.g. K^+ or Br^-, in NaCl have little effect on σ unless present in large concentrations. Impurities that reduce the concentration of mobile species cause a reduction in σ. Therefore, if divalent anions could dissolve in NaCl crystals, the concentration of the less mobile anion vacancies would increase at the expense of the more mobile cation vacancies. This effect is not observed very much in NaCl because divalent anions are not soluble, but it is important in AgCl, as described next.

8.5.2 Silver chloride: interstitial conduction

The predominant defects in AgCl are cation Frenkel defects, i.e. interstitial Ag^+ ions associated with Ag^+ ion vacancies (see Chapter 2). The interstitial Ag^+ ions are more mobile than the Ag^+ vacancies. Two possible mechanisms for migration of interstitial Ag^+ are shown schematically in Fig. 8.19(a). In the direct *interstitial* mechanism (1), an interstitial Ag^+ ion jumps to an adjacent empty interstitial site. In the indirect or *interstitialcy* mechanism (2), a knock-on process occurs. An interstitial Ag^+ ion causes one of its four Ag^+ neighbours to move off its normal site into an adjacent interstitial site and itself occupies the vacant lattice site thereby created. It is possible to distinguish between the interstitialcy and direct interstitial mechanisms if accurate data for both diffusion and conductivity are available. In diffusion measurements, the crystal is doped with radioactive Ag^{+*} ions and their migration is followed. In conductivity measurements, all the Ag^+

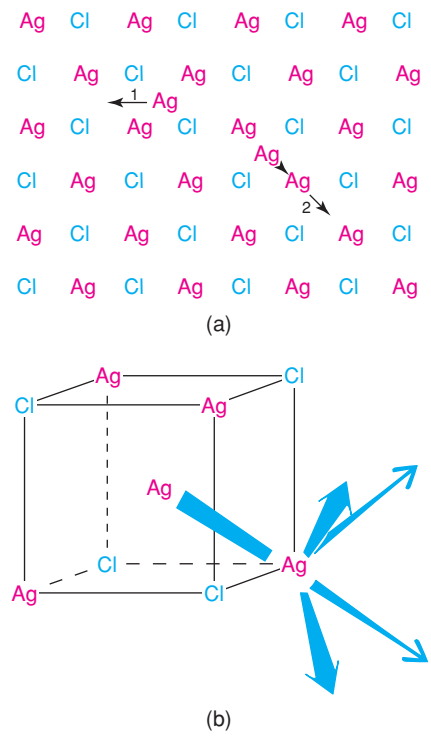


Figure 8.19 *(a) Migration of interstitial Ag^+ ions by (1) direct interstitial jump and (2) indirect interstitialcy mechanism. (b) Pathway for migration of Ag^+ by an interstitialcy mechanism.*

ions, not only the radioactive ones, contribute to the net conductivity. The *Nernst–Einstein equation* relates the coefficient for *self-diffusion*, D, to the conductivity, σ, by

$$D = \frac{kT}{fn(Ze)^2}\sigma \tag{8.31}$$

where (Ze) is the charge on the mobile ions and n is their concentration; f is a correlation factor, the *Haven ratio*, whose value depends on the mechanism of ion migration. The value of f is different for the two mechanisms shown in Fig. 8.19(a). In mechanism (2), the net distance of charge displacement is greater than either of the individual jump distances. In mechanism (1), however, the jump distance of the Ag^+ ion equals the distance of overall charge migration. Since diffusion measurements determine the movement of the ions themselves and conductivity measures the overall charge displacement, the Haven ratio is different for the two mechanisms. It has been found experimentally that the interstitialcy mechanism (2) is operative in AgCl.

The difference between the vacancy migration mechanism in NaCl and the interstitialcy mechanism in AgCl may be seen by comparing Fig. 8.17(a) and Fig. 8.19. Both NaCl and AgCl have the same rock salt structure. In the vacancy mechanism, an Na^+ ion moves from one corner of the cube to another, via an interstitial site at the cube centre which is occupied only transiently. In the interstitialcy mechanism, an Ag^+ ion moves from the interstitial site in the centre of one cube to the site in the centre of an adjacent cube by knocking on an Ag^+ ion placed at one of the corner sites.

The effect of aliovalent cation impurities on the extrinsic conductivity region of AgCl is different from that observed in NaCl. The presence of, say, Cd^{2+} again increases the number of cation vacancies, but because the product of the concentrations of cation vacancies and Ag^+ interstitials is constant, equation (2.15), the interstitial Ag^+ concentration must decrease with increasing Cd^{2+} concentration. Addition of Cd^{2+} therefore reduces the concentration of the more mobile species; the resulting Arrhenius conductivity plot is shown schematically in Fig. 8.20. An extrinsic region at lower temperatures is again observed but is displaced *downwards* to lower σ values. The degree of downward displacement increases with increase in Cd^{2+} concentration until a minimum conductivity is reached at which the conductivity due to the more numerous but less mobile cation vacancies equals that due to the less numerous but more mobile interstitial Ag^+ ions. At still higher defect concentrations, cation vacancy migration predominates and σ increases. This is shown in (b) as a plot of conductivity versus defect concentration for two temperatures. In the intrinsic

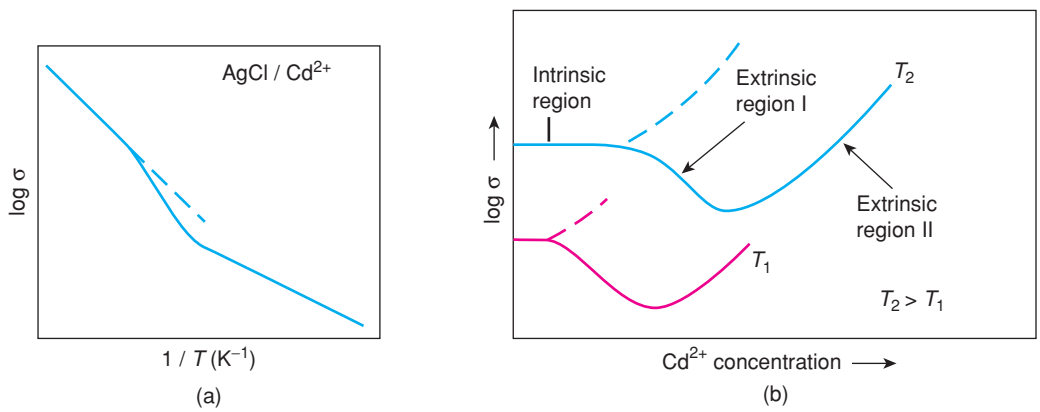


Figure 8.20 ***(a) Effect of Cd^{2+} on conductivity of AgCl crystals. (b) Effect of Cd^{2+} impurity on isothermal conductivity of AgCl. Dashed lines represent the effect of adding a divalent anionic impurity.***

region, the conductivity is independent of $[Cd^{2+}]$. In extrinsic region I, interstitial conduction predominates and σ decreases as $[Ag_i^+]$ decreases with increase in $[Cd^{2+}]$. In extrinsic region II, vacancy conduction predominates and σ increases as $[V_{Ag}]$ increases with increase in $[Cd^{2+}]$.

The relevant equations for the dependence of σ on impurity concentration, c, are obtained as follows. In order to preserve charge balance:

$$x_c = c + x_i \tag{8.32}$$

where x_i and x_c are the concentrations of interstitial Ag^+ ions and Ag^+ ion vacancies, respectively. The conductivity in the extrinsic region is given by

$$\begin{aligned}\sigma &= e(x_c\mu_c + x_i\mu_i) \\ &= e\mu_c\frac{c}{2}\left[1+\left(1+\frac{4x_0^2}{c^2}\right)^{\frac{1}{2}}\right] + e\mu_i\frac{c}{2}\left[\left(1+\frac{4x_0^2}{c^2}\right)^{\frac{1}{2}} - 1\right]\end{aligned} \tag{8.33}$$

Accurate measurements of σ for AgCl show that Fig. 8.20(a) is somewhat idealised. At high temperatures, long-range Debye–Hückel interactions become important and give an upward departure from the intrinsic slope, whereas at low temperatures, 'complexes' form between cation vacancies and aliovalent cation impurities to give a downward deviation in the extrinsic region. Defect energies for AgCl are as follows:

formation of Frenkel defect	1.24 eV
migration of cation vacancy	0.27–0.34 eV
migration of interstitial Ag^+	0.05–0.16 eV.

8.5.3 Alkaline earth fluorides

The most important defect in this group is the anion Frenkel defect in which an interstitial F^- ion occupies the centre of a cube that has eight F^- ions at the corners (Fig. 1.30). Conductivity measurements show the anion vacancy to be more mobile than the interstitial F^- ion. This contrasts with AgCl, in which the interstitial Ag^+ is more mobile than the cation vacancy. In some materials that have the fluorite structure, e.g. PbF_2, the conductivity at high temperatures becomes fairly large.

8.5.4 Solid electrolytes (or fast ion conductors, superionic conductors)

8.5.4.1 General considerations

Most crystalline materials, such as NaCl and MgO, have low ionic conductivities because, although the atoms or ions vibrate, they usually cannot escape from their lattice sites. The small group of solid electrolytes is an exception. Some characteristic features of solid electrolytes are summarised in Table 8.5(c). In these, one component of the structure, cationic or anionic, is essentially free to move throughout the structure. Solid electrolytes are, therefore, intermediate between normal crystalline solids with regular 3D structures and immobile atoms or ions and liquid electrolytes which do not have regular structures but do have mobile ions. Often, solid electrolytes are stable only at high temperatures. On cooling, they may transform to a polymorph with a low ionic conductivity and a more usual type of crystal structure, Fig. 8.21. For example, Li_2SO_4 and AgI are both poor conductors at 25 °C but at 572 and 146 °C, respectively, their structures change to give

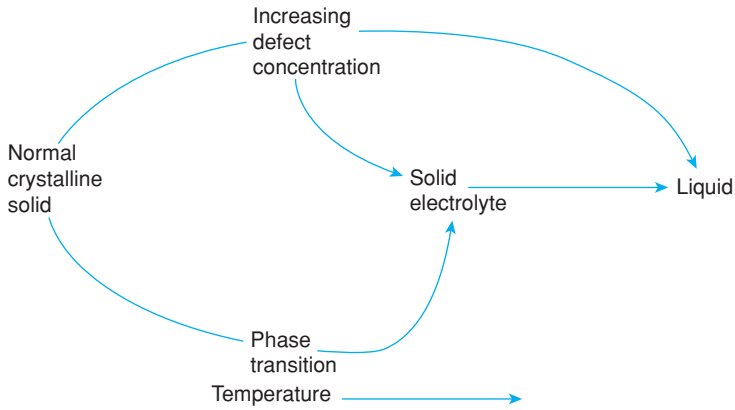


Figure 8.21 ***Solid electrolytes as intermediate between normal crystalline solids and liquids.***

polymorphs, α-Li_2SO_4 and α-AgI, that have mobile Li^+ and Ag^+ ions ($\sigma \approx 1\ \Omega^{-1}\ cm^{-1}$). On heating, the conductivity therefore increases dramatically at the phase transition.

Other solid electrolytes form as a consequence of a gradual increase in defect concentration with temperature. For example, in ZrO_2, the concentration of anion vacancies above ~600 °C is sufficiently large that zirconia is a good, high-temperature oxide ion conductor. The distinction between normal ionic solids and solid electrolytes is often not well defined, especially for materials such as ZrO_2, the behaviour of which changes gradually with increase in temperature.

It is now apparent from both theoretical and experimental results on a wide variety of materials that ionic conductivities of 0.1–10 $\Omega^{-1}\ cm^{-1}$ are the maximum that are likely to be obtained for any solid material. These values are obtained when a large proportion of the ions move at any one time. Some authors reserved the names 'superionic conductor' and 'fast ion conductor' for materials in this latter category of optimised conductivity. Although these names are in common usage, they are misnomers because the mobile ions do not have any 'super' properties, nor is their movement exceptionally fast. Rather, the high conductivities are associated with a high concentration of mobile species and a relatively low activation energy for ion migration.

The classification of solid electrolytes as intermediate between normal ionic solids and ionic liquids, Fig. 8.21, is supported by data on the relative entropies of polymorphic transitions and melting. For typical monovalent ionic materials such as NaCl, disordering of both cations and anions occurs on melting; a typical entropy of fusion is 24 J mol^{-1} K^{-1}. For AgI, the $\beta \rightarrow \alpha$ transition at 146 °C may be regarded as quasi-melting of the silver ions. It has an entropy of transition of 13.5 J mol^{-1} K^{-1}. At the melting point of AgI, only the iodide atoms are left to become disordered. This fits in with a much reduced value for the entropy of fusion, 11.3 J mol^{-1} K^{-1}. The combined entropies of transition and fusion in AgI are close to the fusion entropy of NaCl. Similar effects are observed in fluorides of some divalent metals, e.g. PbF_2 has an entropy of fusion of 16.4 J mol^{-1} K^{-1} whereas for MgF_2, a typical ionic solid, it is ~35 J mol^{-1} K^{-1}. This is because the F^- ions in PbF_2 are disordered above ~500 °C and the entropy of fusion corresponds to disordering of the Pb^{2+} ions alone.

Although the study of solid electrolytes began in earnest in the 1960s following the discovery of high ionic conductivity at room temperature in β-alumina and $RbAg_4I_5$, the possibility of high ionic conductivity in solids was recognised much earlier by Faraday. In 1839, he reported high ionic conductivity at high temperatures in both PbF_2 and Ag_2S. Then, around 1900, Nernst found that mixed oxides of ZrO_2 and Y_2O_3

glowed white-hot when an electric current was passed through them at high temperatures. He attributed this to oxide ion conduction and used it in a lamp known as the 'Nernst glower'. Remarkably, 100 years later the same basic material, now known as yttria-stabilised zirconia, is the key solid electrolyte component currently in various solid oxide fuel cell and sensor applications.

Confirmation that ions can move in solids was obtained by Tubandt in 1910. He passed a direct current through the composite Ag/AgI/Ag at 150 °C and, by weighing the electrodes before and after the experiment, he showed that one electrode increased in weight, the other decreased in weight and, from the quantity of charge passed, the current was carried entirely by Ag^+ ions, which effectively transferred from one electrode to the other by migration through the AgI.

These early experiments on ionic conduction in solids were essentially laboratory curiosities since the phenomena were observed only at high temperatures. In the 1960s, several groups independently discovered solids that exhibited high ionic conductivity at room temperature; notably, these were the groups of Takahashi (various Ag compounds), Bradley and Green, Owens and Argue ($RbAg_4I_5$) and Yao and Kummer (β-aluminas).

In order for significant ionic conduction to occur in crystals, certain conditions must be satisfied:

(a) A large number of the ions of one species should be mobile (i.e. a large value of n in the equation $\sigma = ne\mu$).
(b) There should be a large number of empty sites available for the mobile ions to jump into. This is essentially a corollary of (a) because ions can be mobile only if there are empty sites available for them to occupy.
(c) The empty and occupied sites should have similar potential energies with a low activation barrier for jumping between neighbouring sites. It is no use having a large number of available empty sites if either the moving ion cannot get into them or if the sites are too small.
(d) The structure should have a framework, preferably 3D, permeated by open channels through which mobile ions may migrate.
(e) The anion framework should be highly polarisable.

β-Alumina meets the first four of these conditions, as does stabilised zirconia. The good Ag^+ ion conductors meet all five conditions. Many materials which are not good solid electrolytes may satisfy some of the conditions, but not all. Thus, many silicate zeolites have framework structures but the cations are trapped in relatively deep potential wells. Zeolites with their large cavities would appear to be obvious candidates for solid electrolyte behaviour but they have only moderate conductivities. The cations are usually present as hydrated complexes, which allows for ready ion exchange but not necessarily for high cation mobility. In dehydrated zeolites, the channels are too large and the cations tend to stick at sites in the channel walls. A similar effect occurs in Li^+ β-alumina. Poorly conducting β- and γ-AgI meet condition (e) but, more important, do not meet condition (c).

Conductivity data for a variety of solid electrolytes are given in Fig. 8.22 in the form of Arrhenius diagrams. Data for concentrated H_2SO_4, a typical liquid electrolyte, are shown for comparison. The target for improved solid electrolytes is therefore, the top right-hand corner of Fig. 8.22, i.e. high conductivity at low temperatures. Most materials are found in the bottom left-hand corner, i.e. low conductivities even at high temperatures! A selection of the more important solid electrolytes and their applications is discussed next.

8.5.4.2 β-Alumina

β-Alumina is the name for a family of phases of general formula $M_2O{\cdot}nX_2O_3$ where n is in the range 5–11, M is a monovalent cation [(alkali metal)$^+$, Cu^+, Ag^+, Ga^+, In^+, Tl^+, NH_4^+, H_3O^+) and X is a trivalent cation (Al^{3+}, Ga^{3+}, Fe^{3+}). The most important member is sodium β-alumina (M = Na^+, X = Al^{3+}), which has

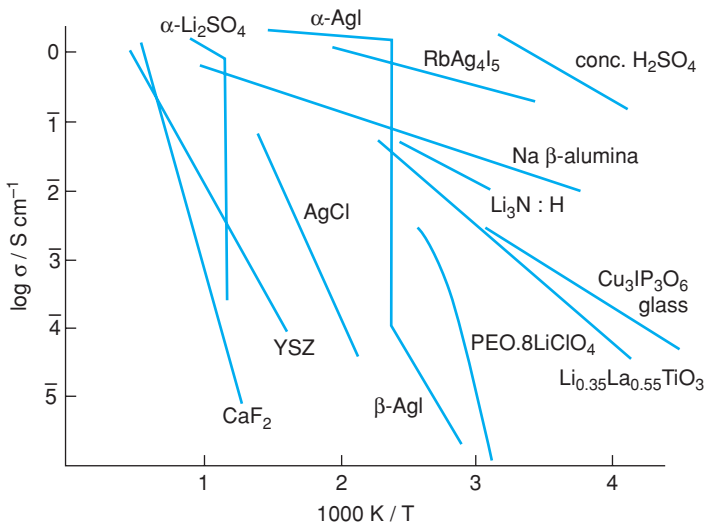


Figure 8.22 ***Ionic conductivity of some solid electrolytes with concentrated H_2SO_4 for comparison.***

been known for many years as a by-product of the glass-making industry: it forms in the refractory lining of furnaces by reaction of soda from the melt with alumina in the refractory bricks. Its name is a misnomer because, although originally thought to be a polymorph of Al_2O_3, it is now known that additional oxides such as Na_2O are essential to stabilise its crystal structure.

Interest in β-alumina as a solid electrolyte began with pioneering work by Yao and Kummer at the Ford Motor Co., who, in 1966, found that the Na^+ ions are very mobile at room temperature and above. They also found that other cations could ion exchange for Na^+ and these cations, too, are mobile. Since then, interest in solid electrolytes has mushroomed. In our energy-conscious society, much impetus for further research comes from the possibility of developing high-density energy storage systems such as the Na/β-alumina/S cell.

The crystal structures and conduction mechanism in β/β''-aluminas are well understood and provide a model system for understanding structure–property relations in solid electrolytes. For this reason, and for their importance in Na/S and related batteries, a detailed description is given next.

Crystal structures of β- and β''-aluminas

The high conductivity of the monovalent ions in β-alumina is a consequence of its unusual crystal structure, Fig. 8.23. It is built of *cp* layers of oxide ions, stacked in three dimensions, but every fifth layer has three-quarters of its oxygens missing. The Na^+ ions reside in these oxygen-deficient layers and are able to move easily because (a) there are more sites available than Na^+ ions to occupy them and (b) the radius of Na^+ is less than that of the O^{2-} ion. β-Alumina exists in two structural modifications, β and β'', which differ in the stacking sequence of the layers, Fig. 8.24. The β'' form occurs with more soda-rich crystals, $n = 5-7$, whereas β occurs for $n = 8-11$. Both the β and β'' structures are closely related to that of spinel, $MgAl_2O_4$, Fig. 1.44; Al^{3+} ions occupy a selection of tetrahedral and octahedral interstices between pairs of adjacent *cp* oxide layers. The β- and β''-alumina structures are built of 'spinel blocks' four oxide layers thick, in which the oxide layers are in the stacking sequence ABCA. Adjacent spinel blocks are separated by the oxygen-deficient layers or 'conduction planes' in which Na^+ ions reside. The unit cells are hexagonal with $a = 5.60$ Å and $c = 22.5$ Å (β), 33.8 Å (β''). In the c direction, perpendicular to the oxide layers, there are two

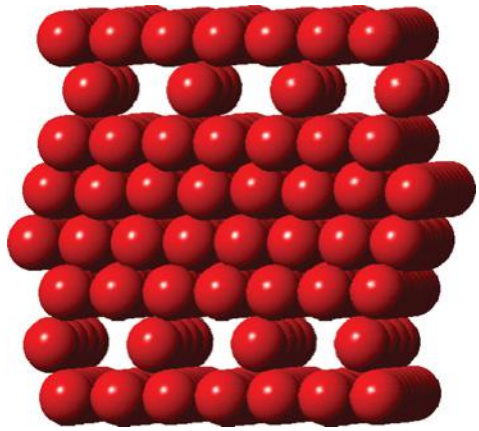


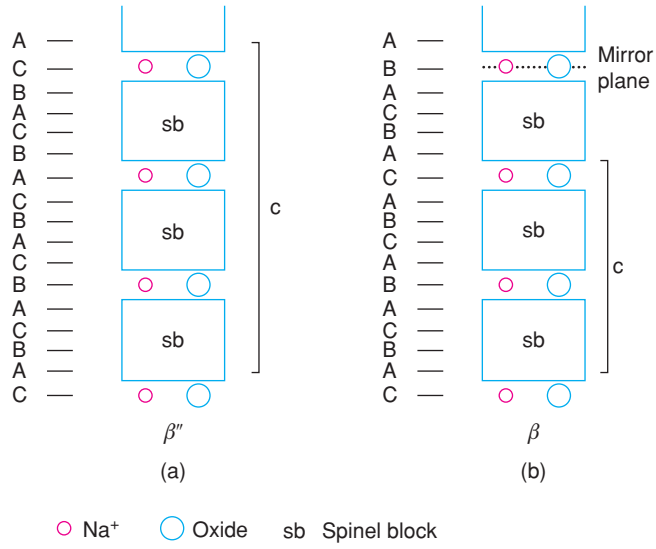
Figure 8.23 ***Oxide layers in β-alumina showing four-layer spinel blocks between adjacent conduction planes.***

spinel blocks in the unit cell of β and three in β''. The structure of the 'spinel blocks' is defective compared with the ideal spinel structure. Thus, spinel, $MgAl_2O_4$, contains both Mg^{2+} and Al^{3+} ions, in the ratio 1:2, but the spinel blocks of β- and β''-alumina contain only Al^{3+}, apart from small amounts of dopants such as Li^+, Mg^{2+} that are often added. For charge balance, Al^{3+} vacancies are also present in the spinel blocks. The overall stacking sequence of oxide ions including spinel blocks and conduction planes is cubic, ABC, in β'' but a more complex, 10-layer sequence in β, Fig. 8.24. In (c), the β structure is shown highlighting the tetrahedra and octahedra that form the spinel blocks and the spacer units of the conduction plane. The 10-layer repeat unit of the β structure arises because the conduction plane is a crystallographic mirror plane and the structure therefore is a mirror image to either side of the conduction plane.

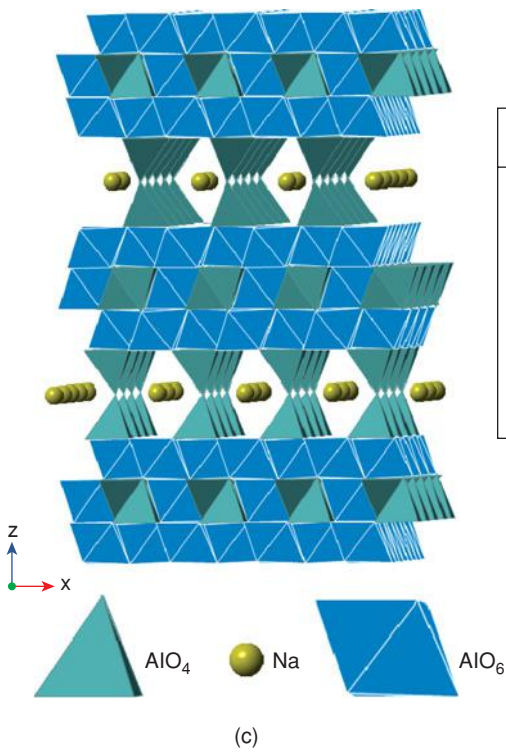
In Fig. 8.25, a layer of *cp* O^{2-} ions that forms one wall of the conduction plane is shown, and the 'column' or 'spacer' O^2 ions (shaded) of the conduction plane in the layer immediately above. In the conduction plane, only one-quarter of the available O^{2-} sites are occupied, i.e. for every shaded O^{2-} ion there are three empty sites, *m*. In the β modification, a mirror plane of symmetry passes through and parallel to each conduction plane (Fig. 8.24); hence the layers of O^{2-} ions on either side of the conduction plane are superposed in the projection in Fig. 8.25. In the β'' modification, the conduction plane does not coincide with a mirror plane and the oxide layers that form the two walls of the conduction plane are staggered relative to each other.

There are three possible sites for Na^+ ions in the β modification: (i) the 'mid-oxygen' positions, *m*; (ii) the 'Beevers–Ross sites', *br*, which were favoured in the original structure determination by Beevers and Ross; and (iii) the 'anti-Beevers–Ross sites', *abr*, Fig. 8.25. Na^+ ions spend most time in *br* and *m* sites, but in order to undergo any long-range migration they must pass through *abr* sites. Both *br* and *m* sites are large, e.g. Na^+ in a *br* site is coordinated to three oxygens in the oxide plane below, three in the plane above and three within the conduction plane. Na–O bond distances are large, ~2.8 Å, compared with more typical values of ~2.4 Å. The *abr* site is much smaller because two oxygens are fairly close, one directly above and one directly below, giving two short Na–O distances of 2.3 Å. Most other monovalent cations also prefer the *br* and *m* sites in β-alumina, with the exception of Ag^+ and Tl^+, which prefer *abr* sites. This is probably because Ag^+ and Tl^+ prefer covalent bonding and sites of low coordination number.

β-Alumina is a 2D conductor. Alkali metal ions move freely within the conduction planes but cannot penetrate the dense, spinel blocks. Because of their high mobility, the Na^+ ions can readily exchange with various cations, not only monovalent ions but also divalent ions such as Pb^{2+} and trivalent ions such as Eu^{3+}. The ion exchange is carried out by immersing crystals in a fused salt of the cation which replaces



Sodium Beta Alumina ($NaAl_{11}O_{17}$)



Atom	Wyckoff positions	x	y	z
Al1	2a	0	0	0
Al2	4f	0.3333	0.6667	0.0248(3)
Al3	12k	0.8333	0.1667	0.1062(2)
Al4	4f	0.3333	0.6667	0.1757(3)
Na	2d	0.6667	0.3333	0.25
O1	12k	0.1667	0.8333	0.0494(4)
O2	4f	0.6667	0.3333	0.0548(8)
O3	4e	0	0	0.1421(8)
O4	12k	0.5	0.5	0.1466(4)
O5	2c	0.3333	0.6667	0.25

Hexagonal : a = 5.602, c = 22.626 Å

Space Group : $P6_3$ / mmc (194)

Figure 8.24 ***Oxide packing in (a) β″- and (b) β-alumina; structure of the β polymorph showing (c) 3D arrangement of tetrahedra and octahedra with Na^+ ions in the conduction planes.***

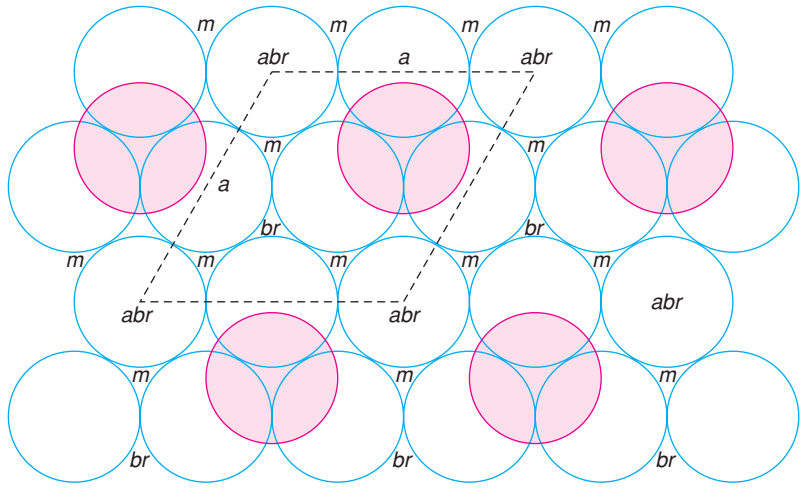


Figure 8.25 ***Conduction plane in β-alumina; the base of the hexagonal unit cell is shown dashed.***

Na^+. For example, Na^+ is replaced by K^+ on immersion in molten KNO_3 at 300–350 °C. The conductivity of ion-exchanged β-aluminas, Fig. 8.26, is highest and the activation energy lowest for Na^+ and Ag^+ β-alumina. With increasing cation size (K^+, Tl^+), the cations cannot move as readily. Ag^+ and Na^+ appear to have the optimum size because Li^+ β-alumina (not shown) also has a higher activation energy and lower conductivity than Na^+ or Ag^+ β-alumina. Li^+ is a small polarising cation and is not happy in large sites of high coordination number.

The conductivity data for the β-aluminas, Fig. 8.26, fit the Arrhenius equation very well over large ranges of temperature (up to 1000 °C) and conductivity (up to seven orders of magnitude). Contrast this with the NaCl data, Fig. 8.18, which show changes of slope several times over a range of ~400 °C. This simple behaviour in β-alumina is characteristic of solid electrolytes and also of many complex oxides, silicates, etc., which are not particularly good ionic conductors. It is remarkable that many materials with simple structures, such as NaCl, exhibit highly complex conductivity behaviour whereas materials with much more complex structures and stoichiometries, such as β-alumina, exhibit simple conductivity behaviour. This simple behaviour is observed independent of the magnitude of σ because β-alumina is equally well behaved at 300 °C, where $\sigma \approx 10^{-1}$ S cm^{-1}, or at −180 °C, where $\sigma \approx 10^{-8}$ S cm^{-1}.

In solid electrolytes, the σ values are usually reproducible between different samples and laboratories and are insensitive to small amounts of impurities. The conductivity of β-alumina has been measured in many laboratories and the data agree very well, with an activation energy of 0.16 ± 0.01 eV. This reproducibility occurs because (i) n is large in β-alumina (all Na^+ ions are potentially mobile) and therefore is not influenced by impurities and (ii) the conduction channels are well defined and not modified by lattice defects.

It is not feasible to apply defect equilibria considerations and the law of mass action to β-alumina, for several reasons. First, the equations that govern the concentration of Frenkel and Schottky defects apply only to very small defect concentrations (<0.1% of the lattice sites defective) whereas most, if not all, of the Na^+ ions in β-alumina are mobile. Second, the absence of doping effects in β-alumina indicates that separation of the conductivity activation energy into components due to the formation and migration of defects cannot be carried out readily and, indeed, it may be a mistake to try to make such a separation. The number of mobile ions is so large that they must be regarded as part of the normal structure rather than as defects and, as such, their number is insensitive to small amounts of impurity. Thus, whereas addition of 0.1% $MnCl_2$ to

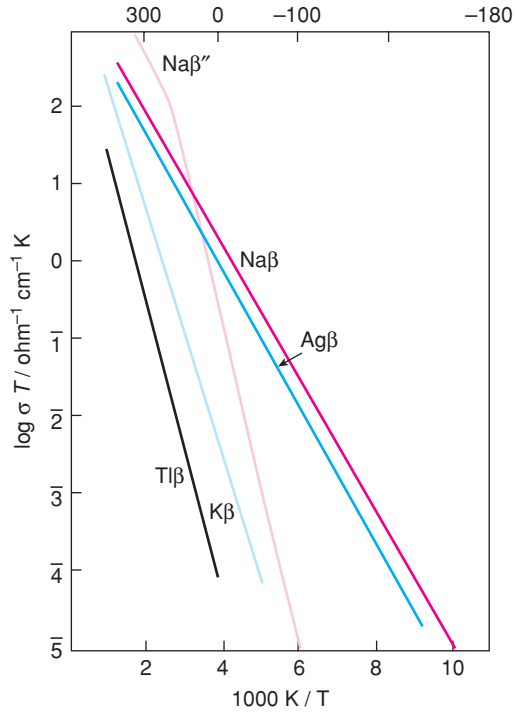


Figure 8.26 ***Conductivity of some single-crystal β- and β″-aluminas.***

NaCl may increase the concentration of cation vacancies by several orders of magnitude and, therefore, have a dramatic effect on σ, the corresponding effect of impurities on the conductivity of β-alumina is negligible. Small amounts of impurity may influence the long-term performance of β-alumina ceramics used in the Na/S battery, especially if they gradually segregate to grain boundaries or surfaces, but that is a different story.

The conduction pathways in β-alumina are shown in Fig. 8.25. The *br*, *abr* and *m* sites form a hexagonal network and, for long-range migration to occur, an Na^+ ion must pass through sites in the sequence *–br–m–abr–m–br–m–*. The activation energy for conduction, 0.16 eV, represents the overall value for migration of an Na^+ ion from, say, one *br* site to the next. Measurements of the Haven ratio indicate that a knock-on or interstitialcy mechanism is operative. In β-alumina crystals of composition $NaAl_{11}O_{17}$, the Na^+ ions occupy *br* sites; *abr* and *m* sites are unoccupied. If we allow one Na^+ ion to leave its *br* site, it must pass through an adjacent *abr* site; on the other side, it encounters an Na^+ ion on its own *br* site. There is not sufficient room for the moving Na^+ to squeeze past the Na^+ on the *br* site. Nor is there room for it to stop short in the nearest *m* site because it would be too close to the Na^+ ion on the *br* site. The moving Na^+ ion has two choices, therefore. It can return to its own *br* site, in which case no net conduction occurs, or it can knock the 'blocking' Na^+ ion out of its *br* site. If the latter occurs, the ejected Na^+ ion may simply move into one of the two adjacent *m* sites, in which case the incoming Na^+ ion remains in the third *m* site; the net result is that a *split interstitial* is created. Alternatively, by an interstitialcy mechanism, the ejected Na^+ ion may escape completely from its *br* and adjacent *m* sites, in which case a chain reaction begins. While the exact mechanistic details are a matter of speculation, it is clear that Na^+ ions cannot move independently of each other in β-alumina; instead, a cooperative process must occur.

An interesting effect observed in β-aluminas containing two alkali metal cations is the so-called *mixed alkali effect*. In this, the mobility of both alkali metal ions, e.g. Na^+ and K^+, is less than that in either single alkali metal end member: the conductivity passes through a minimum and the activation energy passes through a maximum for intermediate compositions. The mixed alkali effect is well known, but poorly understood, in glasses. Its occurrence and the absence of any satisfactory explanation for it shows how difficult it is to understand the nature of the forces that control ion mobility, even in a crystalline material whose structure is apparently well understood.

Much of the work on β-alumina has been directed towards maximising the conductivity of polycrystalline ceramics. Since the conductivity of the β'' form, Fig. 8.26, is several times higher than that of the β form at ~300 °C, which is the temperature of operation of the Na/S cell (Section 8.5.4.11), compositions are desired that both stabilise the β'' phase (it is not very stable in the Na_2O–Al_2O_3 system without additives) and maximise its content relative to that of the β phase. This is best achieved by doping with small amounts of Li_2O and MgO. The conductivity Arrhenius plot of β''-alumina is very different to that of β-alumina, Fig. 8.26. It has two linear regions separated by a region of curvature. In the high-temperature region, Na^+ ion motion is essentially liquid-like, $E \approx 0.05$ eV, but in the low-temperature region, the Na^+ ions prefer to occupy certain lattice sites to give an ordered superstructure. It is then much more difficult for Na^+ ions to move, as reflected in the larger E value, 0.35 eV.

The *magnetoplumbite* structure is closely related to that of β-alumina. It is mentioned here for completeness, although it is not an ionic conductor but has important magnetic applications in the material BAM, barium magnetoplumbite, $BaFe_{12}O_{19}$. The main difference from β-alumina is that the so-called conduction planes contain three oxygens and Ba instead of O and Na in the β-aluminas. Further details are given in Section 9.2.7.

8.5.4.3 Nasicon

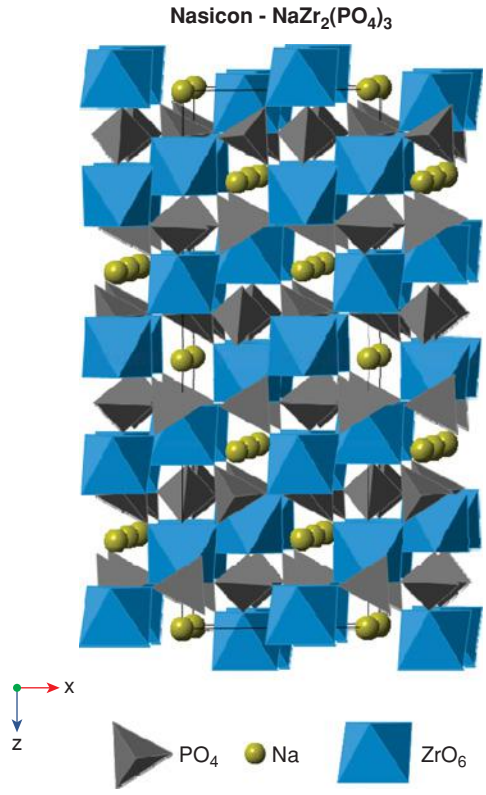
The name Nasicon, taken from Na Superionic Conductor, refers to a non-stoichiometric material with a framework built of ZrO_6 octahedra and $(Si,P)O_4$ tetrahedra which encloses a 3D network of interconnected pathways in which Na^+ ions reside. The Na sites are only partially occupied and easy 3D migration of Na^+ ions is possible. The stoichiometry, and hence the Na content, are controlled by adjusting the P:Si ratio to give the general formula, $Na_{1+x}Zr_2P_{3-x}Si_xO_{12}$, which is a solid solution in the system $NaZr_2(PO_4)_3$–$Na_4Zr_2(SiO_4)_3$. The Na^+ ion conductivity of the Nasicon solid solutions depends greatly on composition x and has a maximum conductivity at $x = 2$ which is comparable to that of β''-alumina at higher temperatures, >~150 °C. At lower temperatures, the structure undergoes a phase transition on cooling from rhombohedral to monoclinic symmetry, which has the effect of reducing the mobility of Na^+ ions and the conductivity at room temperature is only 1×10^{-4} S cm^{-1}.

The crystal structure of the Nasicon end-member, $NaZr_2(PO_4)_3$, is shown in Fig. 8.27(a). This end-member phase has low Na^+ ion conductivity since the Na^+ ions occupy fully one set of crystallographic sites. In Nasicon solid solutions, an additional set of interstitial sites is partially occupied by Na^+ ions and is responsible for the high ionic conductivity.

A wide range of other compositions form Nasicon-related structures. Thus, Zr can be replaced by Ti, V or Co. Na can be replaced by Li, giving an important family of Li^+ ion conductors (see later). The phase $Ca_{0.5}Ti_2(PO_4)_3$, CTP, is the parent structure of a family of materials which have zero, or low, coefficient of thermal expansion and therefore find applications as temperature-resistant ceramics.

8.5.4.4 Hollandites and priderites

Some materials in the hollandite and priderite family have exceptionally high 1D conductivity of Na^+ or K^+ ions. Hollandite is a mineral, $Ba_xMn_8O_{16}$, $x \leq 2$, which has a framework structure consisting of MnO_6



Atom	Wyckoff positions	x	y	z
Na	6b	0	0	0
Zr	12c	0	0	0.1456(1)
P	18e	0.2909(6)	0	0.25
O	36f	0.1860(15)	–0.0144(15)	0.0949(5)
O	36f	0.1913(15)	0.1683(15)	0.0866(5)

Trigonal : a = 8.8043, c = 22.7585 Å

Space Group : $R\bar{3}c$ (No. 167)

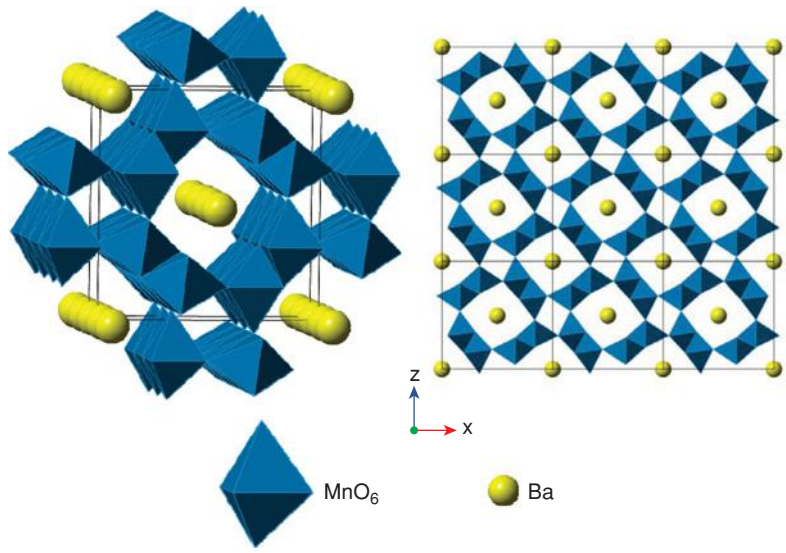
(a)

Figure 8.27 ***(a) Nasicon crystal structure of $NaZr_2(PO_4)_3$. Data taken from L. O. Hogman and P. Kierkegaard (1968),* Acta Chem. Scand.*, 22, 1822–1832. (b) Hollandite. Data taken from S. Ishiwata, J. W. G. Bos, Q. Huang and R. J. Cava (2006),* J. Phys. Condens. Matter*, 18, 3745–3752. (Continued)***

octahedra with Ba^{2+} ions located in channels within the framework, Fig. 8.27(b). The framework consists of double chains of edge-sharing MnO_6 octahedra which link at their corners to form channels parallel to *b*; these are where the large cations such as K^+ and Ba^{2+} reside. The unit cell is very short in the *b* direction, $b \approx 2.86$ Å, which means that the repeat unit in the *b* direction is a single MnO_6 octahedron or a single large cation.

Priderites are isostructural with hollandites but have more complex formulae such as $K_x(Mg,Ti)_8O_{16}$. K priderite has a very high conductivity, 1×10^{-1} S cm^{-1} at 293 K, which is observed only at high frequencies

Hollandite ($Ba_{1.2}Mn_8O_{16}$)



Atom	Wyckoff positions	x	y	z	f_{occ}
Ba	4g	0	0.382(4)	0	0.313(4)
Mn1	4i	0.161(1)	0	0.350(1)	1.0
Mn2	4i	0.347(1)	0	0.839(1)	1.0
O	4i	0.198(2)	0	0.152(2)	1.0
O	4i	0.162(3)	0	0.795(2)	1.0
O	4i	0.161(2)	0	0.542(3)	1.0
O	4i	0.550(2)	0	0.814(2)	1.0

Monoclinic : a = 10.052, b = 2.8579, c = 9.7627 Å, β = 89.96°

Space Group : I2/m (No. 12)

(b)

Figure 8.27 **(*Continued*)**

in the mega- to gigahertz region. At lower measuring frequencies, the conductivity is much less; it is believed that short range motions of K^+ ions are responsible for the conductivity at high frequency but long-range conduction is much more difficult owing to the small size of bottlenecks in the 1D conduction pathway. Long-range conduction in the 1D tunnels is also very sensitive to channel blocking by relatively immobile ions such as stray impurities.

8.5.4.5 Silver and copper ion conductors

AgI undergoes a phase transition at 146 °C and the high-temperature form, α-AgI, has exceptionally high conductivity, ~1 S cm^{-1}, which is four orders of magnitude larger than that of β-AgI at room temperature, Fig. 8.22. The activation energy for conduction in α-AgI is only 0.05 eV and its structure is so suited for easy motion of Ag^+ that the ionic conductivity actually *decreases* slightly on melting at 555 °C. β-AgI has the

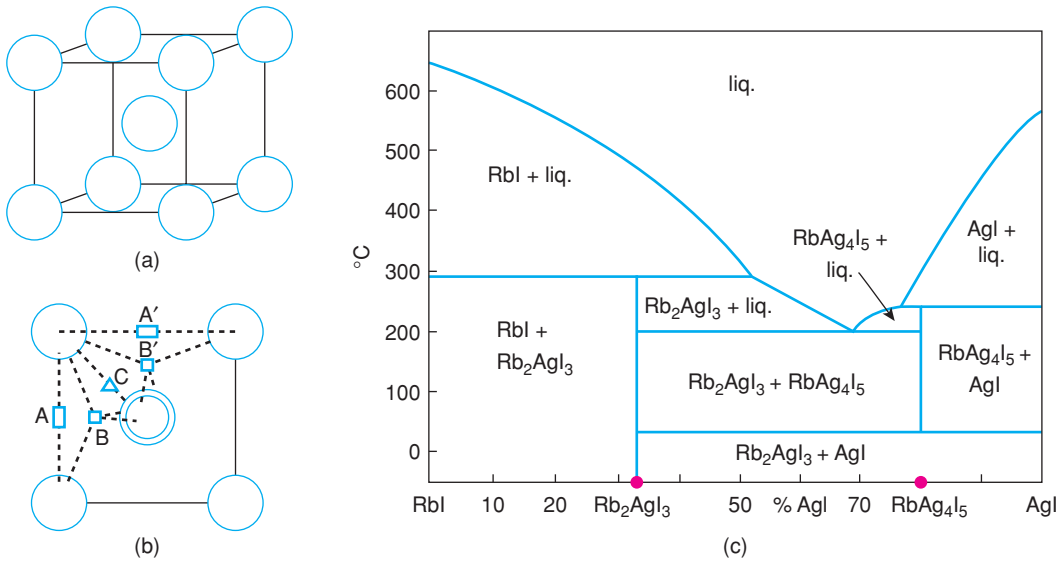


Figure 8.28 ***(a) Crystal structure of α-AgI showing* bcc *arrangement of I^- ions and (b) some of the sites available for Ag^+ ions. (c) Phase diagram for AgI–RbI. Reproduced with permission from T. Takahashi, Solid silver ion conductors,* J. Appl. Electrochem.*, 3, 79 © 1973, Chapman and Hall Ltd.***

wurtzite structure with *hcp* I^- ions and Ag^+ in tetrahedral sites. α-AgI is *bcc*; I^- ions lie at corner and body centre positions, Fig. 8.28(a), and Ag^+ ions are distributed statistically over a total of 36 sites of tetrahedral, trigonal and linear coordination with overall partial occupancy (b). The tetrahedral sites, B, B′, link by sharing common faces, i.e. the trigonal sites C.

Ag^+ ions can readily move from one site to the next in a liquid-like manner. They can also move between tetrahedral sites, B, in adjacent unit cells via the linear sites, A. This easy motion of Ag^+ is related to the nature of the bonding between Ag and I. Ag^+ is a polarising cation since its outer 4*d* electrons are ineffective in shielding nuclear charge. I^- is a large, polarisable anion and so covalent bonds readily form between Ag^+ and I^- that are characterised by structures with low coordination numbers. During conduction, Ag can readily move from one tetrahedral site to the next through the intermediate, trigonal or linear sites; covalent bonding at the intermediate sites helps to stabilise them and reduce the activation energy for conduction. It is interesting that AgCl and AgBr have reasonable conductivities at high temperatures, Fig. 8.22, but much less than α-AgI. The bonding in these is less covalent than in AgI. However, they also have a different crystal structure to α-AgI (they have the rock salt structure; Section 8.5.2) and this is also undoubtedly an important factor.

In attempts to stabilise the highly conducting α-AgI phase at lower temperatures, various anionic and cationic substitutions were tried. The most successful involves partial replacement of Ag by Rb in $RbAg_4I_5$, which has the highest ionic conductivity at room temperature of any known crystalline substance, 0.25 S cm^{-1}. Its activation energy for conduction is 0.07 eV, Fig. 8.22. The amount of electronic conductivity in $RbAg_4I_5$ is negligibly small, $\sim 10^{-9}$ S cm^{-1} at 25 °C.

The RbI–AgI phase diagram is shown in Fig. 8.28(c). Two binary compounds exist, Rb_2AgI_3 and $RbAg_4I_5$. The latter melts incongruently to AgI and liquid at ~230 °C. It has a lower limit of stability at 27 °C and decomposes to AgI and Rb_2AgI_3, but with care, $RbAg_4I_5$ may be cooled below room temperature without decomposing.

The crystal structure of $RbAg_4I_5$ is rather different to that of α-AgI, but also contains Ag^+ ions distributed over a network of face-sharing tetrahedral sites. Again, there are many more available sites than Ag^+ ions to fill them. Rb^+ ions are immobilised in sites that have a distorted octahedral environment of I^- ions. Conductivity data for $RbAg_4I_5$, Fig. 8.22, fall on a smooth curve that shows a small departure from linear Arrhenius behaviour.

A similar disordered, α-AgI-like structure can be stabilised at low temperatures by a variety of cations, notably large alkali metals, NH_4^+, substituted NH_4^+ and certain organic cations. Some examples, all of which have conductivities at 25 °C in the range 0.02–0.20 S cm^{-1}, are $[(CH_3)_4N]_2Ag_{13}I_{15}$, $PyAg_5I_6$ where Py is the pyridinium ion, $(C_5H_5NH)^+$, and $(NH_4)Ag_4I_5$.

A range of anions may substitute partially for I and some of the phases formed have high conductivities, e.g. Ag_3SI, $Ag_7I_4PO_4$ and $Ag_6I_4WO_4$. The latter mixed iodide/oxysalt phases have considerable thermal stability, are unaffected by iodine vapour or moisture and are suitable for use as the solid electrolyte in electrochemical cells. An interesting development is the preparation of glassy solid electrolytes in which molten mixtures of AgI with salts such as Ag_2SeO_4, Ag_3AsO_4 and $Ag_2Cr_2O_7$ can be preserved to room temperature, as glasses, by rapid quenching. The Ag^+ ion conductivities are again very high, e.g. 0.01 S cm^{-1}, for $Ag_7I_4AsO_4$. Obviously, it is not possible to know the detailed structure of the iodide and oxysalt framework in glasses such as these, but it seems reasonable to assume that Ag can again move through a network of face-sharing tetrahedra.

It is also possible to replace simultaneously both Ag^+ and I^-, e.g.

(a) Replace Ag^+ by Hg^{2+} and I^- by Se^{2-} in $Ag_{1.8}Hg_{0.45}Se_{0.7}I_{1.3}$.
(b) Replace Ag^+ by Rb^+ and I^- by CN^- in $RbAg_4I_4CN$.

These AgI derivative phases have high Ag^+ ion mobilities but very small electronic conductivities. They may be used as solid electrolytes without danger of short-circuit through the electrolyte due to electronic conductivity. By contrast, a very desirable if not essential attribute of solid *electrode* materials is that they should have high ionic *and* electronic conductivities. Silver chalcogenides, such as α-Ag_2S exhibit mixed ionic/electronic conductivity and it is possible also to stabilise the high-temperature, disordered phase at lower temperatures by making various substitutions. For example, α-Ag_2Se can dissolve 5–10% Ag_3PO_4 in solid solution formation and the α-phase is then stable to room temperature, where it has an ionic conductivity of 0.13 S cm^{-1} and an electronic conductivity of between 10^4 and 10^5 S cm^{-1}. Mixed conductors are discussed further in Section 8.5.4.10.

Since the chemistries of monovalent Ag and Cu are similar, it is not surprising that a high conductivity of Cu^+ ions occurs in a number of phases similar to those of the Ag compounds, although most conductivities do not achieve the level of conductivity shown by the best Ag^+ ion conductors. CuI undergoes a similar phase transition to AgI at 430 °C and the high-temperature polymorph is a very good Cu^+ ion conductor. The $\beta \rightarrow \alpha$ transition in CuI is also associated with a high increase in entropy and the Cu^+ ions in α-CuI are disordered over partially occupied sites giving an essentially liquid-like sublattice. This disordered sublattice can be stabilised in numerous complex halides such as $Rb_3Cu_7Cl_{10}$ and $Rb_4Cu_{16}I_{17}Cl_{13}$, which have very high Cu^+ ion conductivities.

8.5.4.6 Fluoride ion conductors

PbF_2, which has the fluorite (CaF_2) structure, has high F^- conductivity, especially at high temperatures; its high conductivity was discovered by Faraday and so it has a strong claim to being the first known solid electrolyte. At room temperature, PbF_2 is a typical ionic solid with low conductivity. With rising temperature its conductivity increases smoothly and rapidly until at ~500 °C a limiting value of ~5 S cm^{-1} is reached, Fig. 8.29 (a). Above 500 °C, σ increases only slowly and there is little change in σ on melting at 822 °C. It

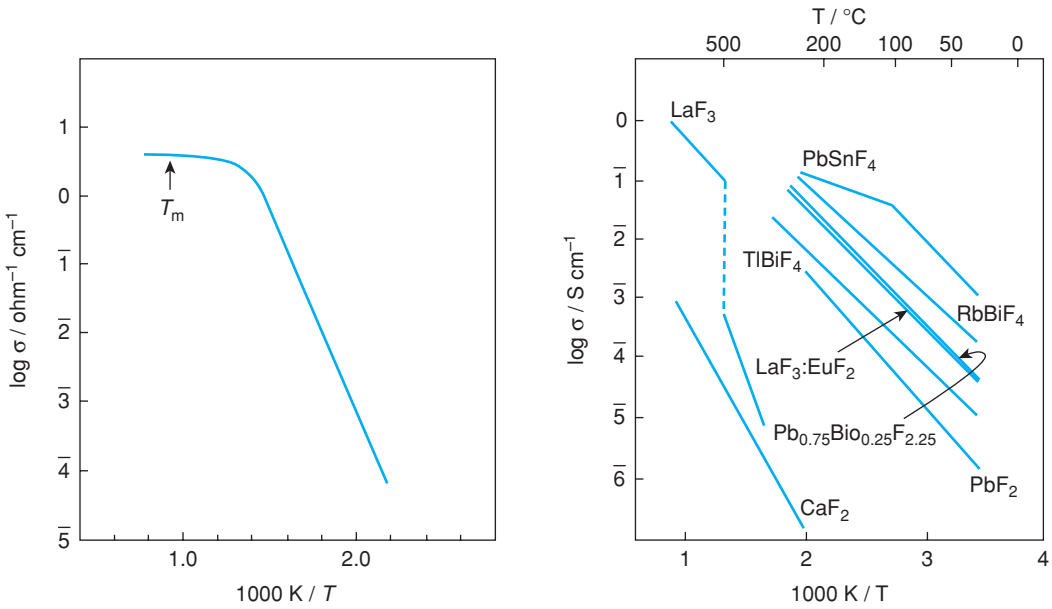


Figure 8.29 ***(a) Conductivity of PbF_2 as a function of reciprocal temperature. Reproduced with permission from Derrington and O'Keeffe,*** **Nature Phys. Sci.,** *246, 44 © 1973 Nature Publishing Group. (b) Conductivity of a selection of fluoride ion conductors. Adapted with permission from T. Kudo 'Survey of types of solid electrolytes' in* **The CRC Handbook of Solid State Electrochemistry,** *Ed. P. J. Gellings and H. J. M. Bouwmeester, © 1997 CRC Press.***

is interesting that PbF_2 arrives in the highly conducting condition on gradually increasing the temperature, whereas AgI does so by undergoing a first-order phase transition and an abrupt change in crystal structure, Fig. 8.22. $SrCl_2$ behaves like PbF_2 and has high σ between ~700 °C and the melting point 873 °C; CaF_2 becomes highly conducting just as the melting point, 1418 °C is approached. Some complex fluorides such as $PbSnF_4$ have high F^- ion conductivity at lower temperatures than PbF_2.

Conductivity data for a selection of fluorides are shown in Fig. 8.29(b). Most of the good fluoride ion conductors have either the simple fluorite structure, or a more complex fluorite structure with ordering of two cations or a fluorite-structured solid solution containing interstitial F^- ions such as $Pb_{1-x}Bi_xF_{2+x}$. LaF_3 with the tysonite crystal structure is a good fluoride ion conductor, especially at high temperatures; its conductivity is increased by the introduction of fluoride vacancies, such as in the solid solution $La_{1-x}Eu_xF_{3-x}$.

The fluorite structure may be described in various ways, one of which is as primitive cubes of F^- ions with Ca^{2+} ions at the body centres of alternate cubes, Fig. 1.34. The sites available for interstitial F^- are at the centres of the empty cubes; these sites are coordinated octahedrally by six Ca^{2+} and by eight F^- ions at the cube corners. In creating an interstitial F^- ion, one corner F^- ion must move off its corner site. The true defect structure is inevitably more complex since structural relaxation involving small displacement of ions occurs in the vicinity of the defects; larger defect complexes form, such as described earlier for $Fe_{1-x}O$, Fig. 2.9, and UO_{2+x}, Fig. 2.10.

8.5.4.7 Oxide ion conductors

Oxide ion conductors have important applications in O_2 sensors, pumps and solid oxide fuel cells, SOFCs. Because of their diverse applications, there is much interest in such materials; a selection is shown in Fig. 8.30.

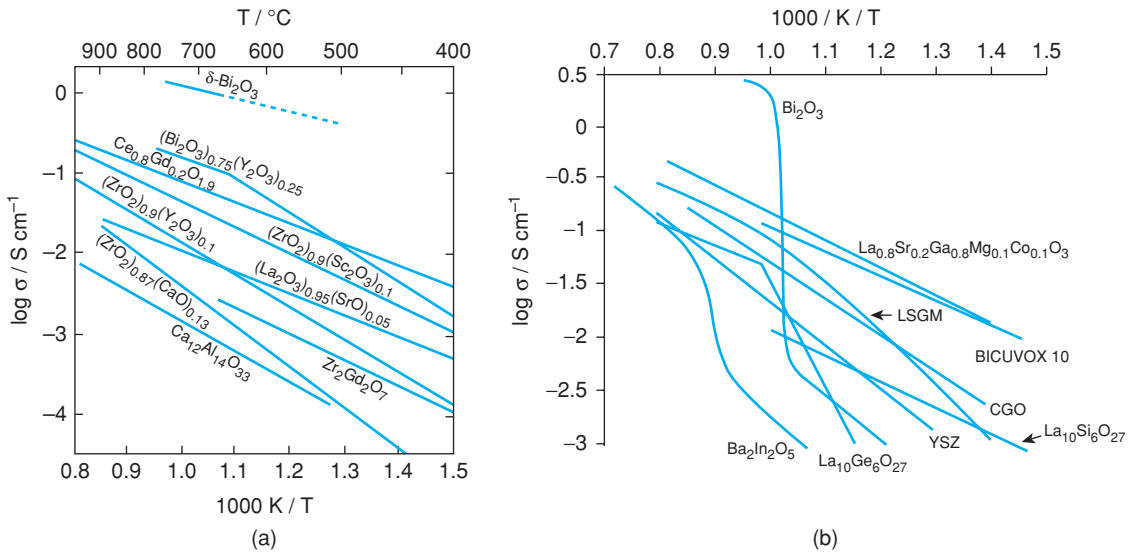


Figure 8.30 ***(a) Some oxide ion conductors. (b) Conductivity data for selected oxygen ion conductors (BICUVOX.10 = $Bi_2V_{0.9}Cu_{0.1}O_{5.5-x}$; LSGM = $La_{0.9}Sr_{0.1}Ga_{0.8}Mg_{0.2}O_{2.85}$; YSZ = $Y_{0.15}Zr_{0.85}O_{1.925}$; CGO = $Ce_{0.9}Gd_{0.1}O_{2-x}$). Reproduced with permission from Slater et al.,* Chem. Record, *© 2004 The Japan Chemical Journal Forum and Wiley Periodicals, Inc.***

The best-known oxide ion conductor is yttria-stabilised zirconia, YSZ. The high-temperature, cubic polymorph of zirconia has the fluorite crystal structure and may be stabilised to room temperature by formation of solid solutions with CaO, Y_2O_3, etc. Such 'stabilised zirconias' are good O^{2-} ion conductors at high temperatures, mainly because the mechanism of solid solution formation involves the creation of vacant O^{2-} sites in order to preserve electroneutrality. Lime-stabilised zirconia has the formula $(Ca_xZr_{1-x})O_{2-x}$: $0.1 \leq x \leq 0.2$; with each Ca^{2+} ion that is introduced 1 O^{2-} ion vacancy is created. YSZ has the formula $(Y_xZr_{1-x})O_{2-x/2}$. These solid solution mechanisms are discussed in Sections 2.3.3.1 (3) and 2.3.5.2 and the phase diagram ZrO_2–Y_2O_3 is given in Fig. 7.28.

Typical conductivities in stabilised zirconias are in the range 10^{-1}–10^{-2} S cm^{-1} at 800 °C with activation energy for conduction in the range 0.8–1.3 eV, Fig. 8.30. At lower temperatures, their conductivities are much lower than those of the good Na^+ and Ag^+ ion solid electrolytes. Their usefulness stems from the fact that they are refractory materials, can be used at very high temperatures, e.g. 1500 °C, and have good oxide ion conductivity, which is an unusual property. ThO_2 and HfO_2 may also be doped and are good O^{2-} ion conductors similar to ZrO_2. Doped CeO_2 is an even better oxide ion conductor but is susceptible to reduction, giving a mixture of Ce^{4+} and Ce^{3+} oxidation states and mixed conductivity of oxide ions and electrons.

The oxide ion conductivity of the fluorite-type solid solutions is associated with the high concentration of oxygen vacancies and the conductivity usually passes through a maximum in the range 5–10% dopant. The oxygen vacancies appear not to be entirely free since they are trapped by the dopant cations to form dipoles and larger defect clusters. This is because the oxygen vacancies have a net double positive charge, represented in Kroger–Vink notation as $V_O^{\bullet\bullet}$, whereas substitutional dopants carry a net negative charge such as Ca''_{Zr} and Y'_{Zr}. There is a natural tendency for oppositely charged species to attract each other and in order for the vacancies to break free, an activation energy equal to the enthalpy of formation of the defect clusters

must be overcome. A feature of some oxide ion conductors is that the defect clusters form over a period of time as the oxygen vacancies become gradually trapped, as shown schematically by

$$V_O^{\bullet\bullet} + Ca_{Zr}'' \rightarrow (Ca_{Zr} - V_O)^x \tag{8.34}$$

and the conductivity therefore gradually decreases, an effect known as *ageing*. With increasing dopant concentration, there is an increasing tendency for defect clustering to occur, which may eventually lead to precipitation of a low-conductivity phase in which the oxygen vacancies are ordered and therefore completely trapped. Thus, in lime-stabilised zirconia, the phase $CaZr_2O_5$ has an anion vacancy ordered structure; similarly, Y_2O_3 in YSZ may be regarded as an oxygen vacancy ordered, C-type rare earth oxide structure.

CeO_2-based solid solutions such as $Ce_{1-x}Gd_xO_{2-x/2}$ have much higher conductivity than stabilised zirconias, especially at lower temperatures, because of their lower activation energy, but suffer from ease of reduction of Ce^{4+} to Ce^{3+}. This is associated with a small amount of oxygen loss at high temperatures in atmospheres of low p_{O_2} to give materials with a more general formula $Ce_{1-x}Gd_xO_{2-x/2-\delta}$. Oxide ions are lost from the lattice, forming O_2 molecules and injecting electrons into the structure, as shown by the following equilibrium:

$$O_O^x \leftrightarrow {}^1/_2O_2 + V_o^{\bullet\bullet} + 2e' \tag{8.35}$$

The law of mass action may be applied to this equilibrium giving:

$$K = \frac{(p_{O_2})^{\frac{1}{2}}\left[V_O^{\bullet\bullet}\right]\left[e'\right]^2}{\left[O_O^x\right]} \tag{8.36}$$

The main variables are p_{O_2} and electron concentration, since the concentration of lattice oxygens is unity and the concentration of oxide ion vacancies is essentially constant for a large dopant concentration. Therefore:

$$\left[e'\right] \propto K^{\frac{1}{2}}\left(p_{O_2}\right)^{-\frac{1}{4}} \tag{8.37}$$

Thus, although doped cerias are good oxide ion conductors, at low p_{O_2} they increasingly become electronic conductors as well.

In a mixed conductor which exhibits both ionic and electronic conduction, the ionic transport number, t_i, is given by

$$t_i = \frac{\sigma_i}{\sigma_i + \sigma_e} \tag{8.38}$$

A useful parameter to characterise the ease of reduction of oxide ion conductors is the oxygen partial pressure, $p_{O_2}{}^*$, at which $t_i = {}^1/_2$ since

$$t_i = \frac{1}{2} = \left[1 + \left(p_{O_2}/p_{O_2}{}^*\right)\right]^{-1} \tag{8.39}$$

The dependence of conductivity on p_{O_2} is shown schematically in Fig. 8.31. At low p_{O_2} the conductivity increases with a slope of –1/4 on a logarithmic scale and the material is an n-type conductor. At higher p_{O_2}, where $t_i = 1$, conductivity is independent of p_{O_2} and the material is said to be in the *electrolytic domain*. At still higher p_{O_2}, p-type behaviour may be observed, which is the mirror image of the n-type dependence of conductivity on p_{O_2}. In this region, the sample absorbs oxygen either at the sample surface or by diffusion

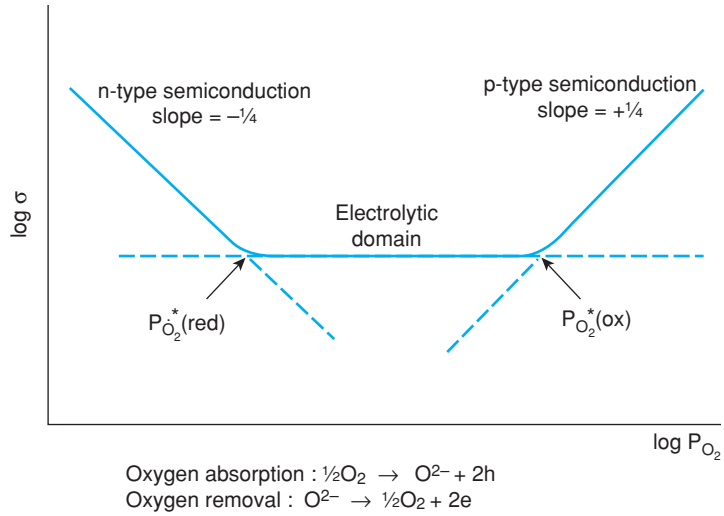


Figure 8.31 ***Ionic and electronic conductivity domains as a function of oxygen partial pressure.***

into vacant anion sites; the absorbed oxygens dissociate and pick up electrons, as shown by the reverse of the equilibrium in equation (8.35):

$$\tfrac{1}{2}O_2 + V_O^{\bullet\bullet} + 2e' \rightarrow O_O^x \tag{8.40}$$

which may alternatively be written as

$$\tfrac{1}{2}O_2 + V_O^{\bullet\bullet} \rightarrow O_O^x + 2h^{\bullet} \tag{8.41}$$

When the holes that are created by this process become the dominant charge carriers, the conductivity dependence on p_{O_2} has a slope of +1/4, Fig. 8.31.

There are two solutions to equation (8.39), which represent the transition from oxide ion conduction to n-type conduction, $p_{O_2}{}^*$(red) and the transition from oxide ion conduction to p-type conduction at $p_{O_2}{}^*$(ox), Fig. 8.31. Values of $p_{O_2}{}^*$(red) are of particular interest in assessing the stability of oxide ion conductors in a reducing atmosphere and therefore their potential applications as an electrolyte in a solid oxide fuel cell. Some data for ceria-based materials are summarised in Table 8.6, from which it can be seen that zirconia-based materials are much more resistive to reduction than ceria-based materials.

The material with the highest oxide ion conductivity at high temperatures is Bi_2O_3, Fig. 8.30. It has a fluorite-related structure but with 25% of the oxide ion sites vacant and, consequently, a highly disordered anion arrangement. Below 729 °C, cubic δ-Bi_2O_3 transforms to monoclinic α-Bi_2O_3 with a vacancy-ordered crystal structure and much lower conductivity. Bi_2O_3 forms extensive solid solutions with oxides such as Y_2O_3, Gd_2O_3 and WO_3, some of which retain the high oxide ion conductivity of δ-Bi_2O_3 to lower temperatures. These materials are, however, easily reduced in low-p_{O_2} atmospheres, which prevents their use as solid electrolytes in SOFC applications.

Several other crystal structures are now known to be good host structures for high levels of oxide ion conductivity at high temperatures. All are characterised by an oxide ion sublattice which has a considerable vacancy concentration. This includes pyrochlores such as $Zr_2Gd_2O_7$, a calcium aluminate, $Ca_{12}A_{14}O_{33}$, (a minor

Table 8.6 ***Some oxygen partial pressures at which*** $t_{O^{2-}} = \frac{1}{2}$ ***and*** $\sigma_{O^{2-}} = \sigma_e$

	p_{O_2}*/atm			
Composition	500 °C	700 °C	850 °C	1000 °C
$Ce_{0.9}Ca_{0.1}O_{1.9}$	10^{-20}	10^{-13}	10^{-9}	
$Ce_{0.905}Y_{0.095}O_{1.95}$	10^{-26}	10^{-17}	10^{-13}	10^{-10}
$Ce_{0.9}Gd_{0.1}O_{1.95}$		1.2×10^{-19}	1.7×10^{-15}	
$Ce_{0.5}Gd_{0.5}O_{1.75}$		1.5×10^{-16}	6.5×10^{-14}	
$Zr_{1-x}Ca_xO_{2-x}$		10^{-34}	10^{-30}	10^{-26}

Source: Data reproduced with permission from T. Kudo and K. Fueki (1990) *Solid State Ionics*, VCH, p. 133.

phase in bags of cement!), the perovskite $(La_{1-x}Sr_x)(Ga_{1-x}Mg_x)O_{3-x}$ and apatites $Ln_{9.33+x}(SiO_4)_6O_{2+3x/2}$: Ln = rare earth. Conductivity data for a selection of these phases are shown in Fig. 8.30(b).

8.5.4.8 Li^+ ion conductors

Materials that have high conductivity of Li^+ ions form the electrolyte in high-density Li batteries. There is enormous world-wide interest in developing advanced Li batteries for both small-scale applications, in mobile telephones, video recorders and laptop computers and on a large scale for transport applications in zero emission vehicles (ZEVs). The interest arises because cells containing Li anodes generally have higher emf than corresponding cells containing, e.g., Na anodes. Commercial Li batteries currently have 4 V single cells with an Li–carbon anode and an intercalation cathode based on Li_xCoO_2. New materials and improved systems are continuously being developed and the current record for the emf of a single cell is 5.0 V; this contains an intercalation cathode based on spinels such as $LiCoMnO_4$ and $Li_2NiMn_3O_8$.

Here we are concerned with crystalline Li^+ ion conductors; as yet, none are good enough to replace non-aqueous electrolytes such as $LiPF_6$ dissolved in an organic solvent, e.g. propylene carbonate, in Li batteries, but there is much interest in developing all-solid-state systems, avoiding liquid containment problems. Conductivity data for some of the most important Li^+ ion conductors are given in Fig. 8.22.

Li_2SO_4 undergoes a phase transition at 572 °C and has $\sigma \approx 1$ S cm^{-1} above this temperature. Many substituted lithium sulfates have been studied in attempts to reduce the temperature of the phase transition and hence preserve the highly conducting α polymorph to lower temperatures, but with limited success, and the α polymorph cannot be stabilised to room temperature. α-Li_2SO_4 is, nevertheless, of continuing scientific interest because, unusually, the SO_4^{2-} ions are able to rotate in the crystal lattice (it is sometimes known as a 'rotator phase'). An interesting question is whether the rotating SO_4^{2-} ions facilitate conduction of Li^+ ions by a 'paddle-wheel' mechanism or whether the two processes, each of which is rare in solids, are unconnected phenomena. The jury is still out!

Li_4SiO_4 (and Li_4GeO_4) is a modest Li^+ ion conductor but, when doped to create either Li^+ vacancies or Li^+ interstitials, dramatic increases in conductivity occur. Examples are the following:

$$Li^+ + Si^{4+} \rightarrow V^{5+} \text{ to give } Li_{4-x}Si_{1-x}V_xO_4$$

$$3Li^+ \rightarrow Al^{3+} \text{ in } Li_{4-3x}Al_xSiO_4$$

$$Si^{4+} \rightarrow Li^+ + Al^{3+} \text{ in } Li_{4+x}Si_{1-x}Al_xO_4$$

Li_4SiO_4 is built of SiO_4 tetrahedra and the Li^+ ions are distributed through a network of face-sharing polyhedral sites. In stoichiometric Li_4SiO_4, several of these sites are full and the remainder are empty;

hence the mobile ion concentration, n, is very low and the conductivity is low. In the doped materials, the conductivity rises as the mobile ion concentration (vacancies or interstitials) increases.

The Li_4SiO_4 solid solutions containing Al provide a particularly clear example of the importance of mobile ion concentration, Fig. 8.32(a). Solid solutions are very extensive, with Li^+ contents ranging from 4.5 (in $Li_{4.5}Si_{0.5}Al_{0.5}O_4$) to 2.3 (in $Li_{2.29}Al_{0.57}SiO_4$). Conductivity isotherms at, e.g., 100 °C, Fig. 8.32(a), show conductivity minima at Li_4SiO_4 and $Li_{2.5}Al_{0.5}SiO_4$ with maxima at intermediate and extreme compositions.

The structures of Li_4SiO_4, Fig. 8.32(b) and $Li_{2.5}Al_{0.5}SiO_4$ (c) differ in that certain sets of sites that are fully occupied by Li in Li_4SiO_4 are completely empty in $Li_{2.5}Al_{0.5}SiO_4$, thus accounting for the insulating nature of $Li_{2.5}Al_{0.5}SiO_4$. These same sites are partially occupied at intermediate compositions and the conductivity passes through a maximum around the composition, $Li_{3.25}Al_{0.75}SiO_4$. For compositions with Li content less than 2.5, some of the framework sites occupied by Li in $Li_{2.5}Al_{0.5}SiO_4$ become empty, and a vacancy conduction mechanism operates with an increase in conductivity; by contrast, for Li contents greater than 4.0, an additional set of interstitial sites begins to fill up with Li^+ ions and the conductivity increases again. Al-doped Li_4SiO_4 is very unusual since the parent structure is retained over such a wide composition range that the conduction mechanism changes several times and leads to the dramatic variations in conductivity shown in (a). Almost all other doped materials have more limited compositional ranges and the conductivity variations are limited to a single type.

A family of Li^+ ion conductors with similar structures to doped Li_4SiO_4 are the so-called *lisicons* (from Li Superionic Conductors). These are interstitial solid solutions such as

$$P^{5+} \rightarrow Li^+ + Si^{4+} \text{ in } Li_{3+x}P_{1-x}Si_xO_4$$

$$Zn^{2+} \rightarrow 2Li^+ \text{ in } Li_{2+2x}Zn_{1-x}SiO_4$$

The highest conductivities are, e.g., 10^{-1} S cm^{-1} at 300 °C in $Li_{3.5}Zn_{0.25}GeO_4$ and 5×10^{-5} S cm^{-1} at 25 °C in $Li_{3.5}V_{0.5}GeO_4$. One problem encountered with lisicons is defect cluster formation and the trapping of mobile species with consequent decrease in conductivity, similar to that described above with stabilised zirconias and aliovalent-doped alkali halides. Thus, in $Li_{2+2x}Zn_{1-x}SiO_4$, the two Li^+ ions that substitute for each Zn^{2+} ion have opposite site charges, i.e. $Li_i^{\bullet}$ and Li'_{Zn}, since one occupies an empty interstitial site and the other substitutes onto a Zn site. At high temperatures, the defect clusters dissociate; on cooling, defect complexes gradually reform and the conductivity decreases. For practical applications, therefore, materials are required that do not exhibit such ageing effects.

More recently, a family of *thiolisicons* has been discovered by Kanno in which O is replaced by S. These have similar crystal structures and solid solution mechanisms to lisicons, have higher ionic conductivity with σ as high as 1×10^{-2} S cm^{-1} at 25 °C, and are being considered for use as the solid electrolyte in all-solid-state lithium batteries. The higher conductivity can be explained by a combination of more covalent bonding, with 'soft' S compared with 'hard' O and a more expanded crystal lattice due to the larger size of S.

Lithium nitride, Li_3N, has exceptionally high conductivity in two dimensions, Fig. 8.22. It has a layered structure, Fig. 8.32(d); sheets of stoichiometry 'Li_2N' alternate with layers of Li. Conductivity occurs primarily in the 'Li_2N' sheets by an Li vacancy mechanism. The conductivity of H-containing Li_3N of formula $Li_{3-x}H_xN$ is higher than that of pure Li_3N since the substitutional H is tightly bonded to N, forming NH units, and leaving Li sites vacant in the 'Li_2N' sheets. Various lithium nitride halide materials are known and some, such as $Li_{3-x}NH_xCl_x$: $x = 0.6$, have Li-deficient antifluorite structures and exhibit high Li^+ ion conductivity.

A family of Li-containing perovskites with unusual structural features and high Li^+ ion conductivity, 1×10^{-3} S cm^{-1} at 25 °C, is based on $Li_{0.5}La_{0.5}TiO_3$, which forms Li-deficient solid solutions:

$$3Li^+ \rightarrow La^{3+} \text{ in } Li_{0.5-3x}La_{0.5+x}TiO_3$$

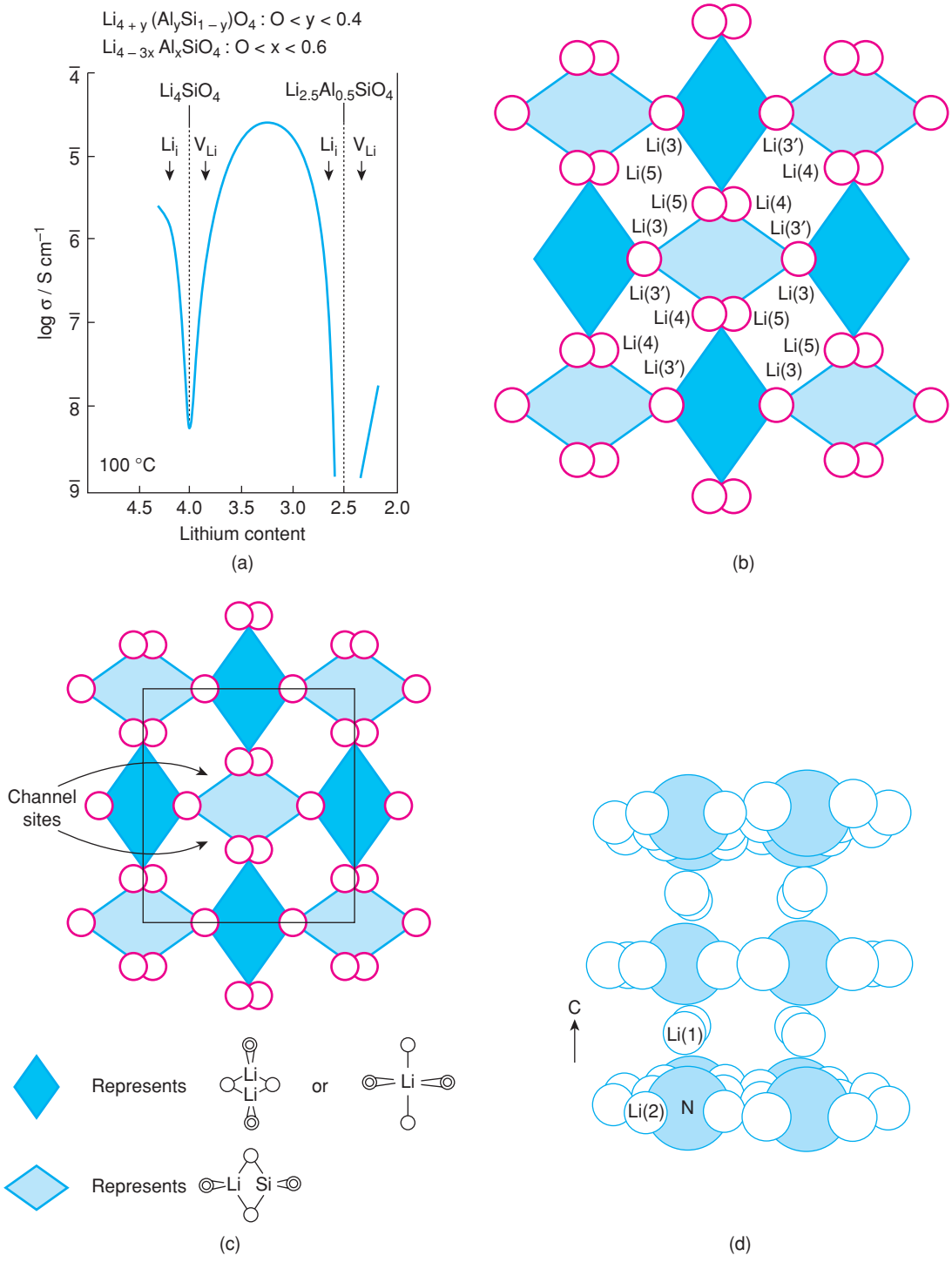


Figure 8.32 *(a) Conductivity at 100 °C of solid solutions based on Li_4SiO_4, as a function of Li content. Crystal structures of (b) Li_4SiO_4 and (c) $Li_{2.5}Al_{0.5}SiO_4$. Li sites labelled Li 3, 3′, 4 and 5 are fully occupied in (b), empty in (c) and partially occupied in intermediate compositions. (d) Crystal structure of Li_3N.*

In a perovskite, Li would probably occupy the smaller B site, with octahedral coordination. In these materials, however, Li clearly occupies the A sites, together with the much larger La^{3+} ion. The explanation for the occurrence of this highly improbable solid solution mechanism was resolved by neutron diffraction (to locate the Li^+ ions): instead of occupying the centre of the A-site cavity, Li^+ is, instead, displaced off-centre to either distorted tetrahedral sites or square planar sites forming the 'window' between adjacent La sites. In these, the Li–O bond distances are much more reasonable, ~2.0 Å compared with ~2.4 Å for the La–O distances.

An unusual effect has been found in mixtures or composites of LiI and Al_2O_3. Although there is no evidence that LiI and Al_2O_3 react chemically, the conductivity of equimolar mixtures of LiI and Al_2O_3 is several orders of magnitude higher than that in pure LiI (pure Al_2O_3 is an insulator). The conductivity at 25 °C is $\sim 10^{-5}$ S cm^{-1}. The origin of this effect in such *composite conductors* is surface conductivity at the interface between LiI and Al_2O_3 grains. It may also involve the presence of moisture since LiI is extremely hygroscopic.

8.5.4.9 Proton conductors

Most H^+ ion conductors contain H_2O and decompose readily on heating. Some are hydrates, such as HUP, hydrogen uranyl phosphate, $HUO_2PO_4{\cdot}4H_2O$, which has $\sigma \approx 4 \times 10^{-2}$ S cm^{-1} at 25 °C. Others are made by ion exchange, such as $Na^+ \leftrightarrow H_3O^+$ in hydronium β-alumina, $H_3O^+Al_{11}O_{17}$. Conduction is often by a cooperative (Grotthus) mechanism in which H^+ transfers from one H_2O molecule to an adjacent one, perhaps with reorientation of the water molecules to make the transfer easier. Thus, materials that ionise protons easily, such as solid acid salts like 'antimonic acid', $Sb_2O_5{\cdot}nH_2O$, tend to be good proton conductors.

A completely different class of proton conductors, typified by barium cerate perovskite, $BaCeO_{3-\delta}$, absorbs a small amount of water which is very strongly bound. The H_2O molecules dissociate and the protons are probably attached to oxygens in the region of Ce^{3+} ions ($BaCeO_{3-\delta}$ is non-stoichiometric, containing mixed-valence Ce^{3+} and Ca^{4+}). $BaCeO_{3-\delta}$ has very good proton conductivity at high temperature, e.g. 10^{-2} S cm^{-1} at 600 °C. A selection of conductivity data for proton conductors is shown in Fig. 8.33.

8.5.4.10 Mixed ionic/electronic conductors

So far, we have concentrated on solid electrolytes which have high ionic conductivity but which, electronically, are insulators. For practical uses as solid electrolytes, electronic conductivity must be avoided since this would lead to internal short circuits in batteries or sensors and therefore, loss of performance or accuracy. In materials that are *mixed conductors*, however, both ionic and electronic conduction occur simultaneously; these find applications as electrodes in a variety of devices.

For the ionically conducting component, the structural, compositional and doping features are the same as for ionic conductors, discussed above, i.e. open structures with partial occupancies of the mobile ion sites. For significant levels of electronic conduction, mixed valence of one of the atoms present is a necessary but not sufficient prerequisite.

In mixed conductors, an intriguing question is whether the ionic and electronic conduction components occur independently of each other or whether they somehow influence each other. In order to consider how ionic conduction could be enhanced by the occurrence of simultaneous electronic conduction, we note that whenever an ion hop occurs, local electroneutrality is disturbed. Thus, a hopping cation leaves behind a site that carries a negative charge and its new site carries a positive charge. The negative charge on the original site acts as a drag to further hopping of the cation, known as the *Debye–Hückel 'ion-atmosphere*' or *Debye—Falkenhagen effect*, which is familiar to electrochemists who usually study liquid electrolytes. The ion atmosphere can be reduced by subsequent displacements of ionic species surrounding the charged

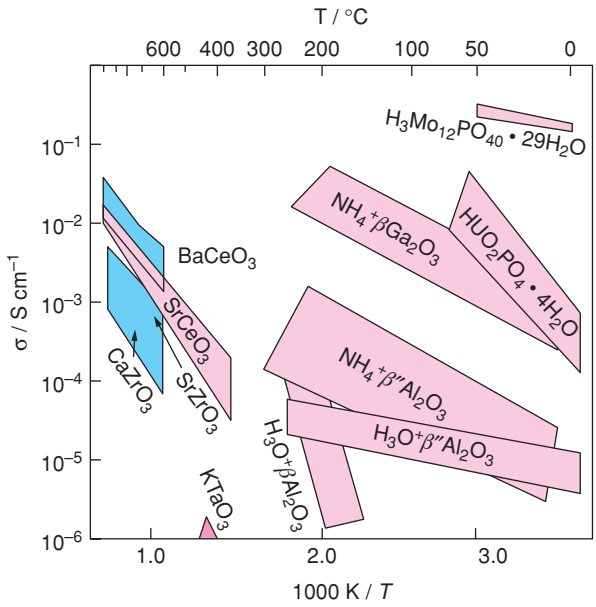


Figure 8.33 ***Conductivity data for a selection of proton conductors. Adapted with permission from T. Kudo 'Survey of types of solid electrolytes' in* The CRC Handbook of Solid State Electrochemistry*, Ed. P. J. Gellings and H. J. M. Bouwmeester, © 1997 CRC Press.***

vacancy, so as to restore electroneutrality to the original site. Alternatively, it could be reduced if the material were a mixed conductor, perhaps by a simple migration or hopping of electrons alongside the migrating cation. The time between successive hops of an ion (i.e. the *site residence time*, which is the inverse of the average, *ion hopping frequency*) may be reduced in the presence of electronic conduction and, effectively, the mobility of the ions increases, as does the conductivity.

For oxide ion conductors, the transition between ionic conduction and mixed conduction can be induced in the *same* material by varying the partial pressure of O_2 in the surrounding atmosphere, Fig. 8.31. For lithium cathode materials, electronic conductivity is associated with mixed valence of transition metal ions in the framework of the cathode material. Thus, $Li_{1-x}CoO_2$ in its partially charged state has a mixture of Co^{3+} and Co^{4+} ions; similarly, $Li_{1-x}FePO_4$ has a mixture of Fe^{2+} and Fe^{3+} and $Li_{2-x}NiMn_3O_8$ has a mixture of Ni^{2+} and Ni^{4+}.

8.5.4.11 Applications of solid electrolytes and mixed conductors

Cells that contain solid electrolytes have a wide range of applications, many of which are not possible with liquid electrolyte-containing cells. The schematic cell shown in Fig. 8.34(a) contains a solid electrolyte membrane separating two electrode compartments. The latter may contain solids, liquids or gases, which may be similar or dissimilar, e.g. oxygen gas at two different pressures or the two components of a cell, e.g. sodium and sulfur. Mixed conductors find varied applications in electrochemical cells but these take the place of one, or both, of the electrode compartments in Fig. 8.34(a). In such cells, the ion that is mobile in the electrolyte is also mobile in the electrode, but additionally, the electrode is an electronic conductor. Electrochemical cells have many applications, as power sources, sensors, smart windows and for thermodynamic measurements. Some examples follow.

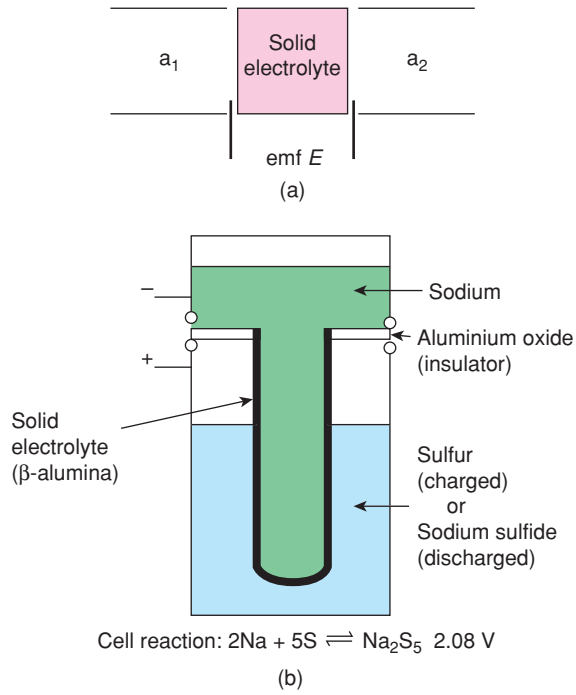


Figure 8.34 ***(a) Electrochemical cell containing a solid electrolyte. (b) The sodium–sulfur cell.***

Thermodynamic measurements

The emf of a cell reaction is given by the *Nernst equation*:

$$E = E_0 + \frac{RT}{nF} \log_e \frac{[\text{Ox}]}{[\text{Red}]} \tag{8.42}$$

Two such equations are usually needed, one for the reaction occurring at each electrode. Suppose that at the anode the following oxidation occurs:

$$\text{M} \rightarrow \text{M}^+ + \text{e}$$

Then

$$E_1 = E_{0_{\text{M/M}^+}} + \frac{RT}{F} \log_e \frac{[\text{M}^+]}{[\text{M}]} \tag{8.43}$$

where $E_{0_{\text{M/M}^+}}$ is the standard redox potential for that reaction, $n = 1$ (number of electrons transferred), $[\text{M}^+]$ and [M] are the concentrations of the two species and F is the Faraday constant, 96 500 C. The M^+ ions generated at the anode diffuse through the solid electrolyte and react at the cathode with X^-, produced according to

$$\text{X} + \text{e} \rightarrow \text{X}^-$$

and for which

$$E_2 = E_{0_{X/X^-}} + \frac{RT}{F}\log_e\frac{[X]}{[X^-]}$$

Summation of E_1 and E_2 gives the overall emf for the reaction $M + X \rightarrow MX$, which is related to the free energy of formation of MX by

$$\Delta G = -nEF \tag{8.44}$$

Electrochemical cells incorporating solid electrolytes may be used to obtain thermodynamic data about materials, e.g. the cell Ag(s)/AgI(s)/Ag_2S(s), S(l), C(s) has been used to measure the free energy of formation of Ag_2S. The cell reaction is $2Ag + S \rightarrow Ag_2S$ and since the reactants, Ag and S, are in their standard states

$$\Delta G^\circ(Ag_2S) = -2EF \tag{8.45}$$

From the temperature dependence of the emf, the entropies and enthalpies of the reaction can be determined.

Sodium–sulfur and Zebra batteries

Much of the impetus for research on solid electrolytes has come from their possible use in new types of battery. The Na/S cell utilises Na^+ β-alumina solid electrolyte, Fig. 8.34(b). It is a high-density secondary battery, i.e. it has a high energy/power to mass ratio and is undergoing extensive development for use in electric vehicles and, particularly, for power station load levelling. It consists of a molten Na anode and a molten S cathode separated by β-alumina solid electrolyte. Usually, the β-alumina is fabricated as a tube closed at one end with Na inside and S outside (or vice versa). Since molten S is a covalently bonded solid, it is an electrical insulator; the cathode material that is used, therefore, consists of conducting graphite felt impregnated with S. The outer casing of the cell, made of stainless steel, serves as the current collector. The cell discharge reaction is

$$2Na + xS \rightarrow Na_2S_x$$

where x depends on the level of charge in the cell. In the early stages of discharge, x is usually given as 5, which corresponds approximately to the formula of the most S-rich sodium sulfide, Na_2S_5, but in fact is one of the limiting compositions of an immiscibility dome shown on the phase diagram for the Na–S system, Fig. 7.24.

The Na/S cell operates at 300–350 °C, which is the lowest temperature at which the discharge products are molten for a large range of compositions. From the phase diagram, when discharge reaches the stage at which $x \leq 3$ (i.e. 60% S, 40% Na), the liquidus curve rises rapidly and crystalline Na_2S_2 begins to form; further discharge would lead to gradual freezing of the cathode.

The open-circuit voltage, *OCV*, of the cell depends on both the level of charge and temperature. The maximum *OCV* for a fully charged cell is 2.08 V at 300 °C. During the first stage of discharge, liquid 'Na_2S_5' gradually forms and the cell *OCV* is constant at 2.08 V whilst the cathode is a mixture of two liquids: 'liquid S' and 'liquid Na_2S_5'. Once the composition 'Na_2S_5' is reached, the cathode enters a region of single phase liquid and the *OCV* gradually drops to 1.78 V until, at approximate composition 'Na_2S_3', the cathode enters the region (Na_2S_2 + liquid) and again, a region of constant *OCV* is encountered. At this stage, the cathode starts to solidify as Na_2S_2 precipitates and discharge of the cell must be terminated. The theoretical energy

storage capacity is 750 Wh kg^{-1}. Experimental values are usually 100–200 Wh kg^{-1} and are unlikely to be improved upon because most batteries give only a small percentage of their theoretical maximum output.

A variation on the Na/S cell is the Na/$FeCl_2$ or Na/$NiCl_2$ cell, also known as the Zebra cell or 'salt and iron filings' cell; it also contains a β-alumina tube as the solid electrolyte. The overall cell reaction is

$$2Na + FeCl_2 \rightarrow 2NaCl + Fe$$

The cell operates at 250 °C and has an *OCV* of 2.35 V. In order to maintain the electrodes in the molten state at 250 °C, the cathode (Fe/$FeCl_2$) is impregnated with molten $NaAlCl_4$. The cell is usually assembled in the discharged state, i.e. with a mixture of NaCl and Fe in the cathode compartment, hence the name, salt and iron filings cell.

Miniature cells, heart pacemakers

Na/S and Zebra cells represent one extreme of scale: for load levelling, units of 1 MW, or greater, are required! At the other extreme, miniature cells which operate at room temperature and have long life rather than high power output find applications in medicine and microelectronic circuits. Various cells have been used satisfactorily, e.g. Ag/$RbAg_4I_5$/I_2 (0.65 V) and Li/LiI/I_2 (2.8 V). In these, iodine alone cannot be used as the cathode because it does not have sufficient electronic conductivity to sustain a discharge current; instead, complexed iodides, e.g. $(CH_3)_4NI_5$ containing polyiodide anions and iodine–poly(2-vinylpyridine) charge-transfer complex, are used. The Li/I_2 cell is widely used for heart pacemaker applications. When operated at 37 °C and current densities of 1–10 μA cm^{-2}, energy densities as high as 0.8 Wh cm^{-3} are possible and the batteries operate in intermittent discharge mode for at least 10 years.

Lithium batteries

Perhaps the most exciting developments at present, both scientifically and commercially, are in Li batteries. The market is expanding hugely, particularly for portable computers, mobile phones and video recorders where secondary Li batteries provide a suitable lightweight, high-density power supply, but increasingly for electric vehicles, EVs, and hybrid EVs.

The components of a rechargeable Li battery are shown in Fig. 8.35(a), with a selection of candidate materials for use as the electrodes and electrolyte. For optimum cell voltage, an anode of high Li activity is required; ideally, Li metal would be used but there are safety issues caused by dendrite formation and internal short-circuiting during cycling. This is believed to be the source of the problems associated with Li batteries used in recent, high profile applications. Instead, various Li-based alloys have been considered and there is currently most interest in Li–C intercalation compounds (the C is in the form of disordered, graphite-like sheets and fragments).

Electrolytes in current commercial usage are liquids based on a non-aqueous solvent such as propylene carbonate, PC, or ethylene carbonate, EC, in which is dissolved a salt such as CF_3COOLi (Li triflate), $LiClO_4$ or $LiPF_6$. Next-generation Li batteries may utilise a *polymer electrolyte*, which again contains an Li salt such as triflate, but now dissolved in a polar polymer such as poly(ethylene oxide), PEO. The temperature of operation of the battery may be above or below the *glass transition temperature*, T_g, of the polymer (this is the temperature at which the polymer changes from being a rigid solid to a very viscous liquid). If it is above T_g, the electrolyte is liquid-like and conduction of both anions and cations occurs. If it is below T_g, or if the polymer electrolyte is crystalline rather than glassy or amorphous, then the polymer exhibits the mechanical properties of a solid and probably the Li^+ ion conductivity will be significantly lower than its value above T_g. A currently fruitful area of research is the development of novel polymers for use as

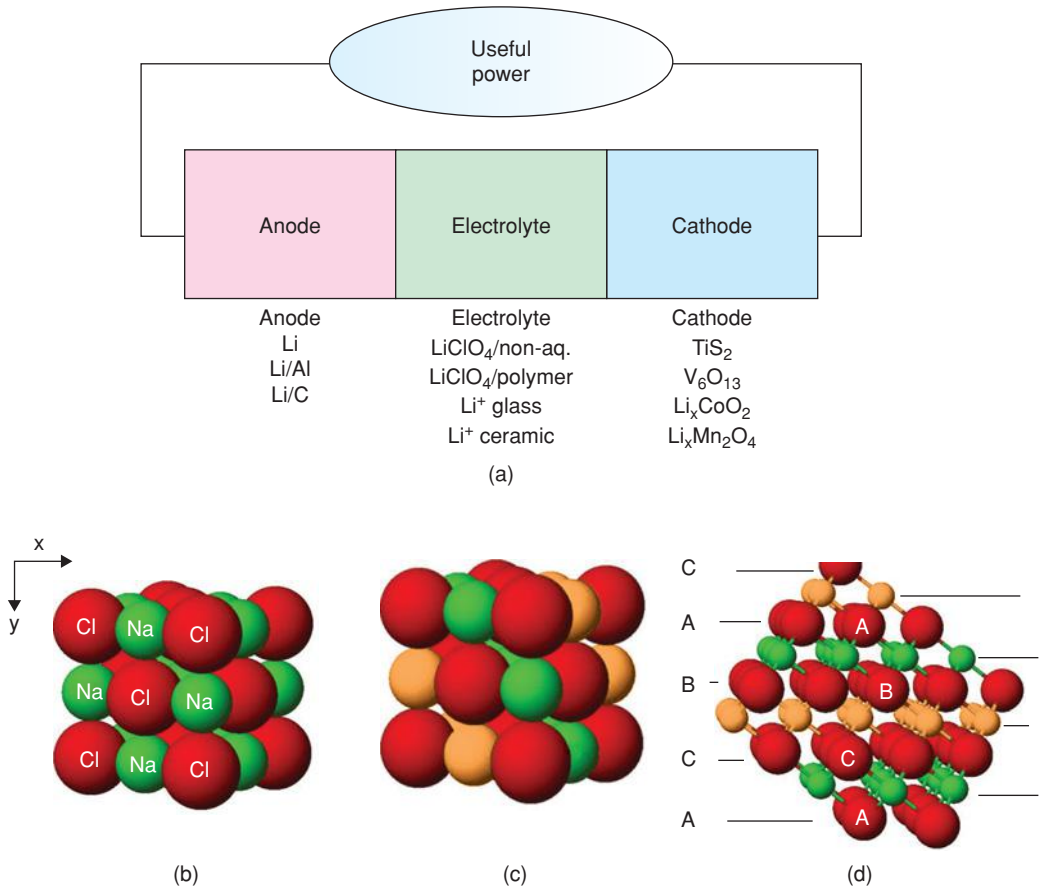


Figure 8.35 ***(a) Components of a secondary lithium battery; (b) crystal structure of NaCl showing layers of Na^+ and Cl^- ions; (c) crystal structure of $LiCoO_2$ showing alternating layers of Li^+ and Co^{3+} ions within a* ccp *oxide ion array; (d) the $LiCoO_2$ structure from a different perspective with alternating layers of Li^+ (orange) and Co^{3+} (green) ions between* ccp *layers of oxide ions (red).***

the electrolyte host which have rigid or semi-rigid mechanical properties but high Li^+ ion conductivity. At present, none of the polymer electrolytes available have sufficiently high Li^+ ion conductivity for commercial applications.

Candidate cathodes in Li batteries are all *intercalation host structures*, which have empty channels or layers capable of accepting Li^+ ions and a mixed-valence species in the host framework that is able to accept, and release, electrons. The first materials to be considered, also named *solid solution electrodes*, were proposed by Whittingham and were layered chalcogenide structures, such as TiS_2. TiS_2 has the CdI_2 structure (Section 1.17.6) and can accept Li^+ ions into empty layers of octahedral sites that separate adjacent TiS_2 sheets. At the same time, and to preserve electroneutrality, electrons enter the TiS_2 sheets and populate the overlapping d_{xy} orbitals on adjacent Ti atoms. The overall intercalation reaction may be written as

$$Ti^{4+}S_2 + xLi \rightarrow Li_x^+ Ti_{1-x}^{4+} Ti_x^{3+} S_2$$

Many other intercalation compounds can function as cathodes. The most important commercially at present is $LiCoO_2$ discovered by Goodenough; it has an 'ordered rock salt structure' in which alternate layers of Li^+ and Co^{3+} ions occur in octahedral sites within the *ccp* oxide ion array, Fig. 8.35(c,d). The formula $LiCoO_2$ refers to the fully intercalated, discharged state. On charging, Li^+ ions are removed and oxidation of Co^{3+} to Co^{4+} occurs. The Co^{3+}/Co^{4+} couple gives a cell voltage of ~4.0 V against Li metal.

Because of the relatively high cost and toxicity of Co, alternative cathodes to $LiCoO_2$ are required. Other layered rock salt materials such as $Li(Ni_{0.5}Mn_{0.5})O_2$ and $Li(Ni_{0.33}Mn_{0.33}Co_{0.33})O_2$, which contain a significant amount of cheap, non-toxic Mn, are possible. Another candidate is $LiMn_2O_4$, with a spinel structure; it is possible to both remove Li from this material and intercalate extra Li into it. Complications arise on insertion of Li, however, since the proportion of Mn^{3+}, which is 50% in $LiMn_2O_4$, i.e. $LiMn^{3+}Mn^{4+}O_4$, increases on Li intercalation, in $Li_{1+x}Mn^{3+}_{1+x}Mn^{4+}_{1-x}O_4$. Mn^{3+} is a Jahn–Teller active ion (d^4) in octahedral coordination and a structural distortion from cubic to tetragonal occurs in $Li_{1+x}Mn_2O_4$. This distortion, and associated change in shape/volume of the $LiMn_2O_4$ crystals, lead to deterioration in the electrolyte–cathode interfacial integrity and consequent decrease in battery performance on cycling. A material of much recent interest as a cathode is $LiFePO_4$, also discovered by Goodenough, with the olivine structure, Fig. 1.45, which utilises the $Fe^{2+/3+}$ couple on Li deintercalation. This contains cheap and non-toxic Fe and P.

Many of the difficulties associated with the use of electrochemical cells and batteries concern the interfacial regions between electrodes and electrolyte. With metal electrodes and ionically conducting electrolytes, interfacial polarisation occurs associated with charge build-up at the interface; usually, a continuous current cannot flow across such an interface. These problems may be avoided by the use of *reversible electrodes*, such as Li_xCoO_2, that allow conduction by both electrons and the mobile ions of the electrolyte. Thus, with a reversible lithium electrode, Li ions may move across the interface between electrode and electrolyte with little or no polarisation problems.

Electrochromic devices, smart windows

Tungsten bronzes, e.g. Na_xWO_3, may act as reversible electrodes and are used in electrochromic devices and smart windows. Tungsten bronzes have good Na^+ mobility because the parent WO_3 structure has a framework built of corner-sharing WO_6 octahedra similar to that in perovskite, Fig. 1.41; a 3D network of interconnected channels permeates the structure along which Na^+ (or other alkali metal cations) may migrate. The oxidation state of tungsten may vary between +V and +VI and the chemical reactions involved in the operation of a tungsten bronze solid solution electrode may be written as

$$x\mathrm{Na}^+ + \mathrm{WO}_3 + \mathrm{xe}^- \rightarrow \mathrm{Na}_x\mathrm{W}^{\mathrm{V}}_x\mathrm{W}^{\mathrm{VI}}_{1-x}\mathrm{O}_3$$

Pure WO_3 is almost colourless and transparent when in the form of a thin film. On intercalation of Na^+ and e^-, the electrons enter the 5*d* band of W (which is completely empty in WO_3) and the material darkens due to absorption of light associated with *d–d* transitions. This darkening forms the basis of the device applications, as illustrated in Fig. 8.36. It is a reversible colour change which is driven by an electrical pulse.

Gas sensors

Electrochemical cells containing solid electrolytes may be used for the measurement of partial pressures of gases or the concentrations of gases dissolved in liquids. An oxygen concentration cell that utilises a stabilised zirconia solid electrolyte in the form of an open-ended tube is shown schematically in Fig. 8.37(a). Inside the tube is a reference gas such as air. The tube is coated with porous metal electrodes to catalyse the absorption and liberation of oxygen gas. If the partial pressure of oxygen that is to be measured, P'_{O_2}, is less than the

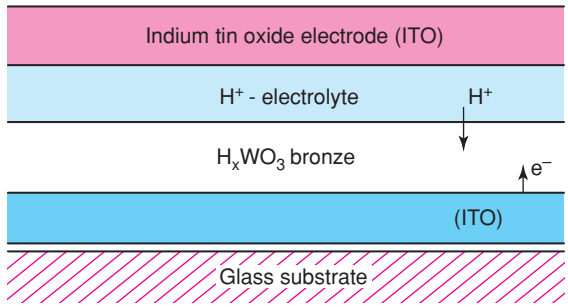


Figure 8.36 ***A thin-film electrochromic device based on a tungsten bronze intercalation electrode.***

reference pressure, P''_{O_2}, the electrode reactions shown in (ii) take place and oxide ions migrate through the solid electrolyte from right to left. The *Nernst equation* relates the difference in oxygen partial pressure to the cell voltage:

$$E = \frac{RT}{4F} \log_e \left(\frac{P''_{O_2}}{P'_{O_2}} \right) \tag{8.46}$$

Gas sensors operate in open-circuit mode, i.e. no current is drawn. Instead, the voltage difference between inner and outer electrodes is used to detect and/or measure P'_{O_2}.

The oxygen sensor is used at high temperatures (in order that transport of O^{2-} ions occurs sufficiently rapidly that equilibrium at the gas–solid interface can be reached) and can measure oxygen partial pressures as low as 10^{-16} atm. Stabilised thoria, ThO_2, solid electrolyte is ionically conducting over a wider range of oxygen partial pressures than YSZ and may be used for $P_{O_2} < 10^{-16}$ atm. Oxygen concentration cells such as the zirconia probe have various uses, e.g. in the analysis of vehicle exhaust gases and pollution, measurement of the consumption of oxygen gas during respiration and study of equilibria such as CO/CO_2, H_2/H_2O and metal/metal oxide. They are also used to probe the oxygen activity in molten steel by dipping the probe into the melt. Oxygen concentrations are read from a calibrated meter; the response of the probe is usually very rapid.

Transport number measurements

For an oxide ion-conducting membrane, the voltage generated when the two electrode compartments contain different activities of oxygen is given by equation (8.46). If, however, the membrane is a mixed oxide ion–electron conductor, the voltage generated, E_{exp}, is smaller and is related to the ionic transport number, t_i, by

$$t_i = E_{exp}/E_{theor} \tag{8.47}$$

Hence cells such as that in Fig. 8.37(a) may also be used for transport number measurements.

Solid oxide fuel cells, SOFCs; steam electrolysers; oxygen pumps

Stabilised zirconia is used in either tube or thin-film form as the solid electrolyte or separator in high-temperature SOFC applications. The two electrode compartments contain (a) air or O_2 and (b) a fuel gas, e.g. H_2, CO. The zirconia is sandwiched between a porous Ni anode and a mixed conducting, lanthanum

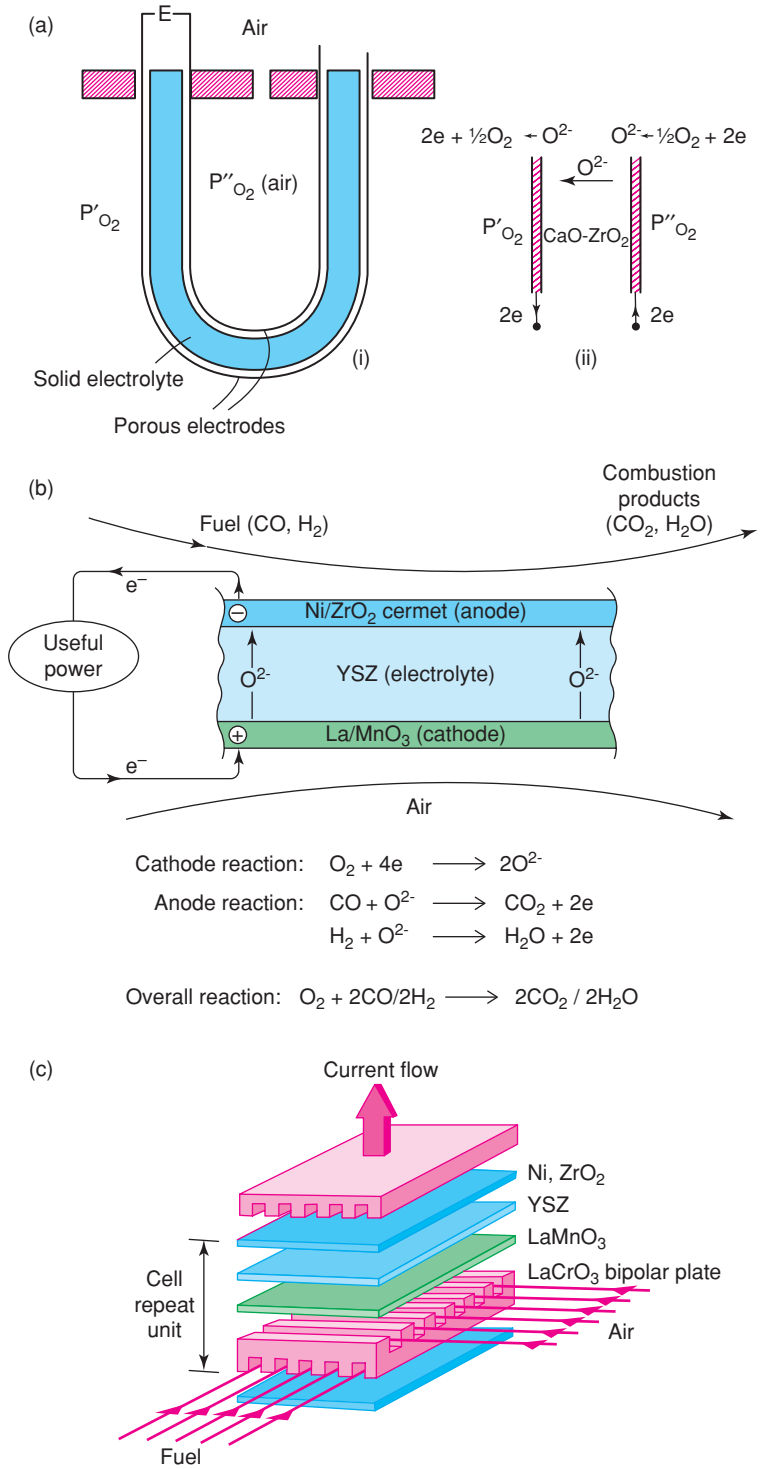


Figure 8.37 ***(a) Oxygen concentration cell with stabilised zirconia solid electrolyte and (b, c) solid oxide fuel cell.***

manganite, $LaMnO_3$, cathode; the cell reaction is, for example, $H_2 + {}^1/_2O_2 \rightarrow H_2O$ or $CO + {}^1/_2O_2 \rightarrow CO_2$. The advantage of this type of fuel cell is that electrode polarisation problems are minimal and high current densities, e.g. 0.5 A cm^{-2}, can be achieved. The operation and design of a multilayer SOFC is shown in Fig. 8.37(b) and (c).

The fuel cell can also be used as an electrolyser and as a means of storing energy by reversing the direction of current flow. Thus, steam can be decomposed and the products, H_2 and O_2, stored. It is also used to regenerate O_2 from CO_2 in, for example, spaceships, to deoxidise liquid metals and to pump O_2 from one compartment to another as a means of gas purification.

8.6 Dielectric Materials

Dielectric materials ideally have zero *dc* electrical conductivity. Dielectric properties may be defined from the behaviour of the material in a parallel plate capacitor. This is a pair of conducting plates, parallel to one another and separated by a distance, d, that is small compared with the linear dimensions of the plates, Fig. 8.38(a). With a vacuum between the plates, the capacitance C_0 is defined as

$$C_0 = \frac{e_0 A}{d} \tag{8.48}$$

where e_0 is the *permittivity of free space*, 8.854×10^{-12} F m^{-1}, and A is the area of the plates. Since e_0 is constant, the capacitance depends only on the dimensions A and d of the capacitor. On applying a potential difference, V, between the plates, a quantity of charge, Q_0, is stored on them, given by

$$Q_0 = C_0 V \tag{8.49}$$

If a dielectric substance is now placed between the plates and the same V is applied, the amount of charge stored increases to Q_1 and the capacitance increases to C_1. This is because the effect of V is to cause a small polarisation of the valence electron charge clouds, although no long range of motion of ions or electrons should occur. The polarisation disappears when V is removed. The *dielectric constant* or *relative permittivity*, ε', of the dielectric is related to the increase in capacitance by

$$\varepsilon' = \frac{C_1}{C_0} \tag{8.50}$$

The magnitude of ε' depends on the degree of polarisation that can occur. For air, $\varepsilon' = 1$. For most ionic solids such as Al_2O_3 and NaCl, $\varepsilon' = 5$–10. For ferroelectric materials such as $BaTiO_3$, $\varepsilon' = 10^3$–10^4.

The terminology associated with the effects of applied voltage on dielectric materials is varied and can be confusing. A summary of terms, symbols and units (mainly SI) is given in Table 8.7. Several of the terms can be grouped together because of their similarities. Thus, an *applied field*, E_0, may be modified locally by the presence of internal fields such as those associated with dipoles, whether permanent or induced, to give a *local field*, E_{local}. The terms *charge* and *capacitance* are related since capacitance is the charge stored for a given applied voltage. Capacitance may be specified in terms of unit area or in terms of volume-related area/thickness. *Polarisation* is related to capacitance as it specifies the separation between positive and negative charges and also the magnitude of the charge.

The response of a material to an applied field is described using three terms: *permittivity*, *polarisability* and *dielectric susceptibility*, which essentially represent the same phenomena but are defined in different ways.

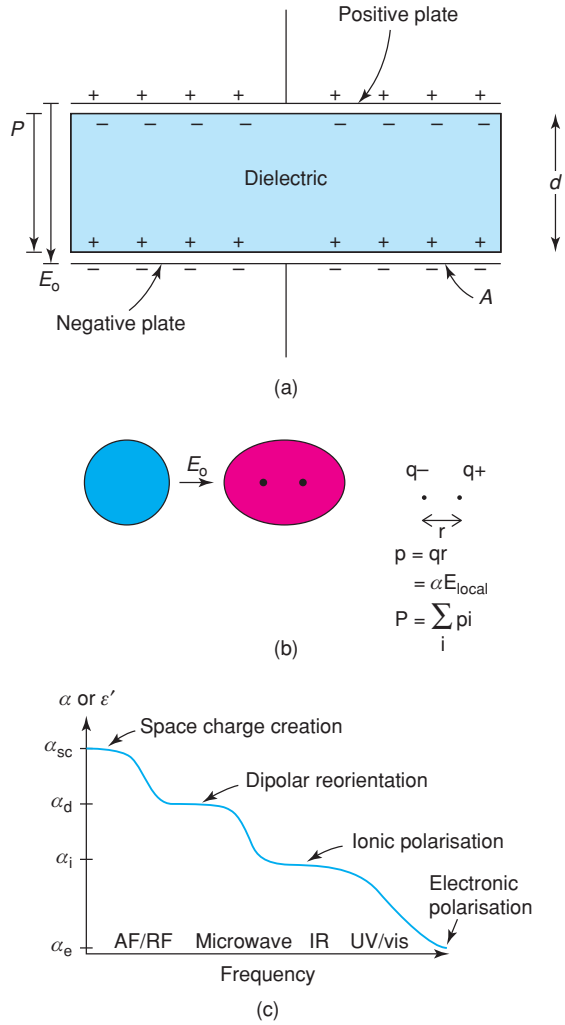


Figure 8.38 ***(a) Dielectric material between the plates of a parallel plate capacitor, (b) creation of induced dipoles in an electric field and (c) frequency dependence of polarisability and permittivity.***

Dielectric susceptibility is the same as (relative) permittivity after subtraction of the contribution from the permittivity of free space, i.e.

$$e_0\chi = e_0\varepsilon' - e_0;\ \chi = (\varepsilon' - 1) \tag{8.51}$$

whereas polarisability, d, is a measure of the dipole moment, p, created in response to E_0, Fig. 8.38(b). *Polarisation* is then the summation of all the individual dipole moments that are created, i.e.

$$P = \sum_i p_i = \sum_i \alpha_i E_{\text{local},i} \tag{8.52}$$

where there are i dipoles per unit volume.

Table 8.7 Terminology and units used with dielectrics

Term	Symbol	Units
Applied field	E_0	V m^{-1}
Local field	E_{local} ($= E_0 + E_1 + \ldots$)	V m^{-1}
Charge	Q	C
Capacitance	C	F or C V^{-1}
	C/unit area/thickness	F m^{-1}
	C/unit area	F m^{-2}
(Relative) permittivity or dielectric constant	ε' or ε_r	Dimensionless
Permittivity of free space	e_0	F m^{-1}
Polarisability	α	C m^2 V^{-1} or F m^2
Dipole moment	p	C m
Polarisation	$P(\Sigma p$/volume)	C m^{-2}
Dielectric susceptibility	χ	m^3

The terms ε', α and χ do not generally have fixed values for a particular material but are dependent on the measuring frequency, unless measurements are made under *dc* conditions. This is because different processes contribute to dipole moment creation and reorientation, which operate on different timescales, as shown schematically in Fig. 8.38(c). At the highest frequencies, in the UV/vis region, small displacements of the electronic charge clouds on individual atoms may occur, giving a difference in the centre of gravity of the electronic charge distribution from that of the positive nucleus, Fig. 8.36(b); this is the *electronic polarisability*, α_e.

With decreasing frequency, small displacements of individual atoms or ions are possible, as shown by the familiar peaks in IR/Raman spectra, which are associated with specific vibrational modes. At these frequencies, *ionic polarisability*, α_i, is detected. At still lower frequencies, molecular motions, rotations and dipole reorientations occur in the microwave region leading to the *orientational polarisability*, α_D.

At the lowest frequencies, in the audio- and radio-wave regions, long-range displacement of species, either ionic or electronic, is possible and, provided that a *dc* current does not flow, separation of charge leading to *space charge polarisability*, α_{sc}, results.

Polarisability and permittivity show a similar frequency dependence, Fig. 8.38(c) and are related to each other by the *Clausius–Mossotti relation*, which is presented in various ways, including

$$\frac{\varepsilon' - 1}{\varepsilon' + 2} = \frac{N\alpha}{3e_0} \tag{8.53}$$

in which N is the number of polarisable species per unit volume in a material that is homogeneous, isotropic and does not contain permanent dipoles. Although this provides useful understanding of the link between ε' and α, in practice the frequency dependence of ε' and α may be less clear-cut than shown in Fig. 8.38(c) and, consequently, equation (8.53) is only semiquantitative. At optical frequencies in the visible region, at which contributions to α from α_{sc}, α_D and α_i should be absent, the permittivity is related to the refractive index, n, by

$$\varepsilon' = n^2 \tag{8.54}$$

Since n for most solids is in the range 1.4–2.9, the limiting high-frequency ε' value should be in the range 2–8. Insulators such as Al_2O_3, MgO and $MgAl_2O_4$ have permittivities in the range 5–10; there is, therefore, very little contribution from α_{sc}, α_D and α_i effects.

8.6.1 From dielectrics to conductors

Dielectrics and conductors represent two extremes of electrical behaviour. In dielectrics, nothing moves apart from possible small displacements of valence electron charge clouds, whereas in conductors, whether ionic or electronic, long-range migration of one or more charged species occurs. Such displacements occur randomly in response to the thermal energy of atoms, ions and electrons; on application of a voltage, a net displacement or drift may occur in the direction of the applied field.

Between these two extremes of behaviour, a range of phenomena involving short-range charge displacement occurs, leading to a wide variety of properties and applications, as summarised in Fig. 8.39. Let us first summarise these before considering some in more detail.

In *low-loss dielectric materials,* nothing moves; they are essential components of electronic circuits where they form insulating substrates upon which the circuit components are deposited, often by thin-film lithographic techniques. They are also used as insulators in spark plugs in engines and to suspend power cables from pylons in the electricity supply grid network. As well as the use of ceramics in insulators, non-conducting polymers such as poly(vinylidene difluoride), PVDF, are used as insulating sheaths to carry conducting wires and cables.

High-permittivity dielectrics also have very low conductivity or dielectric loss but permittivities in the range 50–150 and are used as *microwave dielectrics*, which are essential for the mobile communications industry. This is because the size of microwave receivers used in mobile phones depends inversely on the permittivity of the microwave dielectric; the recent discovery of new, high-permittivity microwave dielectric materials has allowed the miniaturisation of mobile phones.

The crystal structures of *non-polar dielectrics* are centrosymmetric, which means that there are no dipole moments associated with the arrangements of cations and anions in the crystal lattice. A classic,

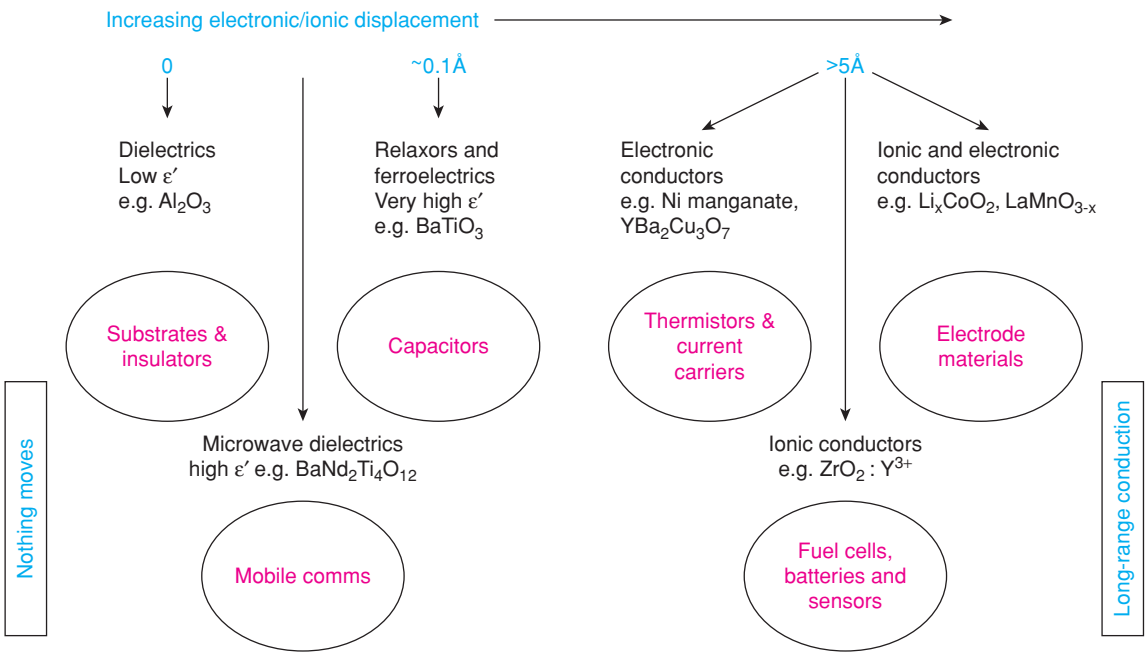


Figure 8.39 ***Response of various electroceramic materials to a small applied voltage.***

non-crystalline, example of a polar dielectric is the water molecule, Fig. 8.40(a), in which the electronegative oxygen carries a residual charge $q-$ and the hydrogens have net charge $q+$. Because the water molecule is non-linear and non-centrosymmetric, it possesses a dipole moment due to the spatial separation of negative and positive charge. In response to an applied voltage V, water molecules can rotate as shown schematically in Fig. 8.40(b) and this gives rise to a permittivity value of 81 for liquid H_2O.

Crystalline dielectric materials with non-centrosymmetric structures are referred to as *polar dielectrics* and may have high permittivities if the net polarisation or dipole moment can be modified in an applied field. The classic example of this is ferroelectric $BaTiO_3$, in which the Ti^{4+} cation in the tetragonal polymorph is displaced off-centre within its TiO_6 octahedron giving rise to a local dipole moment, Fig. 8.40(c) and (d). Effectively, the Ti^{4+} ion can 'rattle' within the octahedron and be closer preferentially to any of the six surrounding oxygens. The facile ability of Ti^{4+} to move position in response to an applied field leads to very high permittivities, as high as 12 000. Such high permittivities give rise to a significant amount of charge storage, equation (8.50), and as such are essential components of capacitors (over 10^{12} are now fabricated annually). In addition to their high permittivity, the orientation of dipole moments of ferroelectric materials can be reversed in an applied voltage leading to hysteresis loops of the kind shown in Fig. 8.40(e) and representing the electrical analogue of the hysteresis loops exhibited by ferromagnetic materials.

Polar dielectric materials also exhibit two other properties if their permittivity is modified under the action of an external stimulus. If the permittivity is temperature dependent because, for instance, cation–anion bond lengths or bond angles show small changes with temperature, the material is *pyroelectric*. Pyroelectric materials are essential components of temperature sensors such as are used in night-time cameras. The change in temperature caused by a heat-emitting body is detected by a change in voltage across the pyroelectric detector.

A related phenomenon is *piezoelectricity*, in which a dipole moment or a change in dipole moment is induced by mechanical pressure on the material. Two sets of applications arise, those in which an applied pressure induces a change in voltage such as the sparks generated by mechanical gas lighters or the mild electrical shock generated when attempting to activate a trick ballpoint pen. In the second set, application of an alternating voltage leads to associated mechanical deformation of the piezoelectric material and the generation of sound waves in piezoelectric transducers such as are used in underwater sonar detection devices. The most widely used piezoelectric is a lead zirconate titanate perovskite with approximate formula $Pb(Zr_{0.48}Ti_{0.52})O_3$.

Polar dielectrics can be separated into two groups, those that are ferroelectric, i.e. in which a change in dipole orientation occurs in response to an applied field, and those which are non-ferroelectric, in which a change in magnitude, but not orientation, of the dipolar component occurs. The structural displacements that occur in ferroelectric materials are now reasonably well understood and typically involve displacement of an ion by a small fraction of a bond length, e.g. 0.1–0.3 Å, Fig. 8.40(c,d). There are also other families of materials in which dipolar reorientation is believed to occur but which appear not to show ferroelectric hysteresis loops:

In aliovalent-doped alkali halides such as Ca-doped NaCl, charge compensation associated with substitution of divalent Ca^{2+} ions involves the creation of an equal number of cation vacancies. Overall, electroneutrality is preserved, but because the vacancies carry a net negative charge, V_{Na}', and the substitutional Ca ions carry a net positive charge, $Ca_{Na}^{\bullet}$, the defects associate to form dipolar units. The dipoles may reorient, by hopping of the cation vacancy around the relatively fixed Ca^{2+} ion. Although the vacancies hop at random as a consequence of their thermal energy, in an applied electric field there is a net direction of hopping so as to partially align the dipoles with the electric field.

In the group of *relaxor* materials, long-range ferroelectric domains do not form but instead short-range clusters of parallel-oriented dipoles exist. A good example is the perovskite $Pb(Mg_{1/3}Nb_{2/3})O_3$, PMN, which contains Mg and Nb disordered over the octahedral cation sites of the perovskite structure. However, Mg^{2+} is larger than Nb^{5+} and only the Nb^{5+} cations show a tendency for off-centre displacement within their octahedra. As a consequence, the centrosymmetric MgO_6 octahedra effectively block long-range polar domain

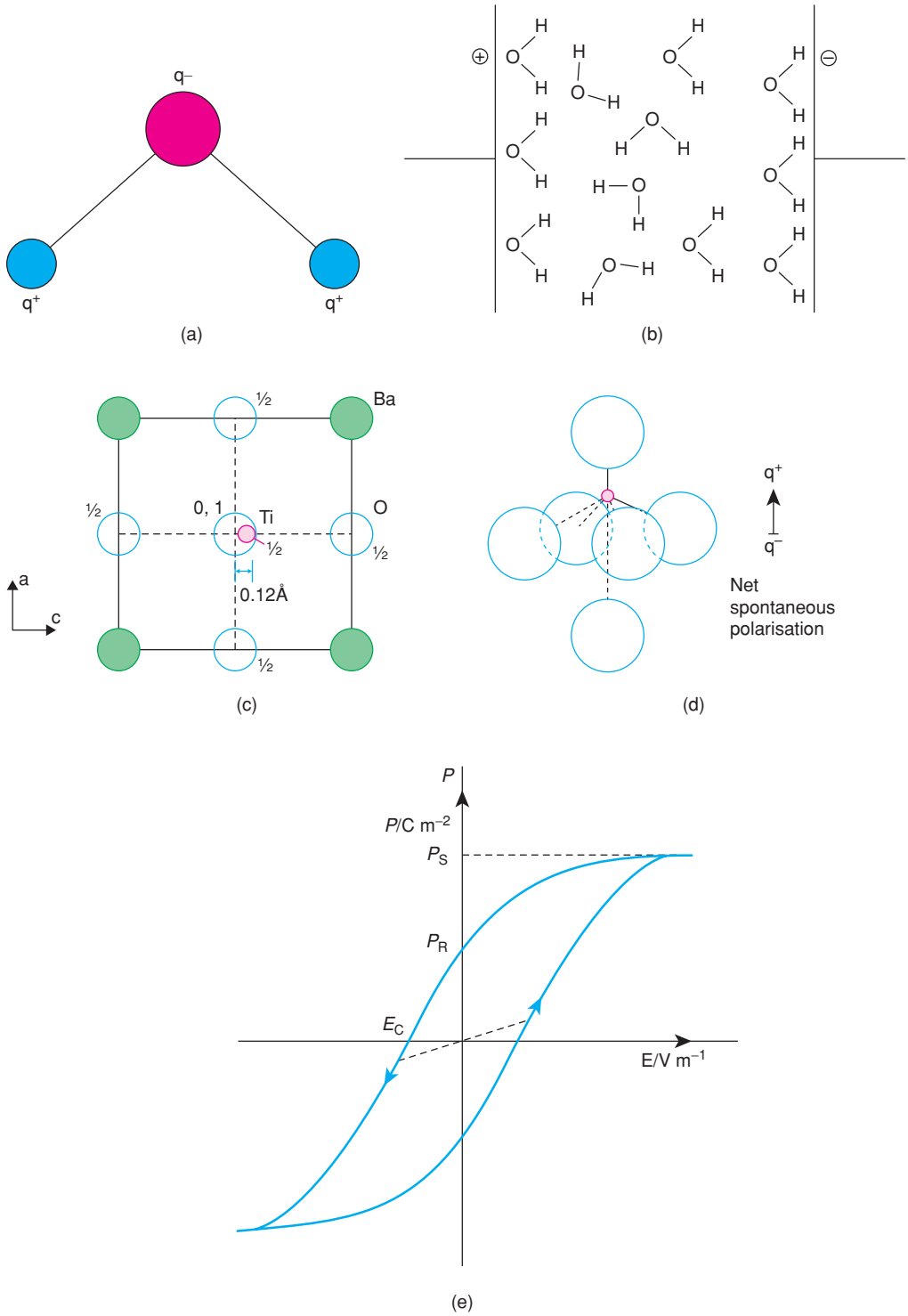


Figure 8.40 *(a) The polar water molecule, (b) reorientation of water molecules in an applied field, (c) polar structure of tetragonal $BaTiO_3$ and (d) spontaneous polarisation of a TiO_6 octahedron in $BaTiO_3$. Note that, in principle, there are six equivalent orientations in which this can occur. (e) Hysteresis loop of a ferroelectric. The dashed line passing through the origin represents the behaviour of normal dielectric materials.*

formation. A useful property of relaxors such as PMN is that the permittivity associated with the ferroelectric to paraelectric phase transition becomes broadened, giving a permittivity that is high over a wide temperature range, unlike the case of $BaTiO_3$ whose permittivity passes through a high but sharp maximum over a limited temperature range. For capacitor applications, it is necessary that the capacitance values change little with temperature over a specified range and therefore this is one of the main applications of relaxors.

Thus far, we have dealt with a range of dielectric and leaky (or lossy) dielectrics in which any charge displacement is limited to small fractions of interatomic distances. For longer range displacement of either, or both, electrons (or holes) and ions, *dc* conduction occurs, Fig. 8.39, and we enter into the realm of semiconductors, metals, solid electrolytes and mixed conductors, as discussed earlier in this chapter.

8.7 Ferroelectrics

Ferroelectric materials are distinguished from ordinary dielectrics by (a) their extremely high permittivities, (b) the possibility of retaining some residual electrical polarisation after an applied voltage has been switched off and (c) the reversibility of the polarisation, P, or dipole moment, under the action of an applied reverse voltage, giving rise to a P versus V hysteresis loop. As the voltage V applied across a dielectric substance is increased, a proportional increase in P, or stored charge, Q, occurs. With ferroelectrics, this simple linear relation between P and V does not hold, as shown in Fig. 8.40(e). Instead, more complicated behaviour with a *hysteresis loop* is observed. The polarisation observed on increasing the voltage is not reversible on subsequently decreasing the voltage. Ferroelectrics exhibit a *saturation polarisation*, P_s, at high field strength (for $BaTiO_3$, $P_s = 0.26$ C m^{-2} at 23 °C) and a *remanent polarisation*, P_R, which is the value retained as V is reduced to zero after saturation. In order to reduce P to zero, a reverse field, the *coercive field*, E_c, is required.

Common ferroelectric materials, Table 8.8, are characterised by structures in which one type of cation, e.g. Ti^{4+} in $BaTiO_3$, can undergo significant displacement, e.g. 0.1 Å, relative to its anionic neighbours. These charge displacements give rise to dipoles and the high dielectric constants characteristic of ferroelectrics.

The unit cell of $SrTiO_3$, which has the same perovskite structure as $BaTiO_3$, is shown in Fig. 1.41. In the primitive cubic unit cell, Ti occupies corner octahedral positions, oxygens occupy the cube edge centres and Sr is at the cube body centre. The TiO_6 octahedra link by sharing corners to form a 3D framework; Sr occupies 12 coordinate cavities within this framework. Remember, the unit cell can also be defined with Sr at the corners and Ti at the body centre.

The same structure is stable above 120 °C in $BaTiO_3$ and does not possess a net dipole moment since the charges are symmetrically positioned. The material therefore behaves as a normal dielectric, albeit with a very high dielectric constant. Below 120 °C, structural distortion occurs in $BaTiO_3$. The TiO_6 octahedra are

Table 8.8 ***Some ferroelectric materials***

Material	T_c/°C
Barium titanate, $BaTiO_3$	120
Rochelle salt, $KNaC_4H_4O_6{\cdot}4H_2O$	Between –18 and +24
Potassium niobate, $KNbO_3$	434
Potassium dihydrogenphosphate, KDP, KH_2PO_4	–150
Lead titanate, $PbTiO_3$	490
Lithium niobate, $LiNbO_3$	1210
Bismuth titanate, $Bi_4Ti_3O_{12}$	675
Gadolinium molybdate, GMO, $Gd_2(MoO_4)_3$	159
Lead zirconate titanate, PZT, $Pb(Zr_xTi_{1-x})O_3$	Depends on x

no longer regular because Ti is displaced off its central position and in the direction of an apical oxygen, Fig. 8.40(c). This gives rise to a spontaneous polarisation (d). If similar parallel displacement occurs in all TiO_6 octahedra, a net polarisation of the solid results.

In ferroelectric $BaTiO_3$, the individual TiO_6 octahedra are polarised all of the time; the effect of applying an electric field is to persuade the individual dipoles to align themselves with the field. When complete alignment occurs the condition of *saturation polarisation* is reached, Fig. 8.40(e). From the magnitude of P_s, it is estimated that Ti is displaced by ~0.1 Å off the centre of its octahedron and in the direction of one of the oxygens. This has been confirmed by X-ray crystallography. This distance of 0.1 Å or 10 pm is fairly small compared with the average Ti–O bond distance of ~1.95 Å in TiO_6 octahedra. Alignment of dipoles is shown schematically in Fig. 8.41(a); each arrow represents one distorted TiO_6 octahedron and all are shown with a common direction of distortion.

In ferroelectric $BaTiO_3$, *domain structures* form in which adjacent TiO_6 dipoles align parallel, Fig. 8.41(d). The domains are of variable size but are usually fairly large, tens or hundreds of ångströms across. Within

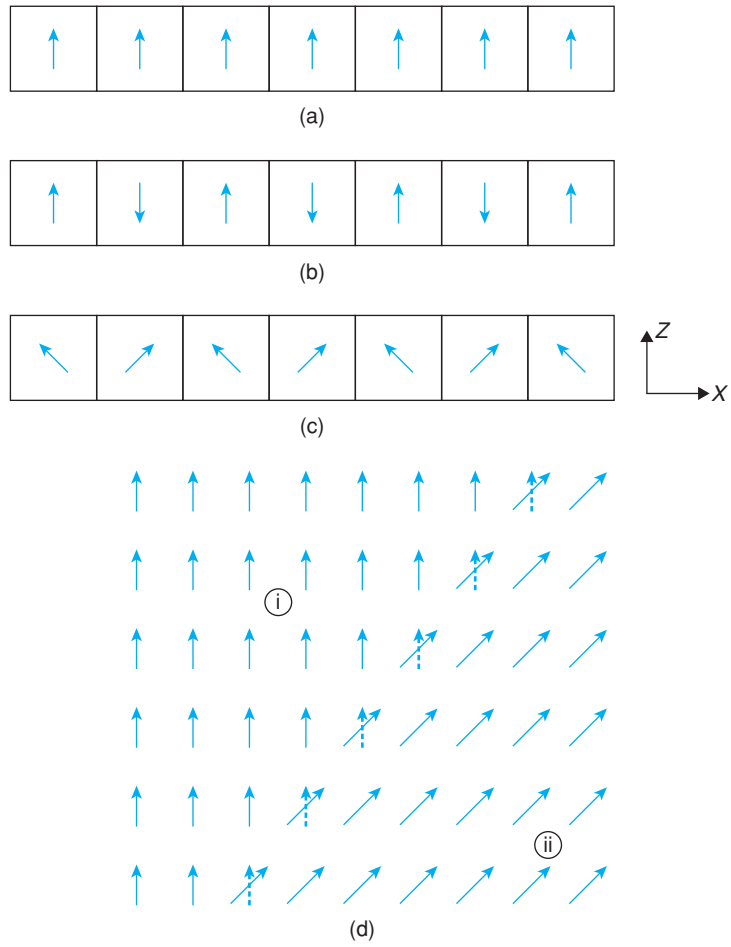


Figure 8.41 ***Dipole orientation (schematic) in (a) a ferroelectric, (b) an antiferroelectric and (c) a ferrielectric. (d) Ferroelectric domains separated by a domain wall or boundary.***

a single domain, the dipoles have a common orientation. The net polarisation is the vector resultant of the polarisations of the individual domains.

Application of an electric field across a ferroelectric leads to a change in P. This can arise from several possible processes:

(a) The direction of polarisation of the domains may change. This would happen if all the TiO_6 dipoles within a domain were to change their orientation, e.g. if all the dipoles in domain (ii) in Fig. 8.41(d) changed their orientation so as to be parallel to the dipoles in domain (i).
(b) The magnitude of P within each domain may increase, especially if some randomness in dipole orientation is present before the field is applied.
(c) *Domain wall* migration may occur such that favourably oriented domains grow in size at the expense of unfavourably oriented domains. For example, domain (i) in Fig. 8.41(d) may grow by migration of the domain wall one step to the right. For this to happen, the dipoles at the edge of domain (ii) change their orientation to the positions shown dashed and become part of domain (i).

The ferroelectric state is usually a low-temperature condition since increasing thermal motion at high temperatures tends to break down the common displacement in adjacent octahedra and destroy the domain structure. The temperature at which breakdown occurs is the *ferroelectric Curie temperature*, T_c, Table 8.7. Above T_c, the material is paraelectric (i.e. non-ferroelectric). High dielectric constants still occur above T_c, Fig. 8.42, but no residual polarisation is retained in the absence of an applied field. Above T_c, ε' is usually given by the Curie–Weiss law:

$$\varepsilon' = C/(T - \theta) \tag{8.55}$$

where C is the Curie constant and θ the Curie–Weiss temperature. Curie–Weiss behaviour is characterised by a linear plot of $(\varepsilon')^{-1}$ against T, Fig. 8.42(b). Usually, T_c and θ either coincide or differ by only a few degrees.

The ferroelectric–paraelectric transition, at T_c, is an example of an order–disorder phase transition. However, unlike order–disorder phenomena in, say, brass, no long-range diffusion of ions occurs. Rather, the ordering that occurs below T_c involves preferential distortion or tilting of polyhedra and is therefore an example of a *displacive phase transition*. In the high-temperature paraelectric phase, the distortions or tilts of the polyhedra, if they occur at all, are randomised.

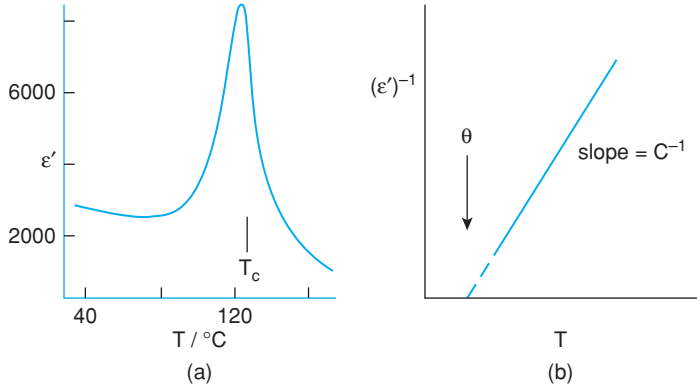


Figure 8.42 ***(a) Dielectric constant of barium titanate ceramic. (b) Curie–Weiss plot.***

A necessary condition for a crystal to exhibit spontaneous polarisation and be ferroelectric is that its space group should be non-centrosymmetric. Often the symmetry of the paraelectric phase stable above T_c is centrosymmetric and the ordering transition that occurs on cooling simply involves a lowering of symmetry to that of a non-centric space group.

Several hundred ferroelectric materials are now known, including a large number of oxides that have distorted (non-cubic) perovskite structures. These contain cations that are happy in a distorted octahedral environment since they are slightly too small to fit snugly inside the octahedra formed by six surrounding oxide ions. Examples of cations commonly found in ferroelectric structures are Ti^{4+}, Nb^{5+} and Ta^{5+}. The asymmetric bonding within the MO_6 octahedron gives rise to spontaneous polarisation and a dipole moment.

Not all perovskites are ferroelectric, e.g. $BaTiO_3$ and $PbTiO_3$ are whereas $CaTiO_3$ is not, and this correlates with the radii of the ions involved. The larger Ba^{2+} ion causes an expansion of the unit cell relative to that of $CaTiO_3$, which results in longer Ti–O bonds in $BaTiO_3$ and allows the Ti^{4+} ions more flexibility to move within the TiO_6 octahedra. Other ferroelectric oxides contain cations that are asymmetrically bonded because of the presence of a lone pair of electrons in their outer valence shell. These are cations of the heavy *p*-block elements in oxidation states two less than the group valency, e.g. Sn^{2+}, Pb^{2+} and Bi^{3+}, which exhibit the so-called *inert pair effect*.

Ferroelectric oxides are used in capacitors because of their high dielectric constants, especially near to T_c, Fig. 8.42. In order to maximise ε' for practical applications, it is necessary to displace the Curie point so that it is close to room temperature. The Curie point of $BaTiO_3$, 120 °C, may be lowered and broadened when either Ba^{2+} or Ti^{4+} are partially replaced by other ions. The substitution $Ba^{2+} \leftrightarrow Sr^{2+}$ causes a unit cell contraction and a reduction in T_c; replacement of 'active' Ti^{4+} by 'non-active' tetravalent ions such as Zr^{4+} and Sn^{4+} also causes a rapid decrease in T_c.

A related type of spontaneous polarisation occurs in *antiferroelectric* materials. In this, adjacent dipoles are arranged antiparallel, Fig. 8.41(b). As a result, the net spontaneous polarisation is zero. Above the antiferroelectric Curie temperature, the materials revert to normal paraelectric behaviour. Examples of antiferroelectrics, with their Curie temperatures, are $PbZrO_3$ 233 °C, $NaNbO_3$ 638 °C and $NH_4H_2PO_4$ –125 °C.

The characteristics of antiferroelectrics are rather different to those of ferroelectrics. The antiferroelectric state is non-polar and no hysteresis loop occurs, although a large increase in permittivity may occur close to T_c (for $PbZrO_3$, ε' is ~100 at 200 °C but ~3000 at 230 °C). Sometimes, the antiparallel arrangement of dipoles in the antiferroelectric state is only marginally more stable than the parallel arrangement in the ferroelectric state and a small change in conditions may lead to a phase transition. For example, application of an electric field to $PbZrO_3$ causes it to change from antiferroelectric to ferroelectric, Fig. 8.43(a); the magnitude of the field required depends on temperature. The polarisation behaviour is then as shown in (b). At low fields, no hysteresis occurs and $PbZrO_3$ is antiferroelectric; at high positive and negative fields, hysteresis loops occur and $PbZrO_3$ is ferroelectric.

A related type of polarisation phenomenon in which the structure is antiferroelectric in certain direction(s) only is shown in Fig. 8.41(c); in the *x* direction, the net polarisation is zero and the structure is antiferroelectric, but in the *z* direction, a net spontaneous polarisation occurs. This type of structure is known as *ferrielectric*; it occurs in, for example, $Bi_4Ti_3O_{12}$ and lithium ammonium tartrate monohydrate.

The role of hydrogen bonding in certain ferroelectric and antiferroelectric materials is shown in Fig. 8.44. Ferroelectric KH_2PO_4 and antiferroelectric $NH_4H_2PO_4$ are both built of isolated PO_4 tetrahedra that link via K^+ and NH_4^+ ions and hydrogen bonds; the hydrogen bonds link oxygens in adjacent PO_4 tetrahedra. The two structures differ mainly in the position of hydrogen in the hydrogen bonds. Each PO_4 tetrahedron forms hydrogen bonds with four adjacent PO_4 tetrahedra. In each of these hydrogen bonds, the hydrogens are displaced so as to be nearer to one oxygen or the other, i.e. in each hydrogen bond, the H atom has a choice of two positions, neither of which is midway along the bond. For each PO_4 tetrahedron, therefore, two hydrogens are close and two are somewhat further away.

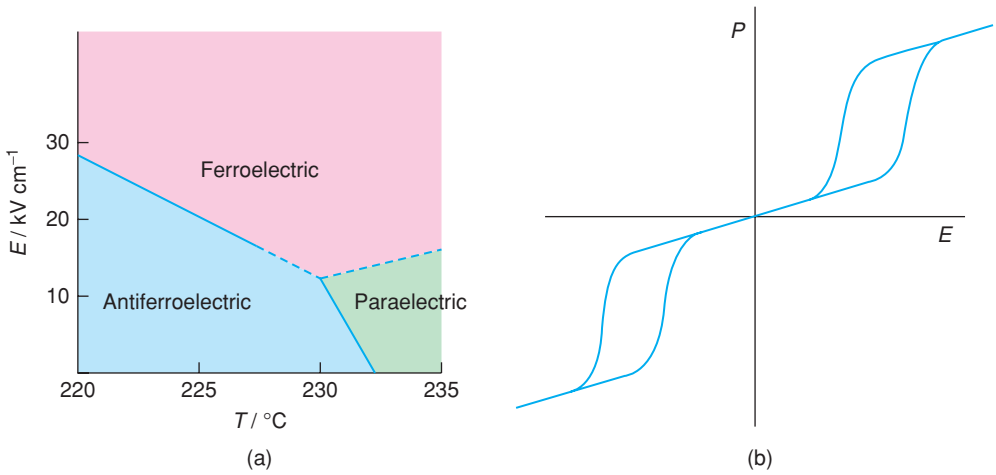


Figure 8.43 ***(a) Antiferroelectric–ferroelectric transition in $PbZrO_3$ as a function of applied field,*** **E.** ***(b) Polarisation behaviour across this transition.***

In the high-temperature, paraelectric forms of KH_2PO_4 and $NH_4H_2PO_4$, the H positions are randomised over the two positions in each bond and a disordered structure obtains. In low-temperature, ferroelectric KH_2PO_4, the hydrogens order themselves so that both are associated with the upper edge of each PO_4 tetrahedron, Figure 8.44(a). The hydrogens are responsible indirectly for spontaneous polarisation within the PO_4 tetrahedra since the P atoms are displaced downwards away from the H atoms. This generates dipoles whose directions are parallel to the *c* crystallographic axis. In order to reverse the directions of the dipoles, it is not necessary to invert the tetrahedra bodily. Instead, a simple movement of H atoms within the H bond achieves the same effect. The two hydrogens associated with the upper oxygens in (a) move away laterally to associate themselves with the lower oxygens of adjacent tetrahedra. At the same time, two hydrogens move

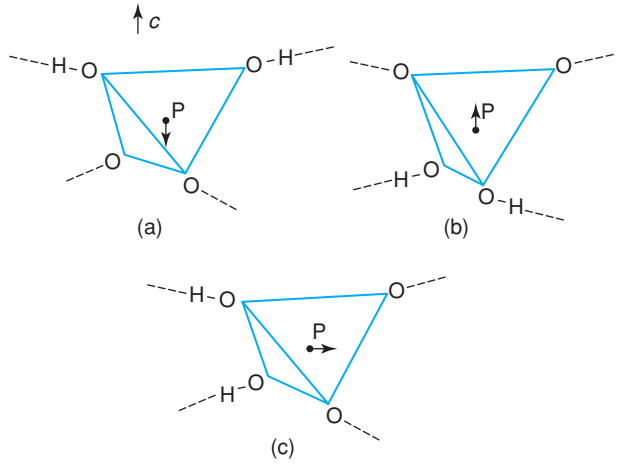


Figure 8.44 ***Displacement of phosphorus within a $PO_2(OH)_2$ tetrahedron giving rise to spontaneous polarisation.***

in to associate themselves with the lower oxygens (b). This motion of H atoms *perpendicular* to c leads to dipole reversal *parallel* to c.

In antiferroelectric $NH_4H_2PO_4$, the two hydrogens of each tetrahedron are associated with one upper and one lower oxygen, Fig. 8.44(c); this creates dipoles in a direction perpendicular to c. The dipole directions are reversed in adjacent tetrahedra and, therefore, the net polarisation over the whole crystal is zero.

8.8 Pyroelectrics

Pyroelectric crystals are related to ferroelectric ones in that they are also non-centrosymmetric and exhibit a net spontaneous polarisation, P_s. Unlike ferroelectrics, however, the direction of P_s cannot be reversed by an applied electric field. P_s is usually temperature dependent:

$$\Delta P_s = \pi \Delta T \tag{8.56}$$

where π is the *pyroelectric coefficient*. This arises because the thermal expansion that occurs on heating changes the sizes (i.e. lengths) of the dipoles. A good example of a pyroelectric crystal is ZnO, which has the wurtzite structure, Fig. 1.35. It contains an *hcp* array of O^{2-} ions with Zn^{2+} ions located in one set of tetrahedral sites, say the T_+ sites. The ZnO_4 tetrahedra all point in the same direction and since each tetrahedron possesses a dipole moment, the crystal has a net polarisation. Opposite, (001) surfaces of a ZnO crystal must contain, respectively, Zn^{2+} and O^{2-} ions as the outermost layers of ions. Usually, however, polar impurity molecules are absorbed onto the crystal in order to neutralise the surface charges. Consequently, the pyroelectric effect in a crystal is often not detectable under constant temperature conditions but becomes apparent only when the crystal is heated, thereby changing P_s.

8.9 Piezoelectrics

Under the action of an applied mechanical stress, piezoelectric crystals polarise and develop electrical charges on opposite crystal faces. As for ferro- and pyroelectricity, the point group symmetry must be non-centrosymmetric. The occurrence of piezoelectricity depends on the crystal structure and the direction of applied stress, e.g. quartz develops a polarisation when subjected to a compressive stress along [100] but not when stressed along [001]. The polarisation, P, and stress, σ, are related to the piezoelectric coefficient, d, by

$$P = d\sigma \tag{8.57}$$

Many crystals that contain tetrahedral groups, e.g. ZnO and ZnS, are piezoelectric since application of a stress distorts the tetrahedra. One of the most important piezoelectrics is PZT, lead zirconate titanate, which is a composition of the solid solution between $PbZrO_3$ and $PbTiO_3$. These solid solutions are also antiferroelectric and ferroelectric at certain compositions, as shown by a partial phase diagram, Fig. 8.45. The best piezoelectric compositions occur at $x \approx 0.5$.

8.10 Applications of Ferro-, Pyro- and Piezoelectrics

The main application of ferroelectrics is in capacitors because they have high permittivity, ε', usually in the range 10^2–10^4. The main commercial materials are $BaTiO_3$ and PZT, which are used in the form of dense,

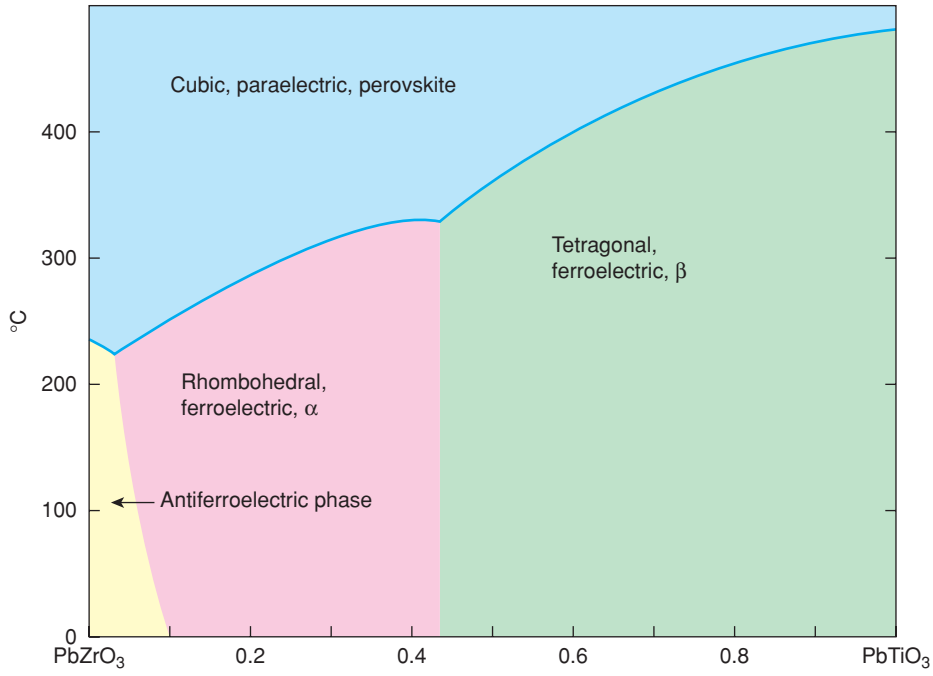


Figure 8.45 ***Phase diagram for the PZT system. Data from E. Sawaguchi,* J. Phys. Soc. Japan*, 8, 615, © 1953 Physical Society of Japan.***

polycrystalline ceramics or thin films. Conventional dielectrics such as TiO_2 or $MgTiO_3$ have ε' in the range 10–100. Hence, for a given volume and geometry, a $BaTiO_3$ capacitor has 10–1000 times the capacitance of a TiO_2-based capacitor. The makeup of a $BaTiO_3$-based multilayer ceramic capacitor, MLCC, is shown in Fig. 8.46. The technology to fabricate these devices is incredibly advanced, with several hundred alternating layers of $BaTiO_3$ and nickel base metal electrodes, each with a thickness of barely 1 μm.

An important use of certain ferroelectrics such as $BaTiO_3$ and $PbTiO_3$, which does not depend directly on their ferroelectricity, is in PTC thermistors (i.e. *p*ositive *t*emperature *c*oefficient *therm*ally sensitive res*istors*), Fig. 8.47. In most non-metallic materials, electrical resistivity decreases with increase in temperature, i.e. the resistivity has a *negative temperature coefficient* (NTC). However, some ferroelectrics, including $BaTiO_3$, show an anomalous and large increase in resistivity on heating as the temperature approaches the ferroelectric–paraelectric transition temperature, T_c. This increase in ρ matches approximately the large increase in ε' close to T_c, although the reasons for the increase in ρ are complex. PTC thermistors are used as switches; when a current is passed through any resistive material, joule heating losses, given by I^2R, cause the material to heat up. With a $BaTiO_3$ thermistor, the resistivity increases dramatically as it heats and consequently the current switches off. Applications include (a) thermal and current overload protection devices, in which the thermistor acts as a reusable fuse, (b) time delay fuses and (c) hair dryers.

Pyroelectric crystals are used mainly in infrared radiation detectors. If desired, they can be made spectrally sensitive by coating the crystal surface with appropriate absorbing material. For detectors, it is desirable to maximise the ratio π/ε', which means that ferroelectric materials with high dielectric constant are not suitable. The best detector material currently is triglycine sulfate.

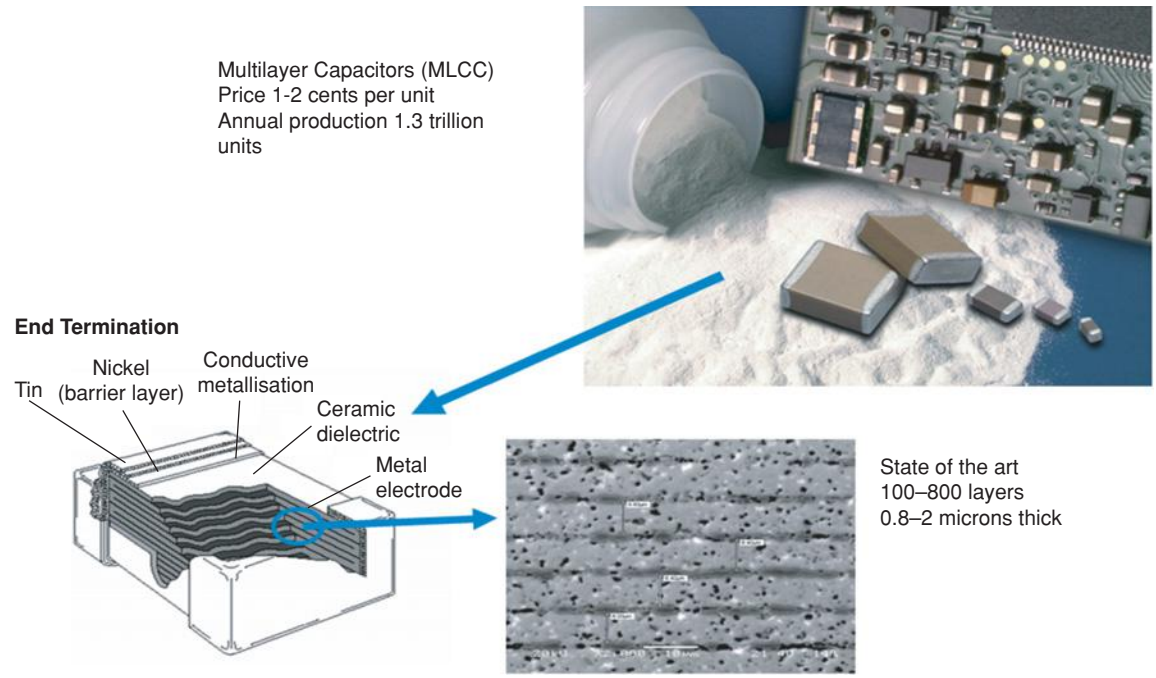
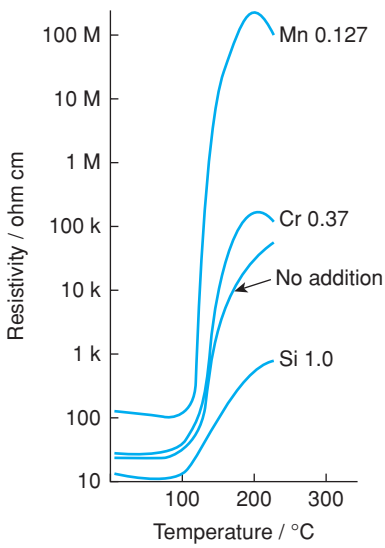


Figure 8.46 ***A multilayer ceramic capacitor.***

Ceramic switches that turn off and on with changing temperature; cheap widgets; small profit margins.



Reversible semiconductor to insulator transition on heating.

Typical PTC thermistor:

$BaTiO_3$ with:

- excess TiO_2
- some Ca for Ba
- donor dopant, e.g. Y
- sintering aid: SiO_2
- Mn additive to improve profile
- "magic ingredient"

Processing is critical; final ceramic is far from thermodynamic equilibrium.

Applications in temperature controllers e.g. hairdryers and thermal fuses.

Figure 8.47 ***Positive temperature coefficient resistivity in semiconducting $BaTiO_3$ ceramic with various dopants.***

Piezoelectric crystals have been used for many years as transducers for converting mechanical to electrical energy, and vice versa. Applications are diverse, e.g. as bimorphs in microphones, earphones, loudspeakers and stereo pick-ups; as fuses, solenoid ignition systems and cigarette lighters, sonar generators and ultrasonic cleaners. More complex devices are used in transformers, filters and oscillators. Most of these applications use PZT ceramics, quartz, Rochelle salt or $Li_2SO_4 \cdot H_2O$.

9

Magnetic Properties

9.1 Physical Properties

Inorganic solids that exhibit magnetic effects other than *diamagnetism*, which is a property of all substances, have some unpaired electrons in their outer valence shells: electrons in inner core levels are always paired in fully occupied orbitals. Unpaired electrons are usually located on metal cations. The unpaired electrons can have both spin and orbital motion, which together generate a magnetic moment associated with the electrons. Magnetic behaviour is thus restricted mainly to compounds of transition metals and lanthanoids, most of which possess unpaired *d* and *f* electrons, respectively. Atomic nuclei also possess magnetic moments, but these are several orders of magnitude smaller than electron magnetic moments and are not considered further here. Magnetic oxides, especially ferrites such as $MgFe_2O_4$, are used in transformer cores, magnetic recording and information storage devices.

Several magnetic effects are possible, as shown for a schematic 1D crystal in Fig. 9.1. The unpaired electron spins may be oriented at random on the different atoms, in which case the material is *paramagnetic* (a), or they may interact with each other leading to *cooperative magnetic phenomena*. The spins may be aligned parallel, in which case the material possesses an overall magnetic moment and is *ferromagnetic* (b), or they may be aligned antiparallel, giving zero overall magnetic moment and the material shows *antiferromagnetic* behaviour (c). If the alignment is antiparallel but with unequal numbers in the two orientations, a net magnetic moment results and the behaviour is *ferrimagnetic* (d). *Spin glass behaviour* is observed in certain cases where the prime ordering tendency is antiferromagnetic but full antiferromagnetic order is not able to propagate through the structure; instead, it is blocked by an arrangement where pairs of spins are ordered ferromagnetically (e). In *helimagnetic* behaviour, spins are not confined to be simply parallel or antiparallel but show a gradual angular rotation between adjacent spins, giving a helical magnetic structure (f).

Clearly, there are strong analogies between these magnetic properties and corresponding electrical properties such as ferroelectricity (see Chapter 8). One difference, of course, is the absence of the magnetic monopole, which would be the magnetic equivalent of an electrical charge, as possessed by an ion or an electron.

The theory of magnetic behaviour is rather complicated, with a plethora of terms, symbols and units, and there are different ways of evaluating the magnetic moments of ions containing unpaired electrons which involve the use of different parameters and approximations. For some insulating and semiconducting materials, it is appropriate to use a localised model for the unpaired electrons, whereas in metallic materials, a

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

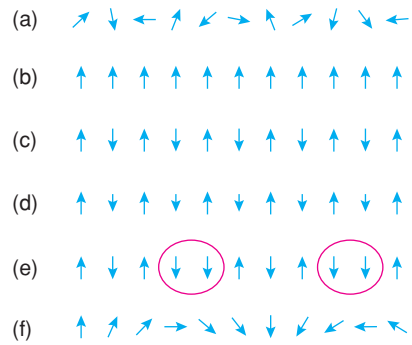


Figure 9.1 ***Schematic magnetic phenomena in a 1D crystal: (a) paramagnetism; (b) ferromagnetism; (c) antiferromagnetism; (d) ferrimagnetism, (e) spin glass behaviour in which an antiferromagnetic array is disrupted or frustrated by enforced ferromagnetic coupling (circled); (f) helimagnetism.***

band model is more appropriate. Here we shall use sufficient terminology and theory to be able to appreciate the different kinds of magnetic behaviour and how they relate to crystal structure.

9.1.1 Behaviour of substances in a magnetic field

When a substance is placed in a magnetic field, H, the density of lines of force in the sample, known as the *magnetic induction* or *magnetic flux density*, B, is related to H by the *permeability*, μ:

$$B = \mu H \tag{9.1}$$

$$B = \mu_0 H + \mu_0 M \tag{9.2}$$

where μ_0 is the *permeability of free space*, $4\pi \times 10^{-7}$ H m^{-1} (H is the symbol for henry) and M is the *magnetisation* or *magnetic moment* of the sample. $\mu_0 H$ is the induction generated by the field alone and $\mu_0 M$ is the additional induction contributed by the sample. The SI unit of magnetic field is amps per metre, A m^{-1}, since magnetic fields are produced whenever an electric current flows in a wire or coil; the unit of B is the tesla, T.

The *magnetic susceptibility*, χ, is defined as the ratio of magnetisation to field:

$$\chi = \frac{M}{H} \tag{9.3}$$

and therefore

$$\mu = \mu_0(1 + \chi) \tag{9.4}$$

The ratio μ/μ_0, which equals $1 + \chi$, is known as the *relative permeability*, μ_r, and is unitless. The susceptibility, which is also unitless, is the main parameter that is usually considered in the characterisation of magnetic properties and basically provides a measure of the response of a sample to an applied magnetic field. It is somewhat analogous to permittivity and polarisability, which refer to the responses in an electrical field.

Table 9.1 ***Magnetic susceptibilities***

Behaviour	Typical χ value	Change of χ with increase in temperature	Field dependence?
Diamagnetism	-8×10^{-6} for Cu; -1 for superconductors	None	No
Paramagnetism	0.1–0.001 for transition metal compounds	Decreases	No
Pauli paramagnetism	8.3×10^{-4} for Mn	None	Yes
Ferromagnetism	5×10^{3} for Fe	Decreases	Yes
Antiferromagnetism	$0–10^{-2}$	Increases	(Yes)

The different kinds of magnetic behaviour may be distinguished by the values of χ and μ_r and by their temperature and field dependences, Table 9.1. For *diamagnetic* substances, χ is very small and slightly negative. Diamagnetism is associated with orbital motion of electrons in atoms. This orbital motion generates a small electric field and in the presence of an external field, the orbital motion is modified slightly to give a magnetic moment that opposes the applied field leading to a slight repulsion effect which is explained by Lenz's law of electromagnetism. Superconductors represent a special, extreme type of diamagnetism since they repel magnetic fields completely (see Section 8.3.2). Since B inside a superconductor is zero, then, from equation (9.2), $H = -M$, and from equation (9.3), $\chi = -1$.

For *paramagnetic* substances, χ is small and positive. Thus, when placed in a magnetic field, the number of lines of force passing through a substance is greater if it is paramagnetic and slightly less if it is diamagnetic than would pass through a vacuum, Fig. 9.2. Consequently, paramagnetic substances are attracted by a magnetic field whereas diamagnetic substances experience a slight repulsion. Since superconductors show

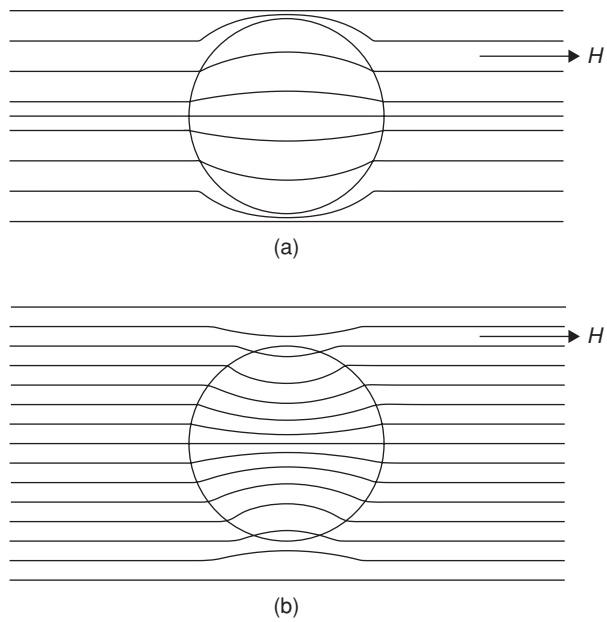


Figure 9.2 ***Variation of flux density or number of lines of force in (a) diamagnetic and (b) paramagnetic substances in a magnetic field.***

perfect diamagnetism, they expel magnetic fields completely, leading to the Meissner effect and applications in levitation.

In *ferromagnetic* substances, $\chi < 1$ and such materials are strongly attracted to a magnetic field. In *antiferromagnetic* substances, χ is positive and comparable to or somewhat less than that for paramagnetic substances.

9.1.2 Effects of temperature: Curie and Curie–Weiss laws

The susceptibilities of different kinds of magnetic material are distinguished by both their temperature dependences and their absolute magnitudes, Table 9.1. Ordered magnetic structures, whether ferro-, ferri-, antiferro-, heli-magnetic or spin glass, lose their ordered structures above a temperature known as either the Curie temperature, T_c (ferro- and ferri-magnets), or the Néel point, T_N (antiferro- and heli-magnets); the spins become disordered and the materials are therefore paramagnetic. At these transition temperatures, there is a balance between thermal energy, which tends to randomise the orientation of the magnetic moments, and internal exchange interactions, which try to maintain the cooperative aligned structures.

At T_c and T_N, an order–disorder transition occurs, therefore, as the structures transform from an ordered structure to the disordered paramagnetic structure. Many, but not all, paramagnetic substances undergo an ordering transition on cooling and evidence for whether or not this is likely to occur is contained in the temperature dependence of the magnetic susceptibility, χ. Paramagnetic substances may obey the simple Curie law, which states that magnetic susceptibility is inversely proportional to temperature:

$$\chi = \frac{C}{T} \tag{9.5}$$

where C is the Curie constant. Such a Curie response occurs when there is no spontaneous interaction between adjacent unpaired electrons. They do tend to align in a magnetic field, but with increasing temperature this alignment is more difficult and χ decreases, equation (9.5).

When there is some spontaneous interaction between adjacent spins, which may develop into an ordered magnetic structure at low temperatures, a better fit to the high-temperature behaviour in the paramagnetic region is often provided by the Curie–Weiss law:

$$\chi = \frac{C}{T - \theta} \tag{9.6}$$

where θ is the Weiss constant. These two types of behaviour are shown in Fig. 9.3, in which χ^{-1} is plotted against T. For paramagnetic substances that show no tendency to magnetic order, the plot extrapolates to 0 K. For paramagnetic substances that show a tendency to ferromagnetic order, there is already some local alignment of spins and χ is higher than for the simple paramagnetic case. Also, χ becomes infinite ($\chi^{-1} \rightarrow 0$) at a positive value of θ which coincides approximately with the *ferromagnetic Curie temperature*, T_c, below which the sample is ferromagnetic.

In paramagnetic substances that show a tendency to antiferromagnetic order, the χ values may be lower than in the simple paramagnetic case. This has the effect of displacing the Curie–Weiss plot to lower temperatures and the extrapolated value of θ is below 0 K. Obviously, such temperatures can never be reached; instead, on cooling, antiferromagnetism develops below T_N and departure from Curie–Weiss behaviour occurs.

The behaviour shown in Fig. 9.3 is for paramagnetic substances or for ferro-, ferri- and antiferro-magnetic substances in the paramagnetic state above T_c and T_N; it is somewhat idealised and departures from ideality do occur in practice. Below T_c and T_N, Curie–Weiss behaviour is not observed. Ferromagnetic substances

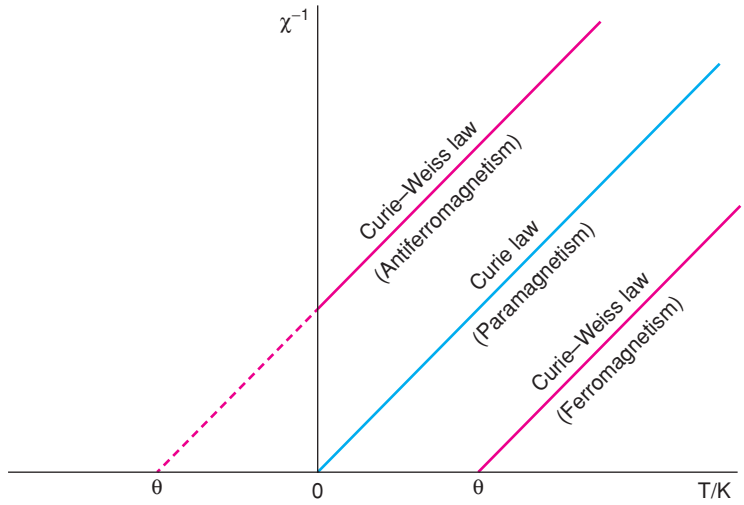


Figure 9.3 ***Reciprocal of susceptibility versus temperature for substances that are paramagnetic but may show magnetic ordering at low temperatures. Slope = C^{-1}.***

show very large susceptibilities at low temperatures, which decrease increasingly rapidly on heating towards T_c; an example is shown in Fig. 9.4(a). For antiferromagnetic substances, χ is very small at low temperatures in well-ordered structures but increases with increase in temperature (unlike for ferromagnets) as some thermally induced disorder is introduced into the antiferromagnetic state; χ passes through a maximum at T_N and then decreases at higher temperatures, following Curie–Weiss behaviour. A selection of Curie and Néel temperatures are summarised in Table 9.2.

A group of materials providing considerable current fascination are antiferromagnets that show *spin glass-like behaviour*. In these, the natural cooperative interaction between adjacent unpaired electrons leads to antiferromagnetism, but this is unable to develop into a full, long-range antiferromagnetic structure, i.e. the alternating 'spin up–spin down' arrangement becomes *frustrated* by the frequent occurrence of unavoidable spin up–spin up (or spin down–spin down) ferromagnetic pairs. An illustration of spin glass behaviour is shown in Fig. 9.1(e).

9.1.3 Magnetic moments

Magnetic properties are often expressed in terms of the magnetic moment, μ, since this relates directly to the number of unpaired electrons. Magnetic susceptibilities and moments may be determined experimentally using a Gouy balance. The sample is placed in the jaws of an electromagnetic and the variation in apparent sample mass is monitored as a function of applied field. For paramagnetic substances, unpaired electrons are attracted by the magnetic field and an apparent increase in mass occurs when the field is switched on. The measured susceptibility is corrected for various factors, including diamagnetism of the sample and sample holder. The approach generally used by spectroscopists and chemists to calculate magnetic moments is as follows.

The magnetic properties of unpaired electrons arise from two causes, electron spin and electron orbital motion. Of most importance is the spin component. An electron may usefully be visualised as a bundle of

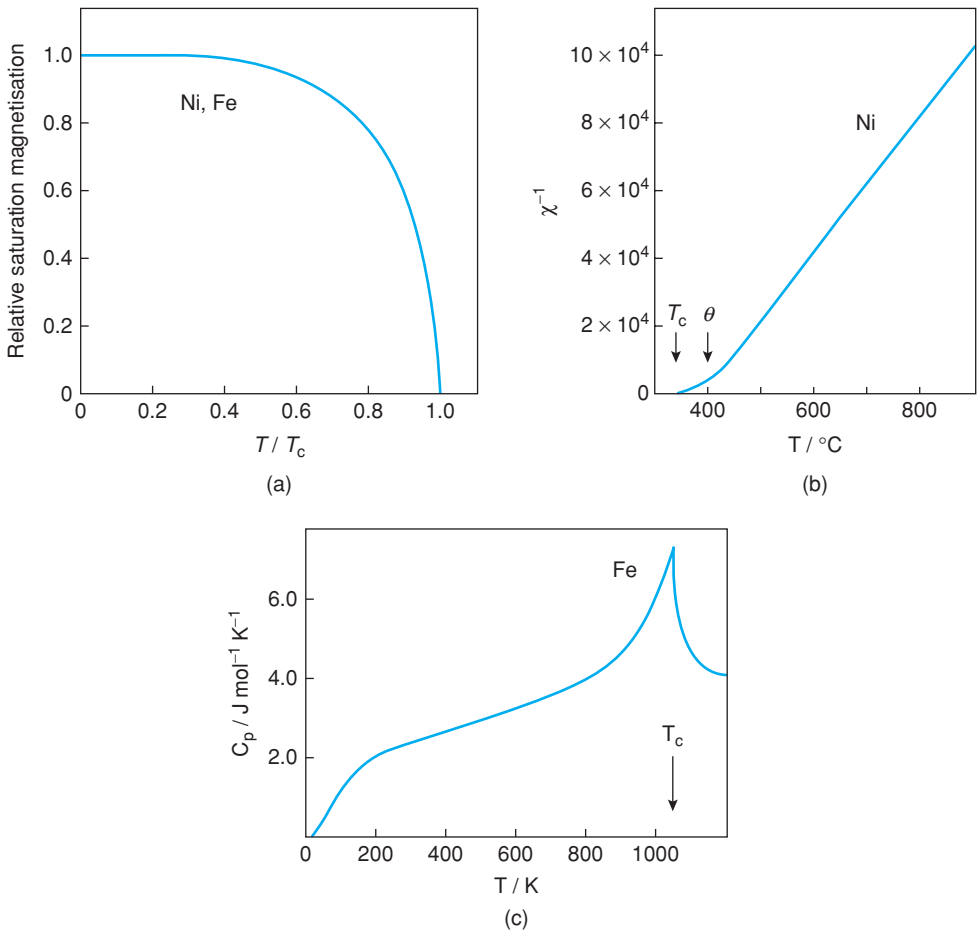


Figure 9.4 ***Some properties of ferromagnetic materials: (a) saturation magnetisation relative to the value at absolute zero, as a function of reduced temperature; (b) Curie–Weiss law plot for nickel of inverse susceptibility against temperature, showing a deviation close to T_c; (c) heat capacity of iron as a function of temperature. Data taken from C. A. Wert and R. W. Thomson,*** **Physics of Solids*****, © 1970 McGraw-Hill.***

negative charge spinning on its axis. The magnitude of the resulting spin moment, μ_s, is 1.73 BM (Bohr magneton); the Bohr magneton is defined as

$$1\ \text{BM} = \frac{eh}{4\pi mc} \tag{9.7}$$

where e is the electron charge, h is Planck's constant, m is the electron mass and c is the velocity of light. The equation used to calculate μ_s for a single electron is

$$\mu_s = g\sqrt{s(s+1)} \tag{9.8}$$

Table 9.2 ***Some Curie and Néel temperatures***

Element	T_c/K	T_N/K
Cr		308
Mn		100
Fe	1043	
Co	1404	
Ni	631	
Ce		12.5
Pr		25
Nd		19
Sm		14.8
Eu		90
Gd	293	
Tb	222	229
Dy	85	179
Ho	20	131
Er	20	84
Tm	25	56

where s is the *spin quantum number*, $^1/_2$, and g is the *gyromagnetic ratio*, ~2.00. Substituting for s and g gives $\mu_s = 1.73$ BM for a single electron. For atoms or ions that contain more than one unpaired electron, the overall spin moment is given by

$$\mu_S = g\sqrt{S(S+1)} \tag{9.9}$$

where S is the sum of the spin quantum numbers of the individual unpaired electrons. Thus, for high-spin Fe^{3+}, containing five unpaired 3*d* electrons, $S = 5/2$ and $\mu_s = 5.92$ BM. Calculated values of μ_s for different numbers of unpaired electrons are given in Table 9.3.

Table 9.3 ***Calculated and observed magnetic moments (BM) for some transition metal ions***

Ion	No. of unpaired electrons	$\mu_{S(calc)}$	$\mu_{S+L(calc)}$	$\mu_{(observed)}$
V^{4+}	1	1.73	3.00	~1.8
V^{3+}	2	2.83	4.47	~2.8
Cr^{3+}	3	3.87	5.20	~3.8
Mn^{2+}	5 (high spin)	5.92	5.92	~5.9
Fe^{3+}	5 (high spin)	5.92	5.92	~5.9
Fe^{2+}	4 (high spin)	4.90	5.48	5.1–5.5
Co^{3+}	4 (high spin)	4.90	5.48	~5.4
Co^{2+}	3 (high spin)	3.87	5.20	4.1–5.2
Ni^{2+}	2	2.83	4.47	2.8–4.0
Cu^{2+}	1	1.73	3.00	1.7–2.2

Source: data taken from F. A. Cotton and G. Wilkinson (1966), *Advanced Inorganic Chemistry: a Comprehensive Text*, 2nd edn, John Wiley & Sons, Inc., New York.

The motion of an electron around the nucleus may, in some materials, especially those containing heavy metal ions such as the lanthanoids, give rise to an *orbital moment* which contributes to the overall magnetic moment. In cases where the orbital moment makes its full contribution,

$$\mu_{S+L} = \sqrt{4S(S+1) + L(L+1)} \tag{9.10}$$

where L is the orbital angular momentum quantum number for the ion. Equations (9.8)–(9.10) are applicable to free atoms or ions. In practice, equation (9.10) may not hold because the orbital angular momentum is either wholly or partially *quenched* in cases where the electric fields on the surrounding atoms or ions restrict the orbital motion of the electrons. In such cases, therefore, the observed moments are similar to or somewhat larger than the calculated, spin-only, values, Table 9.3.

The methods outlined above to calculate magnetic moments have their origin in quantum mechanics. The details of the methods are actually fairly involved but, even so, agreement between theory and experiment is often not good, Table 9.3. An alternative, much simpler method is often used, especially when working with applications of ferro- and antiferromagnetism. In this, the magnetic moment of a single unpaired electron is set equal to 1 BM. For an ion that possesses n unpaired electrons, the magnetic moment is n BM. Thus high–spin Mn^{2+} and Fe^{3+} would both have a magnetic moment of 5 BM. This method may be quantified by the simple equation

$$\mu = gS \tag{9.11}$$

where $g \approx 2.00$ and S, the spin state of the ion, equals $n/2$. The values obtained in this way often underestimate the true values (compare columns 2 and 5 of Table 9.3), but do provide a rough and useful guide to the magnitude of μ. A modification is to allow g to become an adjustable parameter; by allowing g to exceed 2, a contribution from orbital momentum is effectively included. Thus, for Ni^{2+}, a g value in the range 2.2–2.3 is often used. In discussing the magnetic properties of ferrite phases, we shall use equation (9.11) to evaluate μ because of its simplicity.

9.1.4 Mechanisms of ferro- and antiferromagnetic ordering: superexchange

In the paramagnetic state, individual magnetic moments of ions containing unpaired electrons are oriented at random. Partial alignment occurs only on application of a magnetic field. The energy of interaction between dipoles and a magnetic field may be calculated readily, although details are not given here. It is generally greater than the thermal energy, kT, possessed by the ions or dipoles.

In the ferro- and antiferromagnetic states, alignment of magnetic dipoles occurs spontaneously. There must therefore be some positive energy of interaction between neighbouring spins that allows this to occur, either parallel or antiparallel. The origin of this coupling of spins or cooperative interaction is quantum mechanical. Qualitatively, the effect is understood, although a complete rationale for the behaviour of, for example, ferromagnetic Fe or Co requires the results of band structure calculations.

One process, by which spins couple to give antiferromagnetism in, e.g. NiO, is *superexchange*, Fig. 9.5. The Ni^{2+} ion has eight d electrons. In an octahedral environment, two of these occupy singly the e_g orbitals, d_{z^2} and $d_{x^2-y^2}$. These orbitals are oriented parallel to the axes of the unit cell and therefore point directly at adjacent oxide ions. The unpaired electrons in these e_g orbitals couple with electrons in the p orbitals of the O^{2-} ions. This coupling may involve the formation of an excited state in which the electron transfers from the e_g orbital of Ni^{2+} to the oxygen p orbital. Or it may be regarded as partial covalent bond formation involving pairing of electrons between adjacent Ni and O atoms. The p orbitals of the O^{2-} ion contain two electrons each, which are also coupled antiparallel. Hence, provided that Ni^{2+} and O^{2-} ions are sufficiently

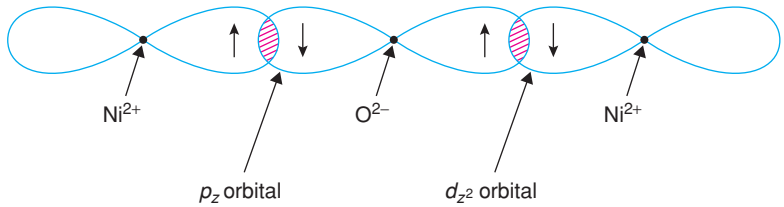


Figure 9.5 ***Antiferromagnetic coupling of spins of* d *electrons on* Ni^{2+} *ions through* p *electrons of oxide ions.***

close that coupling of their electrons is possible, a chain coupling effect may occur which propagates through the crystal structure, Fig. 9.5. The net effect is that neighbouring Ni^{2+} ions, separated by intervening O^{2-} ions, are coupled antiparallel.

9.1.5 Some more definitions

Ferromagnetic materials have a *domain* structure, similar to the domain structure of ferroelectrics. Within each domain, the spins align parallel, but unless the material is in the *saturation condition*, different domains have different spin orientations.

The response of ferromagnetic materials to an applied magnetic field, H, is similar to that of ferroelectrics in an applied electric field. A *hysteresis loop* occurs in the plot of magnetisation, M, or induction, B, against H, Fig. 9.6. A similar loop, Fig. 8.40(e), is observed for the polarisation of a ferroelectric plotted against voltage. During magnetisation and demagnetisation in an alternating magnetic field, energy is dissipated, usually as heat. During one complete cycle, this amount of energy, the *hysteresis loss*, also known as the BH product, is proportional to the area inside the hysteresis loop. The *remanence*, B_r or M_r, is the residual magnetisation at zero applied field; the *saturation magnetisation*, M_s, is the maximum magnetisation achievable with an applied field.

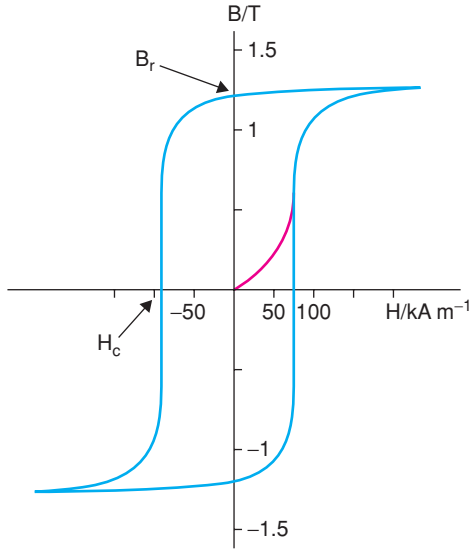


Figure 9.6 ***Rectangular hysteresis loop showing coercivity,* H_c*, and remanence,* B_r *or* M_r*.***

Table 9.4 ***Some soft and hard magnetic materials***

Material	Coercivity/kA m^{-1}	Saturation magnetisation/kA m^{-1}	Curie temperature/°C
γ-Fe_2O_3	~25	~370	600
Co-coated γ-Fe_2O_3	~50	~370	
CrO_2	~60	~500	128
Fe powder	~120	~1700	

An additional source of energy loss in an alternating magnetic field is associated with electrical currents called *eddy currents* that are induced in the material. The varying magnetic field induces a varying voltage and the eddy current losses are given by I^2R or V^2/R [energy loss is given by $IV = (V/R) \times V$, where V is the voltage and I the current). Eddy currents are therefore minimised in highly resistive materials. Alloys such as Fe with 2% Si have much higher resistivity than pure metals. One advantage that most magnetic oxides have over metals is their much higher electrical resistance.

Materials that are magnetically *soft* are those of low *coercivity*, H_c. The coercivity is the reverse field required to achieve demagnetisation. Soft materials also have low *permeability* and a hysteresis loop that is 'narrow at the waist' and of small area. The use of soft materials depends on the frequency, whether *dc* or *ac*. For *dc* applications, a large M_s is needed. With increasing frequency, low H_c becomes important. At high frequency, if it is necessary to reduce eddy current losses, then high resistivity is necessary. Materials that are magnetically *hard* are those with a high H_c and a high M_r or B_r. The latter is the magnetisation that remains after the field has been switched off. Hard materials are not easily demagnetised, therefore, and find use as permanent magnets. Some materials, and their characteristics, that are used for soft and hard applications are given in Table 9.4.

Ferromagnetic materials have a preferred or 'easy' direction of magnetisation; in Fe, this is parallel to the axes of the cubic unit cell, Fig. 9.7. The *magnetocrystalline anisotropy* is the energy required to rotate the magnetisation out of this preferred direction.

Most magnetic materials exhibit *magnetostriction*, i.e. they change their shape on magnetisation. For instance, Ni and Co both contract in the direction of magnetisation but expand in the perpendicular direction. With Fe the reverse happens at low fields, but at high fields Fe behaves like Ni and Co. The dimensional changes involved are small. The coefficient of magnetostriction, λ_s, defined as $\lambda_s = \Delta l/l_0$, increases with H up to a maximum value in the range $1–60 \times 10^{-5}$ for saturation magnetisation. The effect is therefore comparable to changing the temperature of the material by a few degrees.

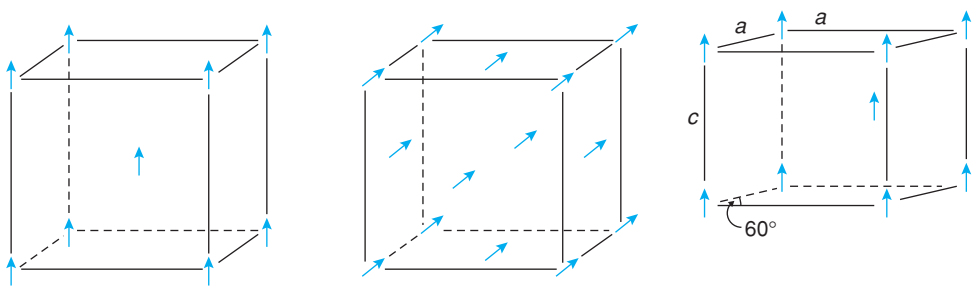


Figure 9.7 ***Ferromagnetic ordering in* bcc *α-Fe,* fcc *Ni and* hcp *Co.***

9.2 Magnetic Materials, their Structures and Properties

9.2.1 Metals and alloys

Five transition elements – Cr, Mn, Fe, Co, Ni – and most lanthanoids are either ferro- or antiferromagnetic, Table 9.2. Several of the lanthanoids are helimagnetic: Dy, Tb, Ho and Er. Some pass through a sequence of magnetically ordered phases, e.g. for Tb, $T_c = 220$ K and $T_N = 230$ K; for Dy, $T_c = 85$ K and $T_N = 180$ K. Hence these metals change from ferromagnetic to helimagnetic to paramagnetic on heating. Many alloys and intermetallic compounds also show some kind of magnetic order.

Fe, Co and Ni are ferromagnetic, as shown in Fig. 9.7. In *bcc* α-Fe, the spins point in a [100] direction, parallel to a cubic cell edge. In *fcc* Ni, they point in a [111] direction, parallel to a cube body diagonal. Co has an *hcp* structure and the spins are oriented parallel to the *c* axis of the unit cell. These examples demonstrate clearly that ferromagnetism is not associated with a particular type of crystal structure!

Cr and Mn are both antiferromagnetic at low temperatures. Mn has a complex structure, but Cr is *bcc*, similar to α-Fe. In Cr the spins are arranged antiparallel along one of the cube unit cell directions.

Some characteristics of ferromagnetic materials are given in Fig. 9.4. In (a) is shown the temperature dependence of the magnetic susceptibility or magnetic moment. The vertical axis represents the saturation magnetisation, relative to its maximum value at absolute zero. The horizontal axis has a 'reduced temperature' scale in which the ratio of the actual temperature to the Curie temperature is plotted. Therefore, at the Curie point, $T/T_c = 1$. Use of reduced axes such as these facilitates comparison between materials with different Curie points and different magnetic moments. When plotted in this way, Fe and Ni are found to behave similarly: with increasing temperature the saturation magnetisation remains almost constant at small T/T_c and decreases increasingly rapidly as T_c is approached.

Above the Curie point, Fe, Co and Ni are paramagnetic. At temperatures well above T_c, Curie–Weiss law behaviour is observed but deviations occur close to T_c, Figure 9.4(b). These deviations are attributed to short-range order between the spins. The long-range order of the ferromagnetic state is lost but residual, short-range order remains just above T_c; hence the Weiss temperature, θ, is somewhat removed from T_c. Data are shown for Ni; Fe and Co behave similarly.

The transition from ferromagnetic to paramagnetic behaviour at T_c has many characteristics of a second-order or lambda phase transition and is a classic example of an order–disorder phase transition. Perfect order is attainable only at absolute zero; at all real temperatures, disorder is present and increases rapidly with increase in temperature. This is shown by the heat capacity, which passes through a maximum at T_c, Figure 9.4(c).

At this stage, let us give an airing to a phenomenon which does not have a simple, intuitive explanation but can be rationalised once the results of electronic structure calculations are known. One of the mysteries of ferromagnetism concerns its dependence on position in the Periodic Table and, in particular, the question of how many unpaired electrons are available to contribute to ferromagnetism. The facts are as follows. Provide your own explanation!

The three ferromagnetic elements in the first transition series have the electronic configurations shown in Table 9.5. Column 2 gives the configuration of the free atom in its ground state; the 4*s* level is fully occupied in each case. In the ferromagnetic state (column 4), the 4*s* band is not full but some electrons are transferred to the 3*d* band. Evidence for this comes from the value of the saturation magnetisation which is proportional to the number of unpaired spins (column 3). Thus, Fe has a net moment of 2.2 BM per atom and hence has, on average, 2.2 unpaired *d* electrons per Fe atom, i.e. of the 7.4 *d* electrons, 4.8 are of one sign and 2.6 of the other.

So far, we are on familiar ground because chemists are used to the idea that electrons are transferable between 4*s* and 3*d* levels, depending on circumstances. The sticking point concerns the maximum number of

Table 9.5 ***Electronic constitution of iron, cobalt and nickel***

Metal	Atomic ground state configuration	Ferromagnetic state	
		No. of unpaired spins	Configuration
Fe	d^6s^2	2.2	$d^{7.4}s^{0.6}$
Co	d^7s^2	1.7	
Ni	d^8s^2	0.6	

unpaired electrons that are available to contribute to ferromagnetism in elements or alloys of the 3*d* series. The answer, apparently, is 2.4 per atom. The effective number of unpaired electrons varies with total electron content across the 3*d* series. The maximum value of 2.4 is found for an alloy of composition $Fe_{0.8}Co_{0.2}$. With increasing total electron content, the number of unpaired spins decreases gradually, passing through Co and Ni before decreasing to zero in the alloy $Ni_{0.4}Cu_{0.6}$. Thus pure Cu is paramagnetic. To the other side of the $Fe_{0.8}Co_{0.2}$ composition, the number of unpaired spins also decreases systematically, passing through Fe, Mn and Cr. Both Mn and Cr are antiferromagnetic at low temperatures.

Some insight into the origins of ferromagnetism is obtained from the calculated density of states, $N(E)$, in the region of the Fermi energy, E_F. $N(E)$ depends on the nature and degree of overlap of the orbitals forming the particular energy band. In the first row transition elements, there are two energy bands to consider resulting from 3*d* and 4*s*, 4*p* levels. 4*s*, 4*p* orbitals in particular are diffuse, or extended, and overlap strongly with those on adjacent atoms, giving rise to a broad band of levels with a wide distribution of energies and a relatively small number of levels in the region of E_F. Electrons in 3*d* orbitals are held more tightly to the individual atomic nuclei with increasing atomic number, due to the incomplete shielding of 3*d* electrons from the (increasing) nuclear charge. This leads to a limited degree of orbital overlap and a narrow *d* band with high $N(E)$, as shown schematically in Fig. 9.8 for a transition element with partly-filled 3*d* and 4*s*, 4*p* bands.

Two factors influence the number of unpaired *d* electrons: first, the position of the Fermi level, and in particular, whether it is located primarily in the middle of the 4*s*, 4*p* band or instead, is in the 3*d* band (for

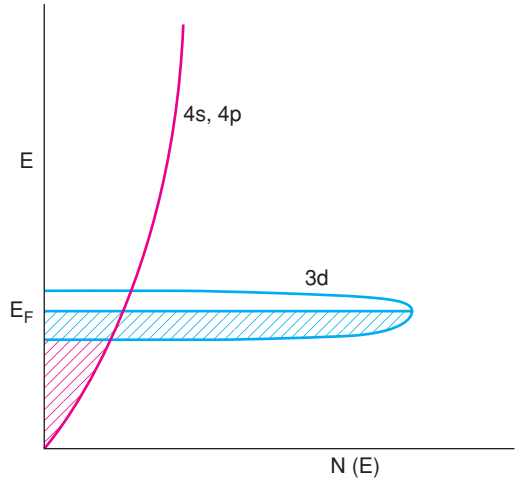


Figure 9.8 ***Occupied energy levels (shaded) and density of states N(E) for 3d and 4s, 4p bands in a first row transition element such as Fe.***

first row transition elements); second, the density of states at E_F. A high $N(E)$ arises if there is limited orbital overlap giving rise to a narrow band with closely spaced energy levels. In order for *d* electrons to be unpaired, they must be in singly occupied levels close to E_F; a high concentration of unpaired electrons, which is necessary for the cooperative interactions required for ferromagnetism, therefore requires a narrow *d* band with E_F located in the middle of the band, as shown schematically in Fig. 9.8.

These two factors, *d* band width and the position of E_F, appear to control the magnetic properties. For early transition elements, E_F is located in the overlapping 4*s*, 4*p* band with some contribution from a broad 3*d* band and the unpaired *d* electron concentration is small. For late transition elements, the 3*d* band is effectively full and E_F is at higher energies, primarily in the 4*s*, 4*p* band. For Fe, Co and Ni, E_F is in the middle of the 3*d* band, with some contribution from 4*s*, 4*p* levels, ensuring an optimised number of unpaired 3*d* electrons, which results in cooperative electron–electron interactions between electrons with parallel spins, leading to ferromagnetism.

For second and third row transition elements, the 4*d* and 5*d* bands are more diffuse, giving insufficient numbers of unpaired *d* electrons at E_F to allow ferromagnetic interactions in these elements.

The lanthanoid elements have magnetically ordered structures associated with unpaired 4*f* electrons. Exceptions occur when the 4*f* shell is empty: La, $4f^0$; or full: Yb, $4f^{14}$; Lu, $4f^{14}$. Most lanthanoids apart from Gd are antiferromagnetic below room temperature. Some, especially the later lanthanoids, form ferro- and antiferromagnetic structures at different temperatures. For these, the sequence with decreasing temperature is always

$$\text{paramagnetic} \rightarrow \text{antiferromagnetic} \rightarrow \text{ferromagnetic}$$

Some Curie and Néel temperatures are given in Table 9.2.

Several antiferromagnetic lanthanoids exhibit *metamagnetism* and switch to a ferromagnetic state in a suitably high magnetic field. For instance, Dy is ferromagnetic below 85 K but antiferromagnetic at high temperatures. By application of a magnetic field, the ferromagnetic state can be preserved above 85 K and up to the Néel temperature, 179 K.

Some rare earth-containing alloys are of great current interest as they form the strongest permanent magnets known to date. Particular examples are $SmCo_5$ and $Nd_2Fe_{14}B$. The origin of their ferromagnetism appears to involve both the *f* and *d* electrons on the lanthanoids, which are able to interact cooperatively. Since with the lanthanoids there are up to seven 4*f* orbitals that can, potentially, contain unpaired electrons, in addition to some of the 5*d* orbitals, much higher values for magnetic moment and magnetisation are possible with lanthanoids than with transition metals.

Electronic structure calculations give insight into band widths and densities of states, especially at E_F, and allow a qualitative explanation as to why magnetism is associated with transition elements and lanthanoids and is not shown by, for example, *s* block and *p* block metals. In these metals, *s* and *p* valence orbitals overlap strongly, giving broad *s* and *p* bands with fully delocalised *s* and *p* electrons. Because the density of states close to E_F is small in comparison with that of narrow 3*d* bands in metals, there are insufficient unpaired electrons in the energy levels close to E_F to give rise to significant ferromagnetic interactions. Under normal circumstances, therefore, *s* and *p* block metals are not ferromagnetic.

In addition to the strong ferro- and antiferromagnetic coupling shown by some transition metals, most metals display a weak paramagnetism in the presence of a magnetic field, known as *Pauli paramagnetism*, that is essentially independent of temperature and does not show the expected Curie law behaviour, equation (9.5). This can be understood using Fig. 9.9 in which (a) the outermost electrons are in a partially-occupied band. All energy levels apart from those close to E_F are fully occupied with spin-up and spin-down electrons and we may imagine that these occupy spin-up and spin-down sub-bands as shown in (a). In the absence of a magnetic field, the two sub-bands have equal energy; in the absence of cooperative magnetic interactions,

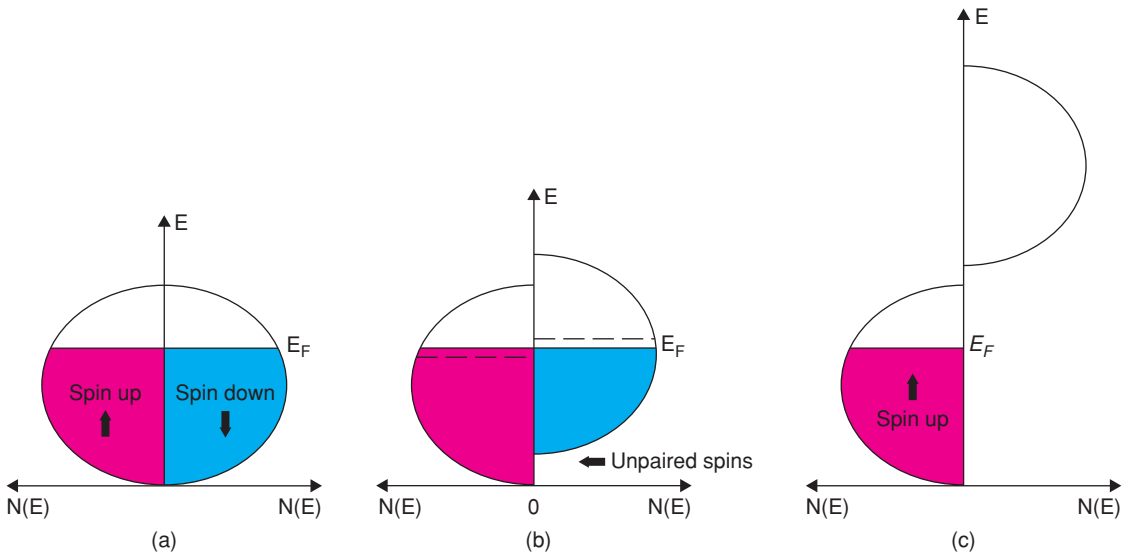


Figure 9.9 ***Schematic splitting of 3d band into two sub-bands: (a) in the absence of a magnetic field; (b) in Pauli paramagnets, an energy difference develops in a magnetic field; (c) in ferromagnetic transition metals, the splitting occurs spontaneously.***

discussed later, there are equal numbers of electrons in the two sub-bands and the metal has no overall magnetic moment.

The effect of applying a magnetic field is to lower the energy of the spin-up electrons aligned parallel to H and raise the energy of the spin-down electrons antiparallel to H, as shown in Fig. 9.9(b). Consequently, spin-down electrons close to E_F, which have slightly higher energy than spin-up electrons close to E_F, spill over between the sub-bands to equalise E_F. The electrons that spill over change their spin orientation, giving rise to a net small spin-up magnetic moment. This leads to an itinerant conduction band type of magnetism known as Pauli paramagnetism. Only a small fraction of the conduction band electrons close to E_F contribute to Pauli paramagnetism, in the same way that few electrons contribute to electronic conductivity and heat capacity. Pauli paramagnetism is mainly independent of temperature because of the small number of unpaired spins in levels close to E_F but is field dependent because of the splitting of the sub-bands into different energies.

Pauli paramagnetism assumes no cooperative interaction between unpaired electron spins. The difference between Pauli paramagnetism and ferromagnetism is that the latter shows a strong exchange interaction in which unpaired electrons align cooperatively even in the absence of an external magnetic field. This has a big effect on the relative energies of the spin-up and spin-down sub-bands, known as *exchange splitting*, but again, the net magnetisation is determined by the difference in occupancy of spin-up and spin-down energy levels. In extreme cases, the exchange splitting leads to complete separation of the sub-bands, as shown in Fig. 9.9(c).

9.2.2 Transition metal monoxides

The oxides of the first transition series show large, systematic changes in properties with atomic number and number of *d* electrons. The divalent oxides, MO, show a complete range of electrical properties, from metallic conductivity in TiO to insulating behaviour in NiO (see Chapter 8). They also show a wide range of magnetic

properties, which parallel the range of electrical properties. TiO, VO and CrO have high conductivity because their *d* electrons are delocalised in a partly-filled t_{2g} band, and the absence of localised, unpaired electrons leads to diamagnetism and Pauli paramagnetism. By contrast, oxides of the later elements, MnO, FeO, CoO and NiO, have localised *d*-electrons which are more tightly held to the individual nuclei and the oxides are paramagnetic, in addition to semiconducting or insulating.

Interesting magnetic effects occur on cooling these latter oxides: cooperative interactions between unpaired electrons on adjacent ions lead to ordered magnetic structures. MnO, FeO, CoO and NiO are all antiferromagnetic below their respective Néel temperatures, T_N, of –153, –75, –2 and +250 °C. All have similar rock salt structures which are cubic at high temperatures and undergo a slight rhombohedral distortion on cooling. In NiO and MnO, a slight contraction of the structure occurs along one of the threefold axes parallel to [111]; in FeO the structure is, instead, slightly elongated along [111] [see Fig. 1.8(c) and discussion of trigonal symmetry].

The cause of the rhombohedral distortion is antiferromagnetic ordering of the M^{2+} ions, which may be explained by the superexchange mechanism, Fig. 9.5. In the rock salt structure, alternate layers of M^{2+} and O^{2-} ions occur parallel to the (111) planes and, therefore, perpendicular to [111]. Within a given layer of Ni^{2+} ions in NiO, the spins of all the Ni^{2+} ions are aligned parallel, whereas in adjacent layers, they are all antiparallel, Fig. 5.25. Ordered magnetic structures such as these may be studied very elegantly by neutron diffraction since neutrons, which possess a magnetic moment, are scattered by unpaired electrons and also by atomic nuclei. Further details are given in Section 5.5.2.

9.2.3 Transition metal dioxides

A similar pattern of behaviour occurs to that seen in the monoxides, although the series is much less complete and only three members, TiO_2, CrO_2 and MnO_2, have the same rutile structure, which permits an easy comparison of structure–property correlations. To the left of Cr in the Periodic Table, the 3*d* orbitals overlap to form a band (TiO_2 is insulating because it has no 3*d* electrons, but on reduction to form TiO_{2-x} it rapidly becomes metallic, showing that good 3*d* orbital overlap occurs). To the right of Cr, the 3*d* orbitals are localised and MnO_2 is a paramagnetic semiconductor. CrO_2 is intermediate; it has a very narrow 3*d* band due to limited orbital overlap, which gives rise to ferromagnetism with $T_c = 392$ K. This example shows very convincingly a connection between magnetic and electrical properties. In particular, cooperative magnetic phenomena, especially ferromagnetism, are associated with narrow *d*-band structures that represent the cross-over between, on the one hand, itinerant electrons and high electronic conductivity and, on the other, localised electrons associated with semiconductivity or insulating behaviour.

9.2.4 Spinels

The commercially important magnetic spinels, known as ferrites, have formula MFe_2O_4, where M is a divalent ion such as Fe^{2+}, Ni^{2+}, Cu^{2+}, Mg^{2+}. They are all *inverse* (see Chapter 1), either partially or completely. This is probably because Fe^{3+}, being a d^5 ion, has no crystal field stabilisation energy (CFSE) in an octahedral site (see Section 3.2.15); hence, the large divalent ions preferentially occupy octahedral sites and Fe^{3+} is distributed over both tetrahedral and octahedral sites. The spinel Fe_3O_4, magnetite or lodestone, was the material in which magnetism was discovered many centuries ago in China and used in the development of compass needles as the first practical application of magnetism.

Magnetite is an inverse spinel and its formula may be written as $[Fe^{3+}]^{tet}[Fe^{2+}, Fe^{3+}]^{oct}O_4$. It is antiferromagnetic with $T_N = 119$ K and, because it contains mixed valence Fe on octahedral sites, it is a very good electronic conductor. A closely related iron oxide, γ-Fe_2O_3, also has a spinel structure but is Fe deficient since the Fe:O ratio is 2:3 instead of the expected 3:4. It contains $21\frac{1}{3}$ Fe^{3+} ions distributed over

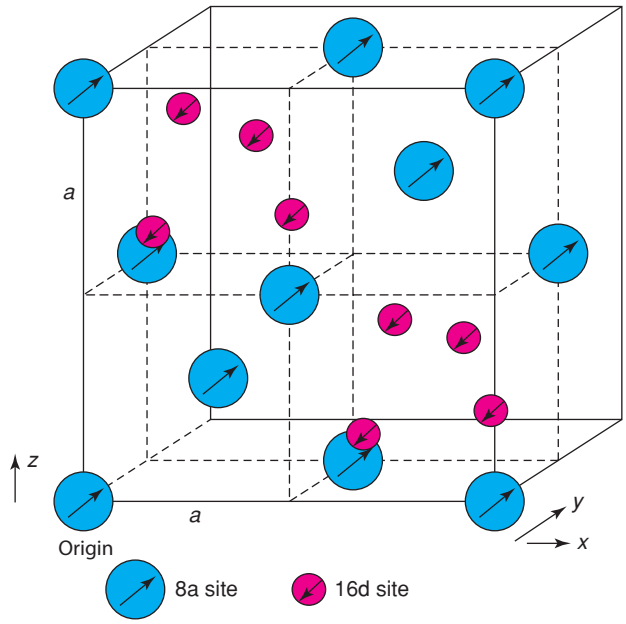


Figure 9.10 ***Magnetic structure of antiferromagnetic and ferrimagnetic spinels. Cations only are shown. Note the antiparallel spins on 8a and 16d sites which form alternate layers perpendicular to [111]. Cations in the front four octants only of the unit cell are shown.***

the 24 tetrahedral and octahedral sites that are usually occupied in spinel. It readily oxidises to magnetite but the reaction is reversible and γ-Fe_2O_3 is recovered on heating magnetite in vacuum at 250 °C. γ-Fe_2O_3 is particularly important in modern magnetic information storage media.

Ferrites are either antiferromagnetic or ferrimagnetic because the magnetic spins of ions on tetrahedral, 8a sites are antiparallel to those on octahedral, 16d sites. The contents of the front four octants of the unit cell are shown in Fig. 9.10; the origin coincides with an 8a ion. Oxide ions are not shown. When the unit cell is drawn in this way, it may be described as an *fcc* array of 8a ions, at corner and face centre positions, with additional 8a ions in the centre of alternate octants of the unit cell. This gives, overall, eight 8a ions per unit cell. The 16d ions are arranged tetrahedrally inside the other four octants, four per octant, giving sixteen 16d ions per unit cell. The magnetic spins of 8a and 16d ions are antiparallel, as shown. Let us now calculate the magnetic moments, μ, of different spinels using equation (9.11).

$ZnFe_2O_4$ is a good example to begin with. It is inverse at very low temperatures (although normal at room temperature) i.e.

$$[Fe^{3+}]^{tet}[Zn^{2+},\ Fe^{3+}]^{oct}O_4$$

Since there are equal numbers of Fe^{3+} ions on 8a and 16d sites, with opposed spins, the net moment of the Fe^{3+} ions is zero. As Zn^{2+} has no magnetic moment, $ZnFe_2O_4$ has zero overall μ but is antiferromagnetic because the two sets of Fe^{3+} ions have antiparallel spins, Fig. 9.10. Experimental measurements confirm this ($T_N = 9.5$ K).

A similar result might be expected for $MgFe_2O_4$, but it has a residual overall μ and is ferrimagnetic. Two explanations are possible: either the spinel is not completely inverse and there are more Fe^{3+} ions on the

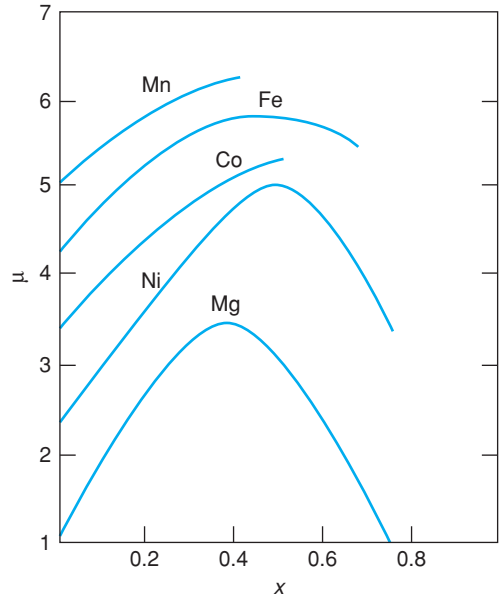


Figure 9.11 ***Variation of magnetic moment with composition for ferrite solid solutions $M_{1-x}Zn_xFe_2O_4$. Reproduced with permission from Gorter,* Nature, *165, 798, © 1950 Macmillan Publishers Ltd.***

octahedral sites than on the tetrahedral sites, giving only partial cancellation of the spins, or the effective μ per Fe^{3+} ion is different for the two types of site. Experimental measurements indicate the first explanation to be correct. With increasing temperature, $MgFe_2O_4$ gradually transforms to the normal spinel structure. The *degree of inversion* at room temperature depends on thermal history, especially the rate of cooling from high temperature after synthesis. Thus, quenched samples have a lower degree of inversion and hence higher μ than slowly cooled samples.

Manganese ferrite, $MnFe_2O_4$, is 80% normal, 20% inverse, but since Mn^{2+} and Fe^{3+} are both d^5, μ is insensitive to the degree of inversion and to thermal history effects. $MnFe_2O_4$ is ferrimagnetic with $\mu \approx 5$ BM.

An interesting example of the importance of cation site occupancies and solid solution effects is provided by the mixed ferrites, $M_{1-x}Zn_xFe_2O_4$: M = Mg, Ni, Co, Fe, Mn. These are largely inverse for $x = 0$, i.e.

$$[Fe^{3+}]^{tet}[M^{2+},\ Fe^{3+}]^{oct}O_4$$

Expected values of μ for purely inverse spinels are Mg 0, Ni 2, Co 3, Fe 4 and Mn 5. Experimental values are larger, as shown by the left-hand axis of Fig. 9.11, but this may represent the oversimplified way in which μ is calculated, equation (9.11). Zinc ferrite, $ZnFe_2O_4$, $x = 1$, is, by contrast, almost entirely normal at room temperature. However, the spins of the Fe^{3+} ions on the octahedral sites are random and $ZnFe_2O_4$ is paramagnetic. On formation of ferrite solid solutions by partial replacement of M^{2+} by Zn^{2+}, a gradual change from inverse to normal behaviour occurs. Introduction of Zn^{2+} into the tetrahedral sites causes Fe^{3+} ions to be displaced onto the octahedral sites, i.e.

$$\left[Fe^{3+}_{1-x}Zn^{2+}_{x}\right]^{tet}\left[M^{2+}_{1-x}Fe^{3+}_{1+x}\right]^{oct}O_4$$

If the solid solutions retained the antiferromagnetic character of the MFe_2O_4 ferrites, $x = 0$, a linear increase in μ should occur and attain the value of 10 for $x = 1$, $ZnFe_2O_4$. Long before $x = 1$ is reached, however, the antiferromagnetic coupling between 16d and 8a sites is destroyed and M values decrease, Fig. 9.11. For small values of x, experimental values of M increase, as expected, consistent with retention of antiferromagnetic/ferrimagnetic ordering, but pass through a maximum for $x = 0.4$–0.5.

In addition to the magnetic moment μ, other parameters affect the magnetic properties of ferrites. These include the saturation magnetisation, M_{sat}, the magnetostrictive constant, λ_s, the permeability, μ, and the magnetocrystalline anisotropy constant, K_1. Suffice it to say that the values of these parameters vary widely. A particular ferrite may be chosen depending on the values of these properties and the application in mind. Variation in magnetic properties may be achieved by making mixed ferrites. For instance, partial replacement of Mn^{2+} by Fe^{2+} in $MnFe_2O_4$, giving a solid solution $Mn^{2+}_{1-x}Fe^{2+}_xFe^{3+}_2O_4$, reduces the magnetic anisotropy to zero. This parameter measures the ease of reorientation of the magnetic moment in an applied magnetic field. Reduction of the magnetic anisotropy yields increased permeability, which is usually desired in commercial ferrites. An undesirable side-effect, however, is an increase in electrical conductivity with increasing Fe^{2+} content.

9.2.5 Garnets

Garnets are a large family of complex oxides, some of which are important ferrimagnetic materials. They have the general formula $A_3B_2X_3O_{12}$ and are described in Section 1.17.13. A is a large ion with radius ~1 Å and has a coordination number of eight in a distorted cubic environment. B and X are smaller ions in octahedral and tetrahedral sites, respectively. The garnets with interesting magnetic properties have A = Y, Sm, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu; B, X = Fe^{3+}. One of the most important is yttrium iron garnet (YIG), $Y_3Fe_5O_{12}$. Many other A, B, X combinations are possible, some of which are described in Section 1.17.13.

Crystallographic data for YIG are given in Fig. 1.49. The *bcc* cell is large, $a = 12.376$ Å, and contains eight formula units. The structure consists of a framework of corner-sharing BO_6 octahedra and XO_4 tetrahedra with A ions in eight-coordinate cavities. In YIG and the rare earth garnets, the B and X ions are the same, Fe^{3+}.

YIG and the rare earth garnets are ferrimagnetic with Curie temperature in the range 548–578 K. In order to evaluate μ, the three types of ion, on sites 24c, 24d and 16a, have to be considered. The spins of the 24d ions are coupled antiparallel to those of both the 24c and 16a ions. If we consider the two types of Fe^{3+} ion first, their spins partially cancel, since there are 16 of one kind and 24 of the other kind in the unit cell. This gives a net moment of one Fe^{3+} ion per formula unit $M_3Fe_5O_{12}$, i.e. 5 BM. Since Y^{3+} is a d^0 ion, $\mu = 0$. Consequently, the μ of YIG is 5 BM.

For the rare earth garnets, the overall moment is expected to be $(3\mu_M - 5)$ BM, where μ_M is the moment of the 24c ion, of which there are three per formula unit. For Gd^{3+}, which is an f^7 ion, $\mu_{Gd} = 7$ BM and the net moment for GdIG is calculated to be 16 BM, which agrees with experiment. For Lu^{3+}, which is f^{14}, $\mu_{Lu} = 0$ and the net moment is 5 BM, in agreement with experiment. For the ions Tb to Yb, it appears that the orbital moment is not completely quenched and the μ_M values are larger than given by the spin-only formula with $g = 2.00$. A comparison between experimental and calculated values is shown in Fig. 9.12. Two curves are given, corresponding to the spin-only formula and the (spin + orbital moment) formula. Experimental values fall between the two theoretical curves and show that the orbital moment is only partially quenched.

The magnetic moments of the rare earth garnets show an interesting and unusual temperature dependence. The spontaneous moment at absolute zero, for DyIG, Fig. 9.13, decreases with increase in temperature, falls to zero at the *compensation temperature*, rises again, but in the opposite direction, and falls to zero a second time at the Curie temperature. This occurs because the spins on the rare earth sublattice randomise more rapidly than those on the Fe^{3+} sublattice.

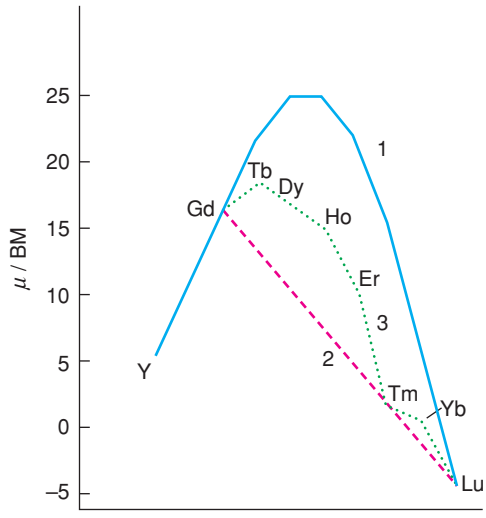


Figure 9.12 ***Variation of magnetic moment at 0 K of garnets. Curve 1, calculated, spin + orbital formula; curve 2, calculated, spin only formula; curve 3, experimental. Data from Standley,*** **Oxide Magnetic Materials,** ***© 1972 Clarendon Press.***

Many ionic substitutions in the garnet structure are possible and the magnetic properties vary systematically. For instance, the large trivalent ion in the 24c site may be replaced partially by Ca^{2+}; to charge compensate, some Fe^{3+} on tetrahedral sites may be replaced by V^{5+} according to

$$\left[Y^{3+}_{3-2x}Ca^{2+}_{2x}\right]Fe^{3+}_{2}\left[Fe^{3+}_{3-x}V^{5+}_{x}\right]O_{12}$$

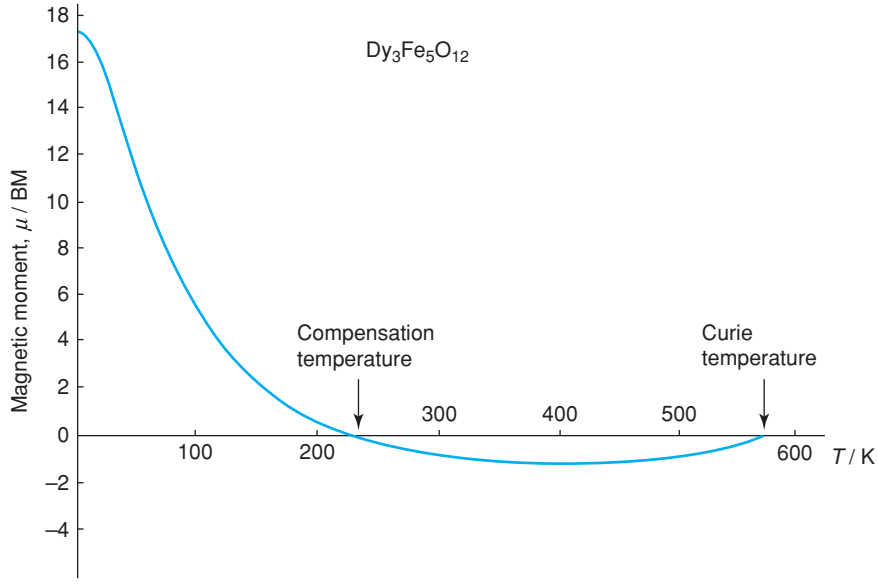


Figure 9.13 ***Spontaneous magnetisation in dysprosium iron garnet.***

9.2.6 Ilmenites and perovskites

Ilmenites, ABO_3: A = Fe, Co, Ni, Cd, Mg; B = Ti, Rh, Mn, have a structure closely related to that of corundum, α-Al_2O_3 (see Section 1.17.11). The symmetry is rhombohedral but the structure can be drawn and visualised more easily using a larger, hexagonal unit cell, Fig. 1.46. It has an *hcp* array of oxide ions. Cations occupy two-thirds of the octahedral sites but are segregated so that, along *c*, there are alternate layers of A and B cations. Another way of looking at the ilmenite structure is as a derivative of NiAs with one-third of the octahedral sites vacant, but with ordering of the two types of cations.

The perovskite structure of $SrTiO_3$ is described in Section 1.17.7. Some oxides containing Fe^{3+} and $Mn^{3+,4+}$ have a perovskite structure and interesting ferromagnetic properties. These are mixtures of $La^{3+}Mn^{3+}O_3$ and $A^{2+}Mn^{4+}O_3$, which form double substitutional solid solutions of formula

$$\left[La^{3+}_{1-x}A^{2+}_{x}\right]\left[Mn^{3+}_{1-x}Mn^{4+}_{x}\right]O_3$$

The larger La^{3+}, A^{2+} (A = Ca^{2+}, Sr^{2+}, Ba^{2+}, Cd^{2+}, Pb^{2+}) ions occupy 12-coordinate sites and $Mn^{3+,4+}$ occupies octahedral sites. The crystal chemistry, phase diagrams and magnetic properties of these systems are, however, complex.

9.2.7 Magnetoplumbites

Magnetoplumbite is the mineral $PbFe_{12}O_{19}$; its barium analogue, $BaFe_{12}O_{19}$, BaM, is an important component of permanent magnets. The structures are closely related to that of β-alumina, '$NaAl_{11}O_{17}$' (see Section 8.5.4.2). Magnetoplumbite has a similar five-layer repeat unit of *cp* layers. Four layers contain *cp* oxide ions, as in β-alumina. The fifth layer contains three-quarters of its quota of oxide ions with a large divalent ion – Pb^{2+}, Ba^{2+} – in the other oxide ion site. Hence the repeat unit contains $(4 \times 4) + (1 \times 3) =$ 19 oxide ions with one Ba^{2+}, Pb^{2+} ion completing the fifth layer, Fig. 9.14.

The magnetic structure of BaM is complex because Fe^{3+} ions occupy five sets of crystallographic sites. However, the net effect is that in the formula unit $BaFe_{12}O_{19}$, eight Fe^{3+} ions have spins oriented in one direction and the remaining four are antiparallel, giving a resultant of four Fe^{3+} ions with $\mu = 20$ BM.

9.3 Applications: Structure–Property Relations

Many parameters affect magnetic properties. By careful control of composition and fabrication procedures, it is possible to carry out 'crystal engineering' and prepare materials with a desired set of properties. There are three main application areas of magnetic materials: soft materials with high μ and low H_c for use in electromagnets in which the objective is to optimise B for a given H; hard materials with low μ, high H_c and high M_r for permanent magnets; and materials with various properties for magnetic information storage. Examples are as follows.

9.3.1 Transformer cores

A major application of ferro- and ferrimagnetic materials is in transformer and motor cores, Fig. 9.15. Materials should be magnetically soft with large power-handling capacity and low losses. Magnetically soft materials have high permeability, are magnetised easily at low applied fields, have a low coercive field and tend to have low hysteresis losses or *BH* products. These properties are favoured in materials that have a

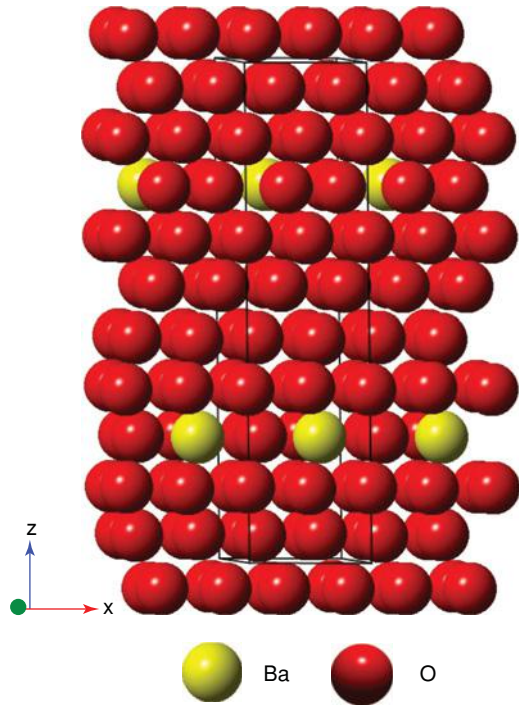


Figure 9.14 ***Packing arrangement of oxygens and Ba in the magnetoplumbite structure.***

small magnetostriction coefficient, λ_s, and a low magnetocrystalline anisotropy coefficient, K_1. Soft magnetic materials are, mechanically, those in which the domain walls migrate easily.

In addition to hysteresis losses, eddy current losses are a serious problem at high frequencies, especially in materials of low resistivity, because eddy currents are proportional to (frequency)^2. Eddy currents in metals such as Fe can be reduced by alloying Fe with, for example, Ni or Si: alloys generally have higher resistivity than the component metals. One great advantage of ferrimagnetic oxides such as ferrites and YIG is their high resistivity and, consequently, negligibly small eddy currents. Thus, in YIG containing trivalent cations only, there is no easy mechanism for electronic conduction to occur and $\sigma_{25\,^\circ C} \approx 10^{-12}$ S cm^{-1}. It is necessary to

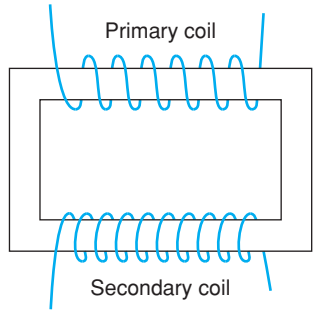


Figure 9.15 ***Primary and secondary coils wound on a transformer core.***

ensure that all the Fe is in the +3 state, otherwise $Fe^{2+} \rightarrow Fe^{3+} + e^-$ redox transfer may give rise to high conductivity. For instance, magnetite, Fe_3O_4 (i.e. $Fe^{2+}\,Fe_2^{3+}O_4$), has a conductivity of 10^2–10^3 S cm^{-1} at room temperature; this is ~15 orders of magnitude higher than that of YIG and is associated with the mixed valency of iron.

9.3.2 Permanent magnets

Permanent magnets require high saturation magnetisation, high energy product *BH*, high coercive field, high remanent magnetisation, high Curie temperature and high magnetocrystalline anisotropy. Suitable materials are metals such as Fe, Co and Ni, new alloys such as $SmCo_5$, and 'hard' oxides such as BaM.

The hardness of magnets can be increased by pinning or reducing the ease of motion of domain walls. This may be achieved in steels by adding dopants such as Cr or W that cause either precipitation of a carbide phase or a martensitic transformation on cooling. In Alnico magnets, ferromagnetic Co, Ni-based material is present as a large number of small crystalline regions embedded in an Al-based matrix. These small regions are magnetised in the same direction and it is difficult to demagnetise them or change their magnetic orientation.

Oxide magnets such as BaM are relatively light and cheap. Their intrinsic properties are generally inferior to those of Alnico magnets, but can be improved if prepared with a magnetically aligned texture. To achieve this, the powdered starting materials are subjected to a magnetic field while they are prepared into a shape and subsequently sintered. The applied field causes magnetic alignment of the grains and this increases the remanent magnetisation of the material.

9.3.3 Magnetic information storage

For magnetically based information storage, memory components should be soft with low eddy current losses and a certain type of hysteresis loop, either square or rectangular, Fig. 9.6. With this characteristic, a reverse field may be applied to a magnetised sample and it should undergo no change until the coercive field, H_c, is exceeded, at which point a sudden switch in magnetisation occurs. The two orientations of magnetisation, + and −, can be used to represent 0 and 1 in the binary digital system.

For cassettes, tapes and floppy disks, sound data are recorded as variations in the amount of magnetisation on the surface of a recording medium, which is usually either a composite consisting of a polymer film embedded with tiny particles of magnetic CrO_2 or γ-Fe_2O_3, or a thin film of magnetic material deposited on a substrate. Each particle or domain can be magnetised in either of two directions, representing 0 or 1. The read–write recording head is a tiny electromagnet of a soft material, such as MnZn ferrite, with a small gap, ~0.8 μm wide, across which a magnetic field can be induced with the poles as N and S, Fig. 9.16. The recording medium is passed across the read–write head. In write mode, the fringing field that escapes from the edge of the gap magnetises the recording medium in the immediate vicinity; the strength of magnetisation is dependent on the field, which is controlled by the current in the coil. In 'read' mode, a change in fringing field at the surface of the recording medium induces a voltage in the coil which is fed to a loudspeaker.

The materials used for read–write heads and the recording media have different requirements for their magnetic properties. The heads are made of high-permeability, soft materials giving high magnetisation and low coercivity. The recording media require hard magnetic properties such as high coercivity to avoid accidental erasure linked to high saturation magnetisation and remanence to allow the fringing field from the surface of the recording media to be detected by the read head.

The storage density of magnetic recording media is an important property given the high storage densities that are achievable with CDs and optical storage systems. Storage density is given by the number of bits

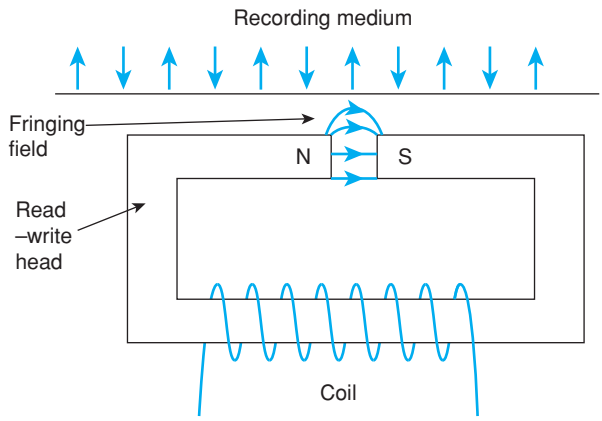


Figure 9.16 ***Read–write magnetic recording head.***

(magnetic domains or particles) per unit area. For domain sizes of 1 μm, storage densities of 10^8 bits cm^{-2} are achievable, especially in perpendicular recording microstructures where the magnetic particles are aligned to be vertical to the tape plane since this leads to higher storage densities and sharper fringing fields.

For CDs, an optical read–write system is used. In 'record' mode, a focused laser beam is used to heat the recording medium either to above its Curie temperature at which its ferromagnetism and associated stored memory are lost, or to a sufficiently high temperature that its coercivity is reduced and the sample can be easily remagnetised. The sample then passes a magnetic recording head where it is either magnetised or remagnetised; on cooling, the coercivity increases again and the stored magnetisation becomes semi-permanent. In 'read' mode, a laser beam of reduced intensity is able to distinguish the different states of magnetisation by making use of the magneto-optic Kerr effect but without melting the ferromagnetic domain structure. In the Kerr effect, the direction of polarisation of a linearly polarised beam of laser light rotates on reflection from the surface of a magnetic medium; the direction of rotation depends on the direction of magnetisation whereas the size of the angle of rotation depends on the material used in the recording medium.

Improvements in information storage densities over the period 1960–1990 are shown in Fig. 9.17. As new technologies have been developed, increasing miniaturisation has been possible and currently the highest densities are achieved with magneto-optic recording systems.

An interesting early development in information storage used thin films of garnet, a few microns thick, deposited epitaxially at high temperature onto a non-magnetic substrate. By careful choice of composition and, in particular, lattice parameters, a slight difference in thermal contraction between film and substrate occurs on cooling. The resulting stresses induce a preferred direction of magnetisation in the garnet film perpendicular to the plane of the film, and a domain structure with spins pointing up or down that appear as bubbles when viewed in a polarising microscope. These *magnetic bubble* materials are used as memory components in binary data storage systems, although they have now been superseded by other technologies.

9.4 Recent Developments

9.4.1 Magnetoresistance: giant and colossal

Magnetoresistance, MR, is the change in electrical resistance of a material in response to an applied magnetic field. Under normal circumstances, it is a small effect involving a change of $<1\%$ in resistance even in high

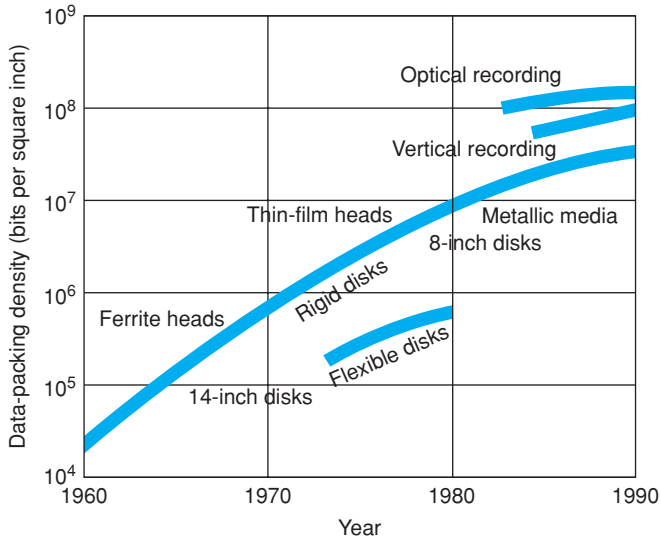


Figure 9.17 ***Improvements in magnetic recording densities for floppy disks, rigid disks, magneto-optical and perpendicular recording techniques. Adapted with permission from David Jiles,*** **Introduction to the Electronic Properties of Materials*****, Chapman and Hall, © 1994 Taylor & Francis Ltd.***

fields of, e.g., 1 T and arises from a small modification to the movement of conduction electrons in an applied field. In *anisotropic magnetoresistance*, AMR, the resistance of a magnetic material changes in an applied magnetic field but depends on the relative orientations of the field and the sample magnetisation. Again, this is a small effect but finds applications in magnetic sensors since the resistance is sensitive to small changes in magnetic field.

Interest in magnetoresistance underwent a step change with the discovery of *giant magnetoresistance*, GMR, in 1988 by Fert and Grünberg, for which they were awarded the Nobel Prize in Physics in 2007. To observe GMR, it is necessary to have a thin multilayer structure with alternating layers of a ferromagnetic metal such as Fe and a non-magnetic metal such as Cr. Depending on the layer thicknesses, exchange coupling between the Fe layers, mediated by the itinerant electrons in the Cr layers, causes adjacent Fe layers to be coupled either antiferromagnetically or ferromagnetically, with antiparallel or parallel spins, respectively, Fig. 9.18. In GMR devices, the Cr layer thickness is controlled to give antiferromagnetic coupling of the Fe layers. In an applied magnetic field, the spins that are antiparallel to H switch orientation to become parallel to H, Fig. 9.18(b).

The key to the GMR effect is that the resistance, R_p, for electrons flowing through a magnetic material whose magnetisation is parallel to the applied field, i.e., with a spin-up arrangement, is significantly lower than that, R_a, for the antiparallel, spin-down arrangement since the spin-down electrons effectively act as strong scattering centres and reduce the mobility of the conduction electrons. For electron flow perpendicular to the multilayers, the overall resistance is controlled by that of the antiparallel layers, (a), giving rise to the field-dependent resistance shown in (c).

Several other methods that generate GMR effects have been developed, in addition to the original anti-ferromagnetic coupling of Fert and Grünberg. One method which avoids the need for non-magnetic spacer layers is to have alternate ferromagnetic layers of different coercivity; in response to an applied field, the soft layer switches first to generate an antiparallel arrangement between adjacent ferromagnetic layers, with

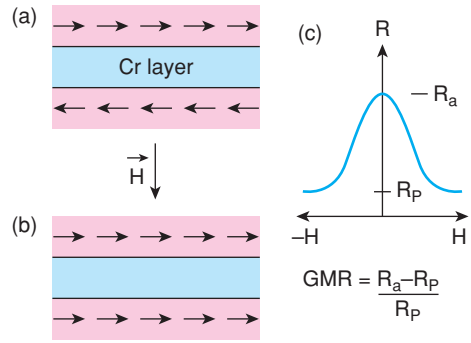


Figure 9.18 ***Giant magnetoresistance in FeCr multilayer structures: (a) antiferromagnetic coupling of adjacent Fe layers; (b) switch to ferromagnetic coupling in an applied field; (c) variation of resistance perpendicular to the layers with*** **H.**

a consequent increase in resistance. At still higher field, the hard layer also switches; the magnetisation in both layers is then parallel and the net resistance decreases.

In order to observe the GMR effect, high-quality multilayer structures are needed in which the magnetic layers are typically only a few nanometres thick. These require fabrication using advanced thin-film deposition techniques such as molecular beam epitaxy, MBE. In order to explain the GMR effect, it is necessary to imagine the existence of spin-up and spin-down sub-bands, as described earlier in accounting for ferromagnetism. This spin-polarised electronic band structure effectively gives rise to two independent channels for electronic conduction. Applications for GMR devices rapidly followed the discovery of the GMR effect. From the variation of resistance with field, as shown in Fig. 9.18(c), applications have followed in areas such as field sensors, magnetic memory chips and hard disk read-heads.

A further example of the diversity of MR phenomena is associated with *colossal magnetoresistance*, CMR. The CMR effect is observed in complex perovskite materials such as $La_{1-x}Ca_xMnO_3$. It is associated with a metal–insulator transition which occurs on application of very high magnetic fields of several tesla at temperatures close to the Curie temperature, T_c. An example is shown in Fig. 9.19 of the change in resistivity of $(Pr_{0.7}Ca_{0.3})MnO_3$ at 4.2 K on application of a magnetic field. In the range 3–6 T, the material changes from an insulating antiferromagnet to a metallic ferromagnet with a change in resistivity of 10 orders of magnitude. Because of the extreme magnetic fields required to create the CMR effect, at present it is still a laboratory curiosity in search of applications.

A different kind of multilayer MR effect is *tunnelling magnetoresistance*, TMR, also known as *junction magnetoresistance,* JMR. In TMR, ferromagnetic metallic layers are separated by thin insulating layers of, e.g., Al_2O_3. The ferromagnetic layers are coupled magnetically by spin-polarised tunnelling through the intervening insulating spacing layer. This contrasts with GMR, in which coupling occurs via the intervening metallic layer, but is reminiscent of the Josephson effect involving coupling between superconductors by tunnelling through an insulating gap.

9.4.2 Multiferroics

Multiferroics are materials which exhibit more than one type of ferroic order. Of most interest are materials that, at the same time, are ferromagnetic and ferroelectric and in particular if there is coupling between the two properties such that one can be induced or modified by the other. Two types of multiferroic are identified, Type I in which the ferroelectricity is largely independent of ferromagnetism and Type II in which

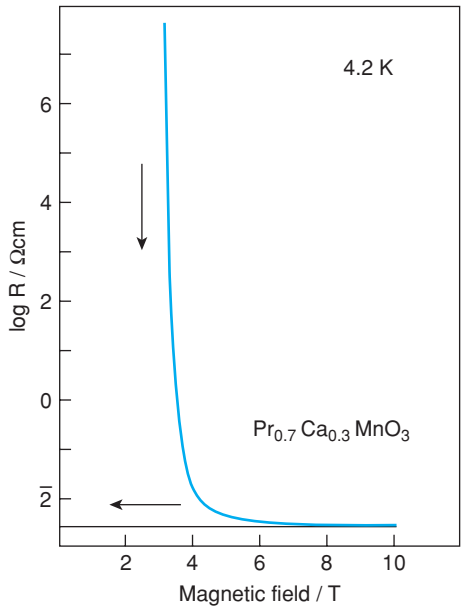


Figure 9.19 ***Dramatic decrease in resistivity of $Pr_{0.7}Ca_{0.3}MnO_3$ in response to an applied magnetic field. Adapted with permission from H. Yoshizawa et al., Neutron-diffraction study of the magnetic-field-induced metal-insulator transition in $Pr_{0.7}Ca_{0.3}MnO_3$,* Phys Rev B, *52 (1995) 13145–13148, © 1995, American Physical Society.***

the ferroelectricity can be stimulated by ferromagnetism; Type II multiferroics give intrinsic magnetoelectric couplings and, therefore, are of more interest for practical applications in which magnetic properties can be switched in an applied electric field or vice versa.

As we have seen, ferroelectric materials exhibit a spontaneous electric polarisation that can be switched by an applied electric field (see Chapter 8), whereas ferromagnetic materials exhibit spontaneous magnetisation that can be switched by an applied magnetic field. A third category is *ferroelastic* materials, which exhibit spontaneous deformation that can be switched by an applied stress, an example of which is $CaAl_2Si_2O_8$. Interest in the field of multiferroics increased greatly after 2003 when two significant examples of multiferroicity were discovered. One was giant ferroelectricity in ferromagnetic films of $BiFeO_3$; the other was magnetic field-controllable ferroelectricity in $TbMnO_3$. Multiferroics are now a hot topic in solid state physics and chemistry.

The possibility of coupling together ferromagnetism and ferroelectricity represents a significant scientific challenge since the requirements for these two properties are fundamentally different. Ferroelectric materials should be electrical insulators and have non-centrosymmetric crystal structures involving polarity in cation–anion bonds. Typically, ferroelectrics are transition metal-containing materials in which the transition metal ions have an empty *d* shell, such as Ti^{4+}, Nb^{5+} and Ta^{5+}. Most ferromagnetic materials, by contrast, are metallic conductors in which the cation–anion bonds do not exhibit a dipole moment and, critically, the magnetism requires unpaired electrons in *d* and *f* shells; ferromagnetism is therefore favoured in late transition elements such as Mn, Fe, Co and Ni and their compounds.

In $BiFeO_3$, one of the most studied multiferroic materials, which has a T_N of 643 K and a T_c of 1100 K, the origin of the ferroelectricity is associated with the Bi^{3+} cation and its $6s^2$ lone pair of electrons which generates

an asymmetric coordination environment for Bi leading to a polar Bi–O bond. The Fe^{3+} ions are responsible for antiferromagnetic order in $BiFeO_3$ through Fe–O–Fe exchange interactions. The antiferromagnetism is, in fact, complex since the spin interactions also incorporate a helimagnetic spiral spin structure.

$BiFeO_3$ has the perovskite structure but in its ferroelectric polymorph below T_c, the structure undergoes a rhombohedral distortion in which there is spontaneous polarisation parallel to the [111] direction of the high-temperature cubic unit cell. Several other perovskite materials are also multiferroic, including $BiMnO_3$ and $PbVO_3$.

10

Optical Properties: Luminescence and Lasers

10.1 Visible Light and the Electromagnetic Spectrum

First, we should remind ourselves 'what is light?'. In spite of its ubiquitous familiarity, as we delve into the different types of light, its properties and its diversity of applications, we realise that it is not the easiest phenomenon to describe. However, once we appreciate that it is just another form of electromagnetic radiation similar to radiowaves, microwaves, ultraviolet radiation and X-rays, then a common approach can be used to understand, describe and analyse these different forms of radiation: they differ only in their range of wavelengths and, therefore, their associated energies. Visible light extends from the lower energy, red end of the visible spectrum with wavelength ~700 nm and energy ~1.7 eV to the higher energy, violet end, ~400 nm and 3.3 eV, Fig. 10.1(a).

Light can be regarded as a travelling wave that propagates through space as a combination of mutually perpendicular oscillating electric and magnetic fields, Fig. 10.1(b). The direction of propagation of the wave is perpendicular to both field directions; it is, therefore, a 3D vector wave in terms of magnitude and direction but for many purposes can be treated as a propagating 1D scalar wave, Fig. 10.1(c). The light that we experience every day is made up of waves that propagate in all directions with a complete spectrum of wavelengths that averages out to white light and with no coherence between different light waves.

When the electric field vector of travelling light waves oscillates or vibrates in a single plane, as shown in Fig. 10.1(c), the light beam is *linearly polarised* if all waves vibrate in the same direction. If there is no relation between the vibration directions of component waves in a beam of light, the beam is *unpolarised.* Other types of polarisation can occur; in *circular polarisation*, the electric field vector rotates continually perpendicular to the direction of propagation, which can be *left* or *right* depending on whether rotation is anticlockwise or clockwise. A more general case of circular polarisation is *elliptical polarisation,* in which the amplitude of the electric field vector changes with rotation.

10.2 Sources of Light, Thermal Sources, Black Body Radiation and Electronic Transitions

Hot objects or materials emit radiation which often has a characteristic colour. Thus, a sodium street lamp is yellow, the embers of a fire may be orange–red and the Sun is white. Depending on the source, these colours may have a broad spectrum of wavelengths or may consist of a series of discrete wavelengths. The spectral

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

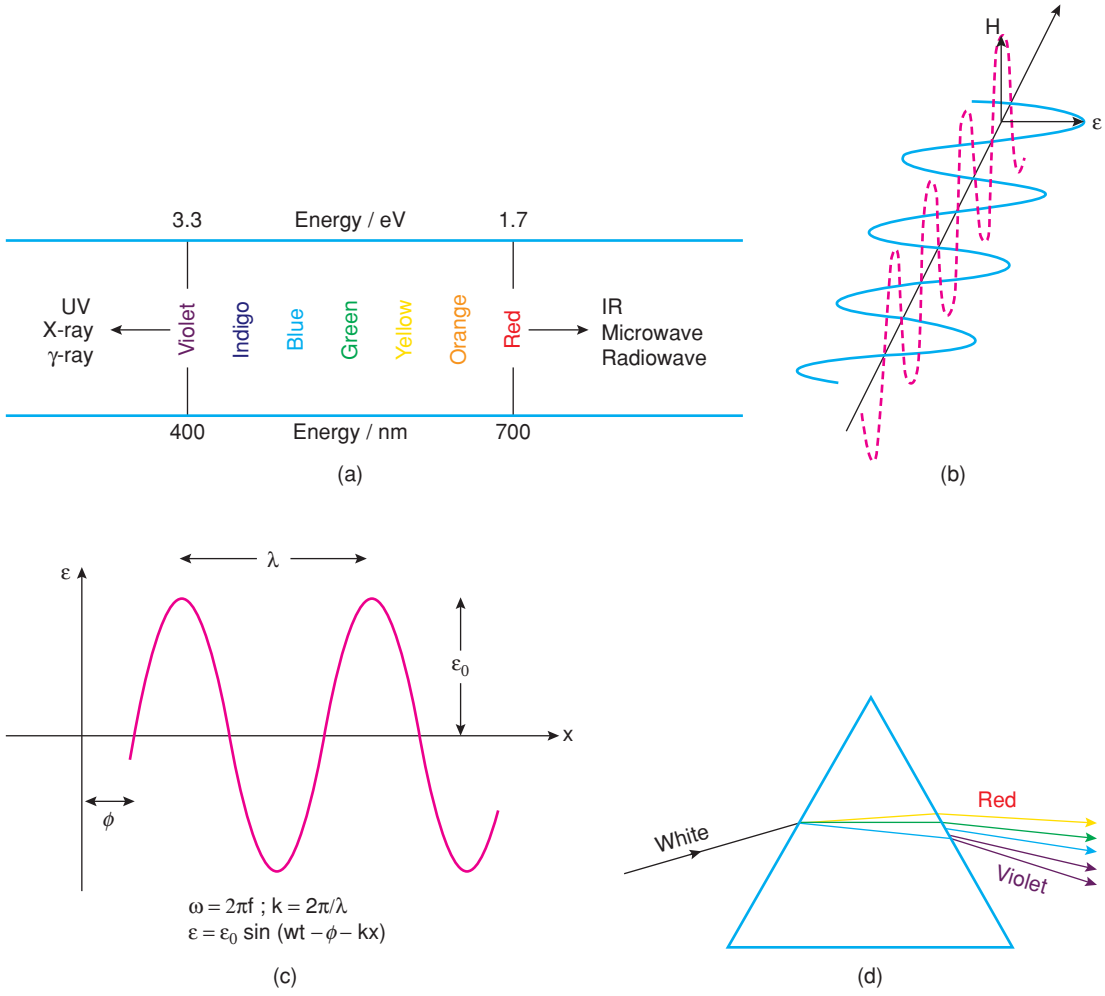


Figure 10.1 ***(a) The electromagnetic spectrum in the region of visible light; (b) mutually perpendicular oscillating electric and magnetic fields that form a propagating light beam; (c) a 1D electromagnetic wave travelling in the x-direction; (d) separation of white light into its component colours on passing through a prism.***

distribution may be determined by passing the light through a prism. The light which emerges is separated according to its wavelength and is at different angles to the incident beam, Fig. 10.1(d).

The origin of such spectra is associated with the existence of allowed energy levels for the electronic structure of atoms, molecules and solids, as discussed earlier for the processes involved in the generation of X-rays. In a hot body, there is a continual process of electrons moving between the various levels and by adjusting the experimental conditions it is possible to determine either *absorption spectra* or *emission spectra*. For a given temperature, an equilibrium develops between the absorption and emission processes which occur spontaneously in the absence of any external stimuli.

For the simple situation with two energy levels, E_0 and E_1, separated by an energy gap, ΔE, and an initial population, N_0, of electrons in level E_0, then depending on temperature, T, a number of electrons N_1 will gain

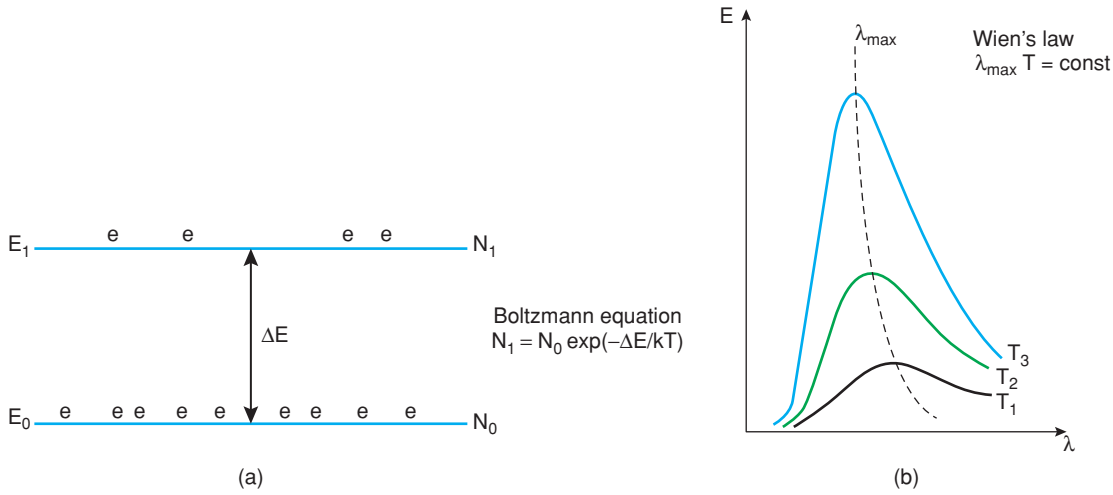


Figure 10.2 ***(a) Population of energy levels by thermal activation. (b) Energy distribution in black body radiation. Wien's constant (determined experimentally) is 2.9 × 10^{-3} m K.***

enough energy to be promoted to level E_1, Fig. 10.2. This shows the importance of temperature in controlling population densities for different energy levels. From the kinetic theory of gases, monatomic species have approximate thermal energy given by $3kT/2$ and an approximate 'rule of thumb' is that this energy equates to 0.025 eV at room temperature. To a first approximation, we may regard electrons as having a similar amount of thermal energy as monatomic gaseous species.

Provided that there are vacancies in energy levels at E_0, electrons in level E_1 spontaneously drop back to the lower levels, emitting energy ΔE in the process. This is known as *spontaneous emission*. Under normal circumstances, an electron will remain in an upper energy level for only a few nanoseconds before it spontaneously relaxes to a lower level. On average, relatively few electrons will be in higher energy levels although this number increases with temperature, as given by the Boltzmann equation in Fig. 10.2(a).

The emission spectra of gases and solids are usually very different in character. In gaseous atoms, such as in an Hg discharge lamp, the number of available electron energy levels is relatively limited, giving rise to discrete line spectra. In solids, by contrast, using the band model description of electron energy levels, their occupancy and subsequent bonding, the outermost valence electrons either occupy partially the conduction band in metals or occupy fully the valence bands which are separated by band gaps from empty conduction bands in non-metallic solids. Either way, electrons are increasingly promoted to higher levels within conduction bands at higher temperatures, giving rise to a continuum of partially occupied levels. This leads to broad band emission spectra associated with the emission component of the absorption–emission processes.

An emission spectrum is shown schematically in Fig. 10.2(b) and is known as *black body radiation.* The spectrum changes with temperature, as given by Wien's law; it displaces to shorter wavelengths with increase in temperature. Materials that form the light sources in arc discharge and incandescent lamps have broad band emission spectra such as this. The spectra consist of a wide range of wavelengths that are emitted at random times and in random directions, giving a light source that is *incoherent*. As discussed later, this is in complete contrast to the operation of lasers in which light emission is essentially monochromatic and *coherent*.

Thus far, we have seen how absorption and emission processes depend directly on electronic transitions between energy levels; if the energies involved fall within the visible range of energies, the transitions are known as *optical transitions*. In addition to absorption and emission, a range of other scattering processes occur when light interacts with matter, as follows.

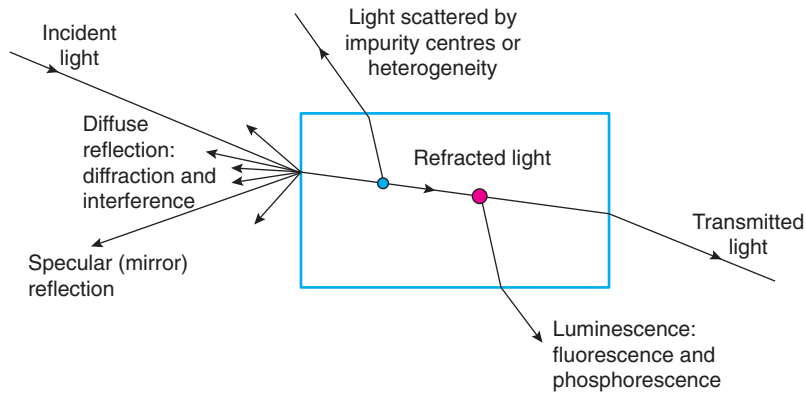


Figure 10.3 ***The interaction of light with a solid. The light can be reflected, transmitted, scattered or absorbed. The appearance of the solid depends on which of these interactions occur. Adapted with permission from Richard Tilley,* Understanding Solids*, p. 442, © 2004, John Wiley & Sons Ltd. Reproduced under the terms of the STM Agreement.***

10.3 Scattering Processes: Reflection, Diffraction and Interference

Reflection is the familiar process in which radiation is reflected off a surface and the angle of incidence equals the angle of reflection in the case of mirror or specular reflection. The reflected beam maintains coherency with the incident beam and for an individual wave, there is no loss of energy in the reflection process. Reflection is a major source of colour of many materials alongside absorption and emission. Thus, *emission colours* constitute those wavelengths that are present in the emitted light beams and are related to the inter-shell electronic transitions responsible for emission. *Absorption colours* are the residual colours after those light wavelengths which are used in absorption are removed from the spectrum of incident light. *Reflection colours* also consist of those colours that are not removed from the incident spectrum by any absorption processes.

Some of the wide variety of processes that can result from the interaction of light with a solid are shown schematically in Fig. 10.3. Reflection processes are dominant in cases where incident light interacts with smooth, flat surfaces whose physical dimensions are much greater than the wavelength of the incident light. A different kind of process, *diffraction*, occurs when the physical dimensions of the material or features on its surface are comparable to those of the incident light. Sample features then act as secondary sources of radiation. This is treated in Chapter 5 with first, the diffraction of light by an optical grating, and second, the diffraction of X-rays (and also neutrons and electrons) by atoms in crystalline solids. Diffraction from secondary sources, or more generally diffuse scattering, effectively scatters light in all directions, leading to a much reduced intensity at any specific angle of reflection for a particular incident beam. This forms the basis of the method used to read information stored optically on compact disks: etch pits of dimensions comparable to the wavelength of light scatter light weakly in all directions compared with the much stronger directional reflection from adjacent mirror surfaces. As also discussed in Chapter 5, whenever diffraction occurs, *interference* between diffracted beams leads to the whole range of possibilities from constructive interference to complete cancellation.

10.4 Luminescence and Phosphors

Luminescence is the emission of light by a material after it has absorbed energy. Various types of excitation source may be used, indicated as a prefix: *photoluminescence* uses photons or light, often UV, for excitation;

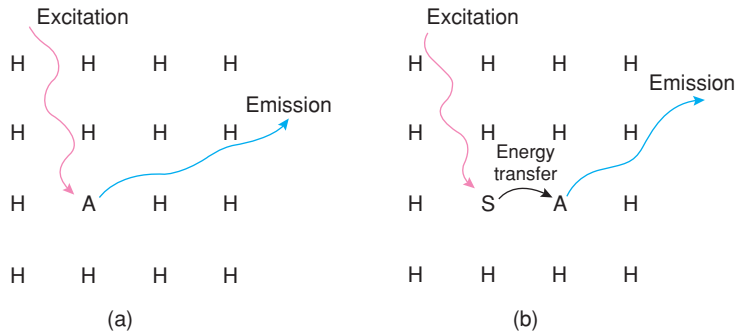


Figure 10.4 ***Schematic representation of luminescence involving (a) an activator, A, in a host lattice, H, and (b) both a sensitiser, S, and an activator, A. Adapted with permission from DeLuca,* J. Chem. Educ.*, 57, 541, © 1980, American Chemical Society. Reproduced under the terms of the STM Agreement.***

electroluminescence uses an electrical energy input; *cathodoluminescence* uses cathode rays or electrons. Two types of photoluminescence may be distinguished. For a short time lapse, $\leq 10^{-8}$ s, between excitation and emission, the process is known as *fluorescence*. Fluorescence effectively ceases as soon as the excitation source is removed. For much longer decay times, the process is known as *phosphorescence*. It may continue long after the excitation source has been removed.

Photoluminescent materials generally require a *host* crystal structure, ZnS, $CaWO_4$, Zn_2SiO_4, etc., which is doped with a small amount of an *activator*, a cation such as Mn^{2+}, Sn^{2+}, Pb^{2+} or Eu^{2+}. Sometimes, a second dopant is added to act as a *sensitiser*. The mode of operation of inorganic luminescent materials, known generally as *phosphors*, is shown schematically in Fig. 10.4. The energy of the emitted light is usually less than that of the exciting radiation and is, therefore, of longer wavelength. This increase in wavelength is known as the *Stokes shift*. In fluorescent lamps, the exciting radiation is UV light from a mercury discharge. Phosphor materials absorb this UV radiation and emit 'white' light. The construction of a fluorescent lamp, Fig. 10.5, consists of a glass tube lined on the inside with a coating of phosphor material and filled with a mixture of Hg vapour and Ar. On passage of an electric current through the lamp, the atoms of Hg are bombarded by electrons, excited into upper electronic energy states and then return to the ground state, accompanied by the emission of UV light of two characteristic wavelengths, 2540 and 1850 Å. This light irradiates the phosphor which subsequently emits white light.

Examples of emission spectra of ZnS phosphors are shown in Fig. 10.6. Each activator gives a characteristic spectrum and colour to ZnS. Some activators and their relevant electronic states are listed in Table 10.1.

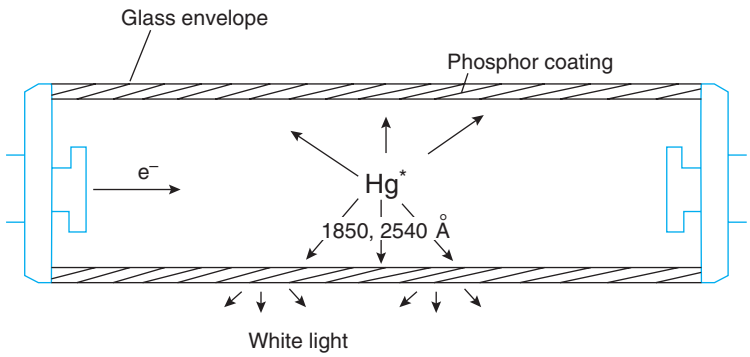


Figure 10.5 ***Schematic design of a fluorescent lamp.***

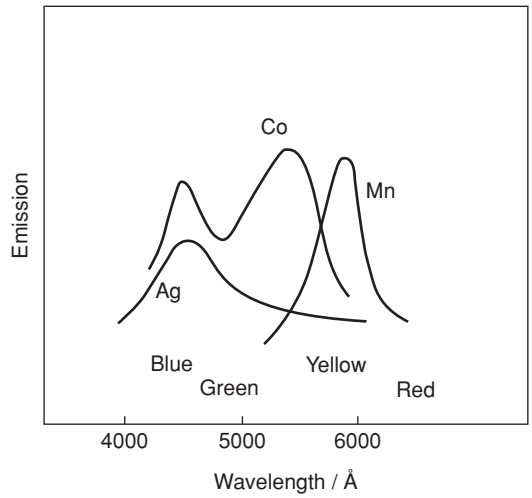


Figure 10.6 ***Luminescence spectra of activated ZnS phosphors after UV irradiation.***

The host materials fall into two main categories:

(i) Ionically bonded, insulating materials, such as $Cd_2B_2O_5$, Zn_2SiO_4 and apatite, $3Ca_3(PO_4)_2{\cdot}Ca(Cl,F)_2$. In these, a set of discrete energy levels is associated with the activator ion and these levels are modified by the local environment of the host structure. For ionic phosphors, the *configurational coordinate model* provides a useful way to represent qualitatively the luminescence processes.
(ii) Covalently bonded, semiconducting sulfides, such as ZnS. In these, the band structure of the host is modified by addition of localised energy levels associated with the activator ions.

10.5 Configurational Coordinate Model

The variation of the potential energy, PE, of the ground electronic state of the luminescent centre is shown schematically against a general coordinate, which is often the internuclear distance, in Fig. 10.7. The PE passes through a minimum at the equilibrium bond length, r_e. Within this electronic ground state, different quantised vibrational excited states of the ion, V_0, V_1, etc., are possible.

Each electronic state of the luminescent centre has a PE curve similar to that in Fig. 10.7, as shown for a ground state and excited state in Fig. 10.8; using this diagram, many features of luminescence can be explained. First, excitation raises the active centre from its ground state, level A, into a higher vibrational

Table 10.1 ***Activators and their electronic states. Data taken from Burrus, 1972.***

Ion	Ground state	Excited state
Ag^+	$4d^{10}$	$4d^9 5p$
Sb^{3+}	$4d^{10}5s^{22}$	$4d^{10}5s5p$
Eu^{2+}	$4f^7$	$4f^6 5d$

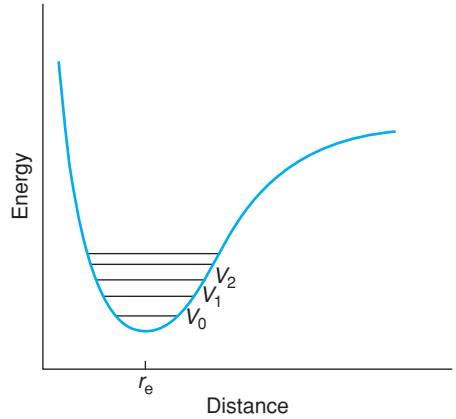


Figure 10.7 ***Ground state potential energy diagram for a luminescent centre in a host crystal.***

level, B, of the excited state. Second, energy is dissipated as the ion quickly relaxes to a lower level, C, in the excited state. This energy is lost as heat to the host lattice. Third, the active centre returns to its ground state, level D or A, and, in so doing, emits light. Since the energy of excitation A → B is greater than that of emission C → D, the emitted radiation is of longer wavelength than the exciting radiation, which accounts for the Stokes shift.

Thermal quenching, in which the luminescence efficiency decreases markedly above a certain temperature can also be explained using Fig. 10.8. The PE curves for the ground and excited states cross over at point E where an ion in the excited state can transfer back to its ground state, at the same energy, and can then return to lower vibrational levels of the ground state by means of vibrational transitions. Point E represents a spillover point, therefore. If an ion in the excited state has the right amount of vibrational energy to reach point E, it can spill over into the vibrational levels of the ground state. All the energy is then released as vibrational energy and no luminescence occurs. Point E is reached as a consequence of increasing the temperature since

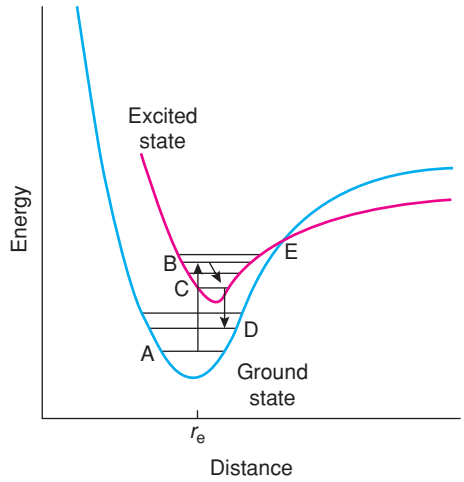


Figure 10.8 ***Ground and excited state PE diagrams for a luminescent centre.***

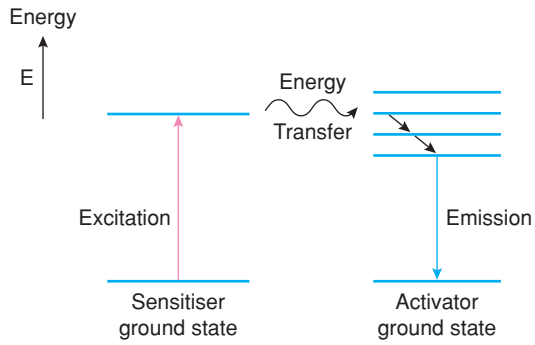


Figure 10.9 ***Non-radiative energy transfer involved in operation of a sensitised phosphor.***

ions have increasing thermal energy and are able to move to progressively higher vibrational levels. The type of transition involved in thermal quenching is *non-radiative* since the excited ion loses its excess energy as heat to the surrounding host lattice.

Another type of transition is *non-radiative energy transfer*, Fig. 10.9. It depends on (a) there being similar energy levels in the excited states of both sensitiser and activator ions and (b) sensitiser and activator ions being relatively close in the host crystal structure. In operation, the exciting radiation promotes sensitiser ions into an excited state. These then transfer energy to neighbouring activator ions, with little or no loss of energy during transfer; at the same time, the sensitiser ions return to their ground state. Finally, the activator ions return to their ground state by emission of luminescent radiation.

Non-radiative energy transfer is also involved in *poisoning* by certain impurities. In this, energy is transferred from either a sensitiser or an activator to a poison site at which the energy is lost to the host structure in the form of vibrational energy. Ions that have non-radiative transitions to the ground state and must be avoided in phosphors include Fe^{2+}, Co^{2+} and Ni^{2+}.

10.6 Some Phosphor Materials

A phosphor used extensively in fluorescent lamps is an apatite, doubly doped with Mn^{2+} and Sb^{3+}. Fluorapatite, $Ca_5(PO_4)_3F$, doped with Sb^{3+} fluoresces blue and with Mn^{2+} fluoresces orange–yellow; the two together give a broad emission spectrum that approximates to white light. A modification of the wavelength distribution in the emission spectrum is possible by partly replacing the F^- ions in fluorapatite by Cl^- ions. This modifies the energy levels of the activator ions and hence their emission wavelengths. By careful control of composition, the colour of the fluorescent light may be optimised. A selection of other lamp phosphor materials is given in Table 10.2.

Trivalent europium is an important activator ion, especially for use in red phosphors for colour television screens. In YVO_4:Eu^{3+}, the vanadate group absorbs energy in the cathode ray tube but the emitter is Eu^{3+}. The mechanism of charge transfer between the vanadate group and Eu^{3+} involves a non-radiative, superexchange process via intervening oxide ions. The effect is related to the superexchange mechanism postulated to explain antiferromagnetic ordering of Ni^{2+} ions in NiO. In order to obtain efficient energy transfer by superexchange, the metal–oxygen–metal bond should be approximately linear, so as to maximise the degree of orbital overlap. In YVO_4:Eu, the V–O–Eu angle is 170° and energy exchange takes place rapidly.

Transitions from and to several different *f* energy levels are possible, in principle, with Eu^{3+}. The ones that are actually observed, and hence the colour, depend on the host crystal structure and, in particular, on the

Table 10.2 ***Some lamp phosphor materials***

Phosphor	Activator	Colour
Zn_2SiO_4, willemite	Mn^{2+}	Green
Y_2O_3	Eu^{3+}	Red
$CaMg(SiO_3)_2$, diopside	Ti	Blue
$CaSiO_3$, wollastonite	Pb, Mn	Yellow–orange
$(Sr, Zn)_3(PO_4)_2$	Sn	Orange
$Ca_5(PO_4)_3(F, Cl)$, fluorapatite	Sb, Mn	'White'

Source: data taken from H. L. Burrus (1972), *Lamp Phosphors*, Mills and Boon, London.

symmetry of the site occupied by Eu^{3+}. If the Eu^{3+} is located at a centre of symmetry, as it is when doped into $NaLuO_2$ and Ba_2GdNbO_6, the favoured transition is $^5D_0 \rightarrow {}^7F$ (these symbols form part of the nomenclature to label spectroscopic states). The resulting colour is orange. When located on a non-centrosymmetric site, as in $NaGdO_2$:Eu^{3+}, the preferred transition is $^5D_0 \rightarrow {}^7F_2$ and the emission is red. The crystal structure of the host material therefore has a major effect on the resulting colour.

Colour television screens require three primary cathodoluminescent colours:

(i) red, mentioned above, for which YVO_4:Eu^{3+} is often used;
(ii) blue – ZnS:Ag^+;
(iii) green – ZnS:Cu^+.

For black and white screens, mixtures of blue-emitting ZnS:Ag^+ and yellow-emitting (Zn, Cd)S:Ag^+ are used.

10.7 Anti-Stokes Phosphors

The anti-Stokes phosphors exhibit the remarkable property of emitting light or photons of higher energy (shorter wavelength) than the incident exciting radiation. Using these, it is possible, for instance, to up-convert IR radiation into higher energy, visible light. There must be a catch in this somewhere, of course: the law of conservation of energy cannot be violated. What happens is that the process of excitation takes place in two or more stages, as shown schematically in Fig. 10.10.

The most studied anti-Stokes phosphors to date are host structures such as YF_3, $NaLa(WO_4)_2$ and α-$NaYF_4$, doubly doped with Yb^{3+} as a sensitiser and Er^{3+} as an activator. These convert infrared radiation into green

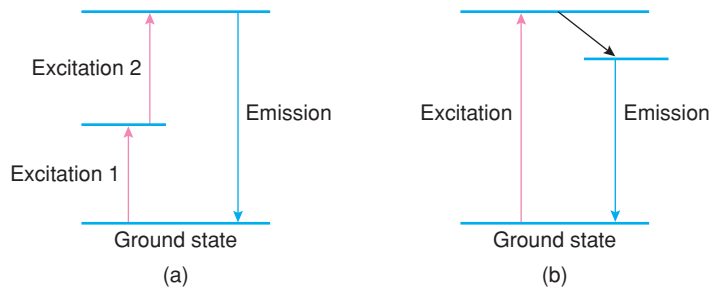


Figure 10.10 ***(a) Anti-Stokes and (b) normal luminescence phenomena.***

luminescence. During irradiation, Yb^{3+} ions transfer two IR photons to nearby Er^{3+} ions, which are raised into a doubly excited state and decay by the emission of visible light.

10.8 Stimulated Emission, Amplification of Light and Lasers

Thus far, we have regarded emission of light as the result of two main processes, absorption and spontaneous emission of energy. A third process, which led to the discovery of lasers and their widespread application, is *stimulated emission.* This has two characteristic features. First, a *population inversion* is created in which decay from the excited state takes place much more slowly than expected for spontaneous emission and allows *pumping* of an excess of electrons into the excited state, giving more electrons in the excited state than in the ground state.

The lifetime of different excited states is closely linked to their electron configuration and whether or not an electron that decays or relaxes to a lower energy level has to change its spin orientation. For transitions in which there is no change in spin, the transition is *allowed*, occurs rapidly and the excited state is short-lived. For *spin-forbidden* transitions, however, such as from parallel to anti-parallel configurations, the excited state is long-lived, allowing a population inversion to build up.

Second, electrons in the excited state are stimulated to decay and emit radiation of a certain energy by an incident photon of the same energy. If there is a large number of electrons in the excited state, a cascade effect occurs, generating an intense beam of monochromatic radiation which is in-phase and coherent. Effectively, the incident light has been amplified by this process, leading to the name *light amplification by stimulated emission of radiation* or *LASER*. The effect is not limited to visible light and can also be achieved with microwave radiation in the *MASER*.

In order to commence lasing action, a few electrons in the long-lived excited state decay by spontaneous emission. Using a system of mirrors in the laser design, the laser material is then re-irradiated with the photons that have just been released by spontaneous emission; this triggers the process of stimulated emission.

To achieve population inversion by pumping, a two-level system is insufficient, since the maximum that can be achieved is equal populations in the excited state by absorption and the ground state by stimulated emission, leading to no effective amplification of light. Instead, three- and four-level systems are needed, as in the case of ruby Cr^{3+} and Nd^{3+} laser systems, respectively. In a three-level system, electrons are pumped to the highest level and decay spontaneously to an intermediate level where their population builds up, Fig. 10.11. The condition then exists to stimulate an avalanche of coherent radiation.

Improved lasing action is obtained with four-level systems, Fig. 10.12, in which a laser transition takes place between the second and first excited states. This allows the lower laser level to be always empty by rapid spontaneous decay and the upper level to be always full by a similar mechanism, involving fast relaxation from the third excited state or pump level.

In the design of lasers and in order to maximise the light amplification, the stimulated beam is constrained to pass back and forth along a tube with mirrors at the ends, one of which is totally reflective and the other partially reflective, by a controlled amount in the range 10–90%, to ensure that some of the laser light escapes, giving useful optical power. The schematic design of a ruby laser, shown in Fig. 10.13, is effectively a longitudinal optical resonant cavity. In the cavity, only certain wavelengths are allowed. This is reminiscent of the ‘particle in a box’ model for the allowed energies of an electron in a confined situation such as a chemical bond; the same concept is used to describe the standing waves in a stretched spring pinned at both ends. Only those light beams, or modes, corresponding to multiples of half a wavelength can be supported within the laser cavity; all other modes die away. This condition is expressed as

$$L = n\lambda/2 \tag{10.1}$$

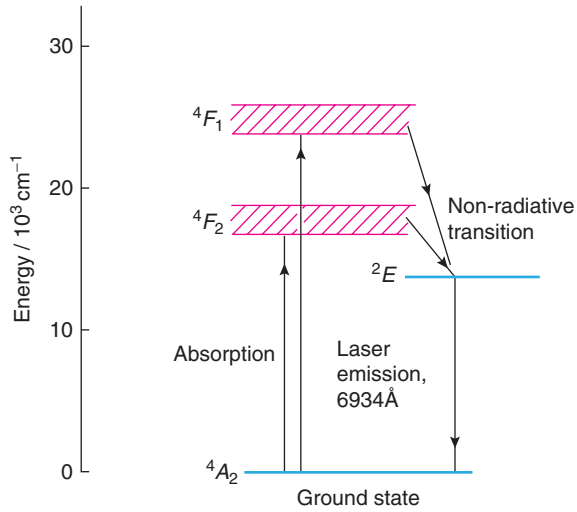


Figure 10.11 ***Energy levels of the Cr^{3+} ion in ruby crystal and laser emission.***

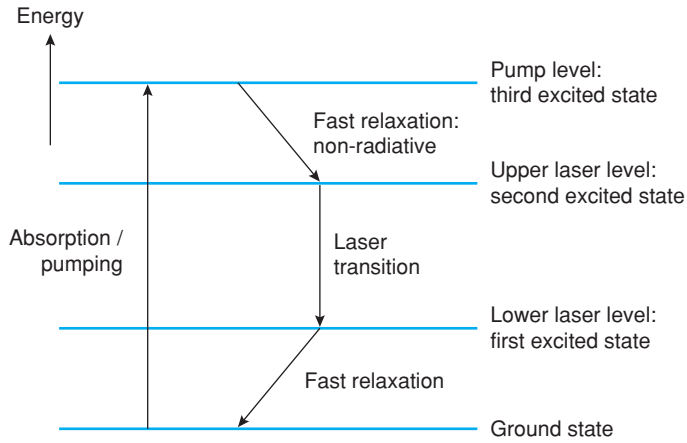


Figure 10.12 ***A four-level laser system.***

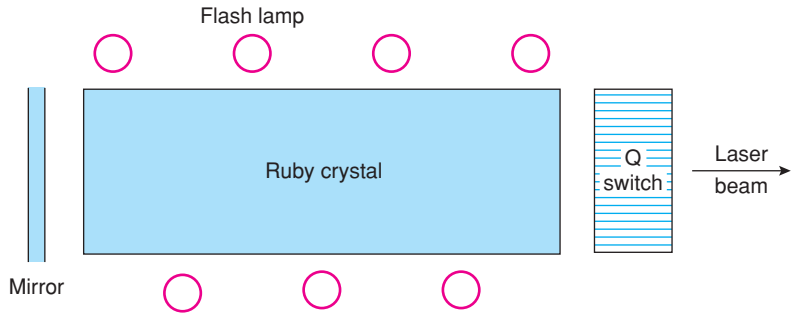


Figure 10.13 ***Design of a ruby laser.***

Table 10.3 ***Some laser systems***

Type	Medium	Principal wavelength/nm	Average output	Mode
Gas	He–Ne	633	0.1–50 mW	cw
	Ar	488, 514	5 mW–20 W	cw
	CO_2	10 600	20 W–15 kW	cw
Solid state	Ruby	694	30 mJ–100 J	Pulse
	Nd:YAG	1064	10 mJ–100 J	Pulse
	Nd:YAG (diode pumped)	1064	1–10 mW	cw
	Nd:glass	1060	100 mJ–100 J	Pulse
Semiconductor	GaAlAs	750–905	1–40 mW	cw
Excimer	ArF	193	50 W	Pulse
	XeF	351	30 W	Pulse
Dye tuneable		300–1000	2–50 W	cw or pulse

where L is the cavity length and n is a particular mode. The frequency, f, of each mode is given by

$$f = nc/2L \tag{10.2}$$

In a typical laser cavity, there may be a large number of possible modes of oscillation, which arise because the spectral line associated with the laser transition has a finite line width. In practice, to obtain a monochromatic laser beam it is necessary to fine-tune the cavity length and this is achieved by effectively tilting the mirrors so that only one mode satisfies the cavity length criterion.

The first laser system, the ruby laser, was reported in 1960 by Maiman; from this has grown a large area of modern science and technology with applications in photography, surgery, communications and precise measurements, to name but a few. Many types of laser system have been discovered; some of the more important ones are listed in Table 10.3. We are here concerned mainly with solid state and semiconductor lasers and the chemistry involved in their operation. It is not intended to give a general review of lasers; modern laser design is complex and many different laser systems are available. Lasers can deliver power in either *pulsed* or *continuous wave*, *cw*, mode. Pulsed systems give power ranging from a few watts with semiconductor lasers to around 10^{18} W delivered by solid state lasers that are used in laser fusion systems; cw systems have power ranges of a few milliwatts with He–Ne lasers to several kilowatts with CO_2 lasers.

10.8.1 The ruby laser

This was the first laser system to be discovered and more than 50 years later it is still important. The essential component is a single crystal of Al_2O_3 doped with a small amount, 0.02–0.05 wt%, of Cr^{3+}. The Cr^{3+} ions substitute for Al^{3+} ions in the distorted octahedral sites of the corundum crystal structure (see Section 1.17.11). On addition of Cr_2O_3 to Al_2O_3, the colour changes from white, in Al_2O_3, to red at low Cr^{3+} levels and to green for larger Cr^{3+} contents; these solid solutions are discussed in Section 1.3.1.

If ever you needed a convincing example of how trace amounts of impurities are entirely responsible for a given property, this is it! Typically, for every 3000 Al^{3+} ions in the corundum crystal lattice, one is replaced, at random, by a Cr^{3+} ion. The colour changes dramatically, from white to red, and stimulated emission of light becomes possible, leading to laser action. It also illustrates the contrast between solid state inorganic and molecular materials, since impurities can be incorporated into a crystal lattice, such as that of alumina, whereas this is not possible with molecular materials: if impurity atoms are introduced into a molecule then

it simply becomes a different molecule, perhaps with a significantly different structure and properties, which can be separated from the other molecules by chemical extraction procedures. The Cr^{3+} ions in ruby are an intrinsic part of the corundum crystal structure and cannot be regarded as forming a separate phase.

The energy levels of the Cr^{3+} ion in ruby are shown schematically in Fig. 10.11. On shining intense visible light from a Xe flash lamp onto a ruby crystal, Fig. 10.13, *d* electrons on Cr^{3+} ions may be promoted from the $^4A_2$ ground state into $^4F_2$ and $^4F_1$ upper states by absorbing in the blue–green region of the spectrum. These decay rapidly, by a non-radiative process, into the ^2E level. The lifetime of the ^2E excited state is fairly long, $\sim 5 \times 10^{-3}$ s, which means that a considerable population inversion has time to build up. Laser action then occurs by transition from the ^2E level to the ground state. During this transition, many ions are stimulated to decay, in-phase with each other, giving an intense, coherent pulse of red light, of wavelength 6934 Å. The laser pulse has a duration of $\sim 250\ \mu$s, with typical energies ranging from a few millijoules to several hundred joules.

The design of a ruby laser, Fig. 10.13, contains a ruby crystal rod, several centimetres long and 1–2 cm in diameter. A flash lamp is wrapped around or placed alongside the ruby rod; the two are then arranged inside a reflection cavity such that the rod is effectively irradiated from all sides. At one end of the rod, a mirror, known as an etalon, reflects the light pulse back through the rod. At the other end is a Q switch, which either allows the laser beam to pass out of the system or reflects it back through the rod for another cycle. The Q switch may simply be a rotating mirror timed to allow the laser beam out when it has reached its optimum intensity: as the light pulse passes back and forth along the rod, it builds up in intensity as more active centres are stimulated to emit radiation that is coherent with the initial pulse. Q switching allows much higher laser powers to be obtained for a given pump energy in comparison with the *free-running* mode of operation.

10.8.2 Neodymium lasers

The host material for the laser-active Nd^{3+} ion is either a glass or yttrium aluminium garnet (YAG), $Y_3Al_5O_{12}$ (see Section 1.17.13), in which a small amount of Nd^{3+} substitutes for Y^{3+}. The energy levels and transitions involved in the operation of Nd lasers, which have superseded ruby as the main solid state laser system, are shown in Fig. 10.14. During irradiation from a high-energy lamp, several absorption transitions occur, although only one is shown. The excited states all decay non-radiatively to the $^4F_{3/2}$ level from which lasing

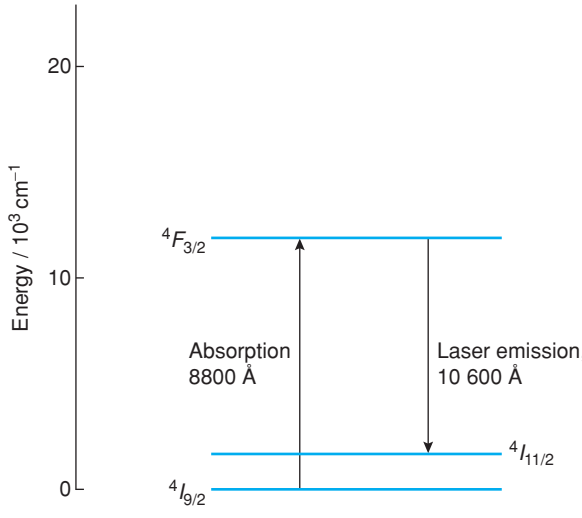


Figure 10.14 ***Energy levels of the Nd^{3+} ion in neodymium lasers.***

occurs to the $^4I_{11/2}$ level with a wavelength of 10 600 Å for Nd:glass and 10 640 Å for Nd^{3+}:YAG. The $^4F_{3/2}$ state is long lived, $\sim 10^{-4}$ s, but depends somewhat on the Nd^{3+} concentration. This again allows a large population inversion to build up and allows Nd^{3+} to be used in high-power lasers.

10.8.3 Semiconductor lasers and the light-emitting diode, LED

Semiconductor lasers are based on p–n junctions and give compact laser sources that are widely used in optical communications and high-density storage systems such as optical disks. Recalling characteristic features of a p–n junction (see Sections 8.4.2 and 8.4.4), a depletion layer develops spontaneously between p- and n-type regions which is empty of charge carriers. This occurs because electrons and holes, in the conduction and valence bands, respectively, have sufficient energy to diffuse across the junction to the other side where they can combine with carriers of the opposite polarity.

The ionised impurities, or dopants, which are fixed in the lattice to either side of the junction, ensure that a voltage gradient develops from the n-side (positive) to the p-side (negative). This gradient prevents further net crossover of mobile carriers, although electron–hole pairs are being generated continuously by thermal activation on either side of the junction. These electron–hole pairs that are created thermally give rise to minority carriers, i.e. electrons on the p-side and holes on the n-side; these minority carriers are pulled across the junction by means of the potential gradient. An equilibrium develops in which, for instance, the number of electrons crossing from n to p by diffusion is balanced by those crossing in the reverse direction by drift. This is illustrated schematically in Fig. 10.15.

On application of a forward dc bias, V_{bias}, the equilibrium at the junction, or diode, is disturbed and the barrier height is reduced by an amount V_{bias}. Majority carriers can now cross the junction more easily and the diffusion current is therefore greater than the drift current. A net flow of electrons, known as the *injection current*, flows across the junction from n to p regions (note that electrical engineers use the opposite convention and refer to the current as if it is flowing in the opposite direction; this can be confusing!). On crossing the barrier, electron–hole recombination occurs, leading to spontaneous emission of radiation whose wavelength is controlled by the energy difference between electron and hole energies. The maximum wavelength is given by electrons dropping from the bottom of the conduction band to the top of the valence band and is given by

$$\lambda_{\text{max}} = hc/Eg \tag{10.3}$$

where E_{g} is the band gap. If the electrons drop from higher states in the conduction band, then shorter wavelength radiation is emitted.

In normal operation of a p–n junction as a current rectifier, such spontaneous emission leads to loss of energy, is undesired and is minimised by attention to junction design. For laser diodes and LEDs, however, it is desired to optimise this *injection luminescence*. In addition, for laser diodes, it is necessary to create a population inversion in the conduction band so as to create the conditions for stimulated emission.

The first laser diode was based on gallium arsenide semiconductor, rather than silicon, because GaAs is a *direct* band gap semiconductor, in contrast to Si and Ge, which are *indirect* band gap semiconductors. This means that in GaAs, conduction electrons can lose energy directly by photon emission and, therefore, are more efficient at producing light than indirect band gap materials for which the electrons first have to lose some of their excess energy before decay. In this sense, indirect band gap materials need additional non-radiative transitions, such as shown in Fig. 10.11, before the spontaneous decay process can occur.

Direct band gap transitions are favoured when the energies of the two states involved in the transition have the same value of the wave vector, k, or electron momentum, since this allows the transition to occur with high efficiency. Equalisation of k vectors can occur across certain p–n junctions when the top of the valence

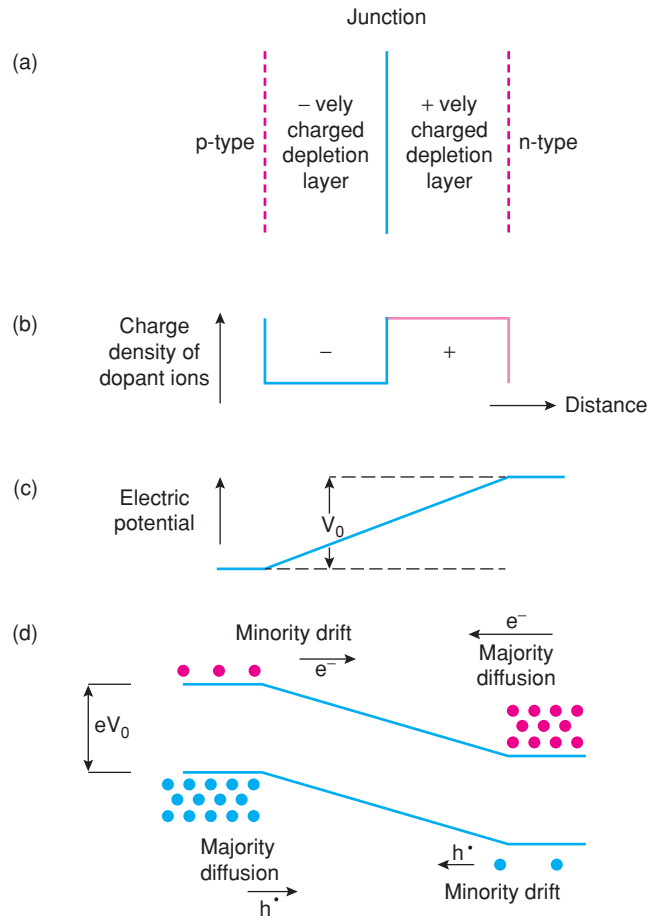


Figure 10.15 ***(a) A p–n junction showing (b) two components of a charged depletion layer, (c) resulting electric potential gradient and (d) diffusion/drift of carriers across the junction.***

band and the bottom of the conduction band have similar energies. Red and yellow LEDs and lasers are based on GaAs since the band gap transition for electrons crossing a p–n junction is direct. For blue light, GaAs is unsuitable but there is much interest in developing systems based on GaN.

GaAs lasers are designed so that a population inversion in the conduction band builds up in a very narrow depletion region known as the *active layer*, ~1 nm wide. To achieve this, a high level of doping is needed to ensure that occupied states in the conduction band are directly above empty states in the valence band. In order to achieve lasing action, the injection current needs to be above a specific threshold current since below this value, light emission is spontaneous and incoherent. The wavelength of the laser light can be controlled by doping the GaAs to give, for example, $Ga_{1-x}Al_xAs$ for which the band gap, and λ, vary with x.

As with other laser systems, for the wide diversity of applications ranging from laser light shows to laser eye surgery and holography, it is desired to have a narrow pencil-beam of laser light in which the component waves are in-phase and do not interfere destructively. Design of laser systems therefore has the critical objective of reducing beam divergence and arranging for the source to be as small as possible. One means to reduce beam divergence in semiconductor lasers is to have more complex junction structures in which, by

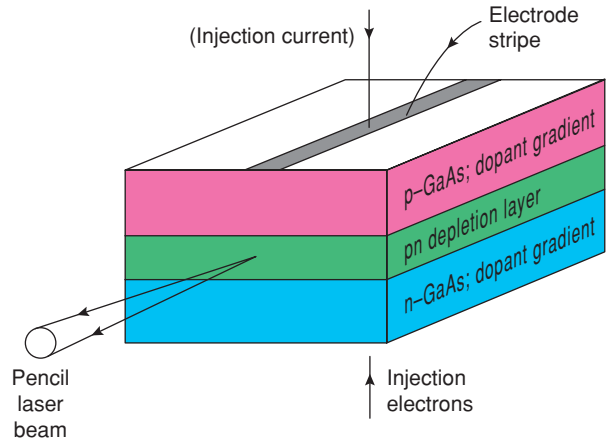


Figure 10.16 ***Construction and operation of a laser diode.***

compositional control of the p- and n-type regions, the refractive index decreases on moving away from the active junction, Fig. 10.16. Laser emission is then confined to be parallel to, and within, the active junction region of higher refractive index; the same principles are used in fibre-optic light transmission, Section 10.10, and indeed, in observing the Becke line effect in transmission optical microscopy.

In order to increase the efficiency of the stimulated emission process, the ends of the laser diode are polished to give them mirror-like characteristics and therefore some back-reflection through the p–n junction; with rough end surfaces, some of the emitted light will be lost by random scattering processes. Semiconductor lasers can be very powerful for their size since the entire p–n junction region can act as a source of stimulated emission rather than the dilute dopants, occupying $\leq 1\%$ of the volume, in Cr, Nd lasers.

The design of an LED is, in principle, simpler than that of a laser diode since it is no longer necessary to have an optical cavity that produces coherent emitted radiation. The radiation emitted by an LED results mainly from spontaneous transitions and the colour is controlled by the composition of the semiconductor host material. For instance, gallium arsenide–gallium phosphide solid solutions of formula $GaAs_{1-x}P_x$ can be prepared whose band gap varies from 2.26 eV for GaP ($x = 1$) to 1.44 eV for GaAs ($x = 0$), giving light wavelengths in the range 5500–8600 Å.

10.9 Photodetectors

So far, we have considered the emission of light from luminescent sources and laser systems in which the key step is energy release resulting from decay or relaxation of electrons from upper to lower energy levels. The reverse process forms the basis of photodetectors, which are used to detect and measure the light, or other radiation, emitted from a source. Various kinds of photodetector are available using several mechanisms for photodetection.

In semiconductor detectors, electrons are promoted from the valence band to the conduction band by absorption of light, which is detected as a change in either the resistance or voltage of the semiconductor detector. The *junction photodiode* is basically a p–n junction whose electrical characteristics are modified on exposure to light. As discussed earlier, in the absence of light or a dc bias, the junction reaches an equilibrium state in which diffusion of majority carriers in one direction is balanced by drift of minority carriers in the opposite direction, Fig. 10.15. The equilibrium can be disturbed either by application of a forward or reverse

bias or upon exposure to radiation. Upon illumination, some electrons in the depletion region are promoted to the conduction band, leading to electron–hole pair formation and because of the spontaneous junction bias that has developed with an energy barrier eV_0, the electron–hole pairs separate and drift in opposite directions. Since electrons drift from the p side to the n side and, vice versa, holes drift from the n side to the p side, this reverse current flows in the opposite direction to the spontaneous diffusion current and has the effect of reducing the energy barrier for spontaneous diffusion from eV_0 to $(eV_0 - eV_{h\nu})$, where $eV_{h\nu}$ is the energy of the incident radiation. Consequently, more majority carriers are able to cross the junction in both directions until equilibrium is re-established in which the photocurrent is exactly balanced by the spontaneous forward current. In open circuit, no net dc current flows but the junction barrier height is reduced from eV_0 to $(eV_0 - eV_{h\nu})$; this energy difference is measured when the junction is operating in *photovoltaic mode*.

In the other mode of operation, the *photoconductive mode*, a reverse bias applied across the junction has the effect of increasing the spontaneous diffusion barrier from eV_0 to $(eV_0 + eV_{\text{reverse bias}})$. Upon illumination, electron–hole pairs are again generated and drift apart, leading to a dc drift current when the junction is in closed circuit. In photoconductive mode, the photogenerated current is measured rather than the junction voltage. The advantage of the photoconductive mode is that the response time is more rapid than in the photovoltaic mode because with the higher junction potential, the transit time for charge carriers to drift across the junction and reach the electrodes is reduced; however, in this mode there is always 'noise' associated with the leakage current of minority carriers. Operation of a junction diode in the photoconductive mode is shown schematically in Fig. 10.17.

In addition to the basic p–n junction diode, more complex junction structures have been devised to decrease further the response time to illumination. In the *pin diode*, n- and p-type depletion regions are separated by a wide layer of undoped, intrinsic material to give a p–i–n junction. The effective depletion layer width is

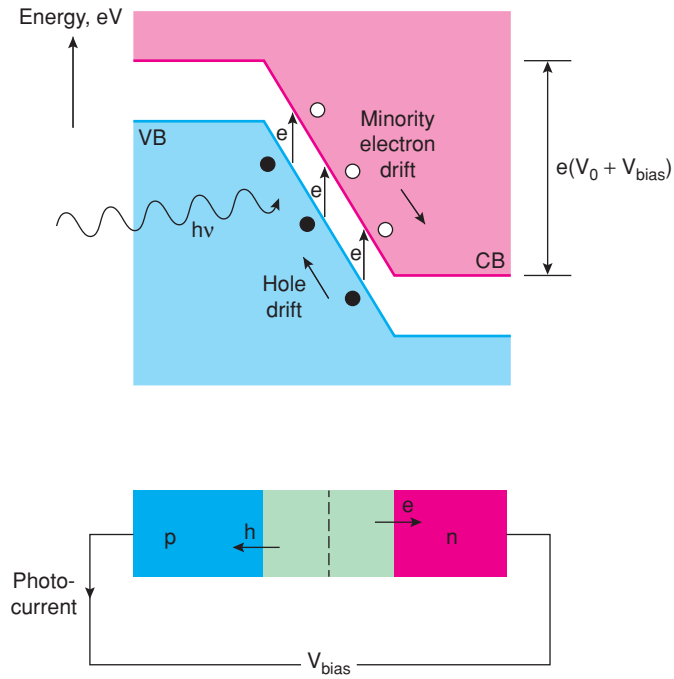


Figure 10.17 ***Operation of a junction diode in the photoconductive mode.***

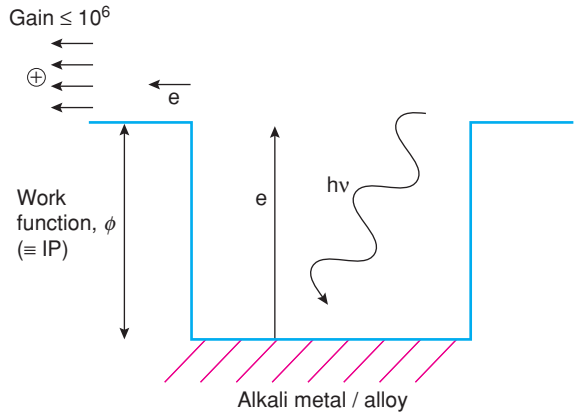


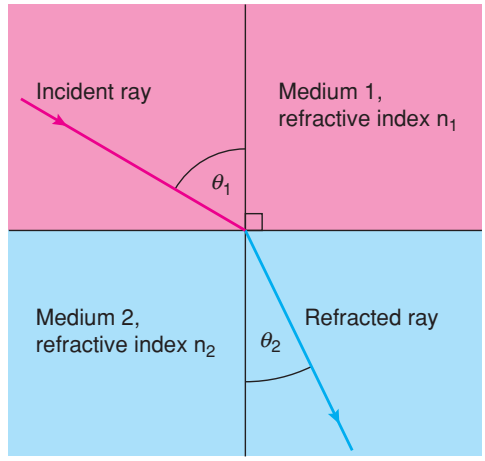
Figure 10.18 ***Operation of a photomultiplier based on photoemission.***

now much higher; if the dopant concentrations in the p- and n-type regions are small, the depletion region can be extended to cover the entire distance between the terminals on application of a small reverse bias potential. This has the effect of giving a very fast response to illumination. In the *avalanche photodiode*, the p–n junction is, by contrast, highly doped; the reverse bias that is applied is high, with the effect that minority carriers drifting across the junction have sufficient energy to activate further carriers and effectively give a current avalanche.

In *photoemissive detectors*, electrons are emitted upon irradiation of a material with a low *work function*, ϕ, which is equivalent to the ionisation potential of an isolated atom. Low-ϕ materials include alkali metals and their alloys and lanthanum hexaboride, LaB_6. Provided that the incident radiation has a minimum threshold energy $h\nu_{min}$ or maximum wavelength λ_{max}, electrons near the top of the conduction band absorb enough energy to be ionised. They are then accelerated towards the anode through an increasingly positive potential gradient which causes other electrons to be emitted leading to a cascade effect and an increase in current, or *gain*, by many orders of magnitude. This conceptually simple process forms the operational basis of the *photomultiplier*, Fig. 10.18.

10.10 Fibre-Optics

When light travels across an interface between two materials of different refractive index, n, it is *refracted* and bends towards the normal in the medium of higher n. The relation between the angle of incidence, θ_1, angle of refraction, θ_2, and the refractive indices of the two media, n_1 and n_2, is given by Snell's law, Fig. 10.19. This is the fundamental principle upon which fibre-optics is based since it is possible for light to pass down a fibre, usually made of glass, without escaping from the sides of the fibre. To achieve this, it is necessary to clad the fibre with a material of lower n and achieve the condition for *total internal reflection*, TIR, so that the light beam is internally reflected as it passes down the fibre and not refracted into the external medium, Fig. 10.20. As the light beam enters the fibre from the glass interface, it is refracted towards the normal and provided the angle, θ_2, is greater than the critical angle, θ_{crit}, light is reflected with 100% efficiency and passes down the fibre, essentially without loss.



Snell's law : $n_1 \sin \theta_1 = n_2 \sin \theta_2$

Figure 10.19 ***Snell's law governing the angles of incidence and refraction when light crosses an interface between media of different refractive index.***

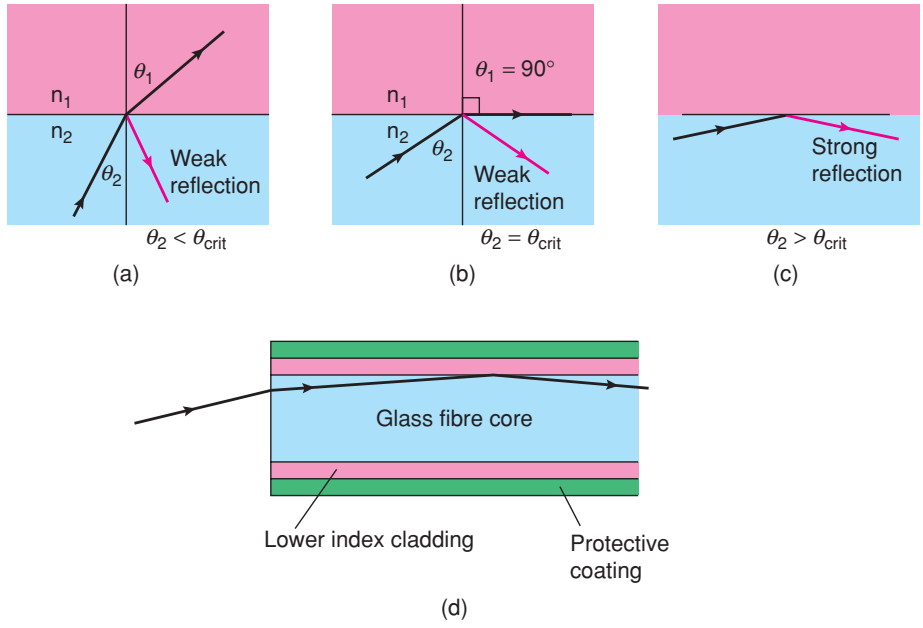


Figure 10.20 ***The confinement of a light beam to pass along an optical fibre by total internal reflection and its application in fibre-optic communications.***

10.11 Solar Cells

Conversion of solar energy into electricity is another, increasingly important, application of the p–n junction. Construction of a solar cell is based on the principles of the semiconductor diode shown in Fig. 10.15 and Fig. 10.17. Irradiation of the p–n junction generates electron–hole pairs which separate because they are attracted in opposite directions by the intrinsic field gradient produced at the junction. To generate a high gradient, the dopant level at the p–n junction is high, giving a depletion region ~500 nm thick. It is not necessary to connect the junction to an external power supply since the holes and electrons generated by irradiation separate spontaneously. The current that is generated is equivalent to application of a reverse bias potential since electrons drift to the n-type region and holes drift to the p-type region.

Further Reading

General

Books

D. M. Adams (1974), *Inorganic Solids. An Introduction to Concepts in Solid-State Structural Chemistry*, John Wiley & Sons, Inc., New York.

R. Cotterill (2008), *The Material World*, Cambridge University Press, Cambridge.

P. A. Cox (1987), *The Electronic Structure and Chemistry of Solids*, Oxford University Press, Oxford.

P. A. Cox (2010), *Transition Metal Oxides*, Oxford University Press, Oxford.

P. K. Davies, A. J. Jacobson, C. C. Torardi and T. A. Vanderah (eds) (1997), *Solid State Chemistry of Inorganic Materials*, MRS Symposium Proceedings Series, Vol. 453, Materials Research Society, Warrendale, PA.

B. D. Fahlman (2011), *Materials Chemistry*, 2nd edn, Springer, Dordrecht.

W. Gao and N. M. Sammes (1999), *An Introduction to Electronic and Ionic Materials*, World Scientific, Singapore.

A. Guinier (1980), *The Structure of Matter*, Arnold, London.

A. Guinier and R. Jullien (1989), *The Solid State: from Superconductors to Superalloys*, International Union of Crystallography, Oxford University Press, Oxford.

P. J. F. Harris (1999), *Carbon Nanotubes and Related Structures: New Materials for the Twenty-First Century*, Cambridge University Press, Cambridge.

J. M. Honig and C. N. R. Rao (eds) (1981), *Preparation and Characterisation of Materials*, Academic Press, New York.

W. D. Kingery, H. K. Bowen and D. R. Uhlmann (1976), *Introduction to Ceramics*, 2nd edn, John Wiley & Sons, Inc., New York.

W. J. Moore (1967), *Seven Solid States*, W. A. Benjamin, New York.

R. E. Newnham (1975), *Structure–Property Relations*, Springer, New York.

G. A. Ozin, A. C. Arsenault and L. Cademartiri (2009), *Nanochemistry: a Chemical Approach to Nanomaterials*, RSC Publishing, Cambridge.

S. K. Pati, T. Enoki, and C. N. R. Rao (2011), *Graphene and its Fascinating Attributes*, World Scientific, Singapore.

C. N. R. Rao (ed.) (1993), *Chemistry of Advanced Materials*, Blackwell, Oxford.

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

C. N. R. Rao and J. Gopalakrishnan (1997), *New Directions in Solid State Chemistry*, Cambridge University Press, Cambridge.
K. J. Rao (ed.) (1995), *Perspectives in Solid State Chemistry*, Narosa Publications, New Delhi.
R. E. Reed-Hill (1964), *Physical Metallurgy Principles*, Van Nostrand Reinhold, New York.
D. Sangeeta and J. R. LaGraff (2005), *Inorganic Materials Chemistry Desk Reference*, 2nd edn, CRC Press, Boca Raton, FL.
L. E. Smart and E. A. Moore (2012), *Solid State Chemistry*, CRC Press, Boca Raton, FL.
R. J. D. Tilley (2004), *Understanding Solids: the Science of Materials*, John Wiley & Sons, Ltd, Chichester.
M. T. Weller (1994), *Inorganic Materials Chemistry*, Oxford University Press, Oxford.
M. A. White (2012), *Physical Properties of Materials*, CRC Press, Boca Raton, FL.
A. Wold and K. Dwight (1993), *Solid State Chemistry: Synthesis, Structure, and Properties of Selected Oxides and Sulfides*, Chapman and Hall, New York.
J. D. Wright (1995), *Molecular Crystals*, 2nd edn, Cambridge University Press, Cambridge.
H. Yanagida, K. Koumoto and M. Miyayama (1996), *The Chemistry of Ceramics*, John Wiley & Sons, Inc., New York.

Reviews and research publications

N. S. P. Bhuranesh and J. Gopalakrishnan (1997), Solid state chemistry of early transition metal oxides containing d^0 and d^1 cations, *J. Mater. Chem.*, **7**, 2297.
F. J. DiSalvo (1990), Solid state chemistry – a rediscovered chemical frontier, *Science*, **247**, 649.

Chapters 1, 2 and 3

Books

F. Agullo-Lopez, C. R. A. Catlow and P. D. Townsend (1988), *Point Defects in Materials*, Academic Press, New York.
S. L. Altmann (1994), *Band Theory of Solids: an Introduction from the Point of View of Symmetry*, Oxford University Press, Oxford.
L. W. Barr and A. B. Lidiard (1970), Defects in ionic crystals, in *Physical Chemistry: an Advanced Treatise*, Vol. 10 (eds H. Eyring, D. Henderson and W. Jost), Academic Press, New York.
J. K. Burdett (1995), *Chemical Bonding in Solids*, Oxford University Press, Oxford.
J. K. Burdett (1997) *Chemical Bonds: a Dialog*, John Wiley & Sons, Ltd, Chichester.
C. R. A. Catlow (1997), *Computer Modelling in Crystallography*, Academic Press, New York.
J. M. Dubois (2005) *Useful Quasicrystals*, World Scientific, Singapore.
L. Eyring and M. O'Keeffe (eds) (1970), *The Chemistry of Extended Defects in Non-Metallic Compounds*, North-Holland, Amsterdam.
N. N. Greenwood (1968), *Ionic Crystals, Lattice Defects and Nonstoichiometry*, Butterworths, London.
D. T. Griffin (1992), *Silicate Crystal Chemistry*, Oxford University Press, Oxford.
R. M. Hazen and L. W. Finger (1984), *Comparative Crystal Chemistry*, John Wiley & Sons, Ltd, Chichester.
R. Hoffmann (1988), *Solids and Surfaces: a Chemist's View of Bonding in Extended Structures*, Wiley-VCH Inc., New York.
B. G. Hyde and S. Andersson (1989), *Inorganic Crystal Structures*, John Wiley & Sons, Inc., New York.
H. W. Jaffe (1988), *Crystal Chemistry and Refractivity*, Cambridge University Press, Cambridge.
C. Janot (1992), *Quasicrystals: a Primer*, Oxford University Press, Oxford.

P. Kofstad (1972), *Non-Stoichiometry, Electrical Conductivity and Diffusion in Binary Metal Oxides*, John Wiley & Sons, Inc., New York.
F. A. Kröger (1974), *The Chemistry of Imperfect Crystals*, North-Holland, Amsterdam.
F. Liebau (1985), *Structural Chemistry of Silicates*, Springer, Berlin.
H. D. Megaw (1973), *Crystal Structures, a Working Approach*, Saunders, Philadelphia, PA.
U. Müller (1992), *Inorganic Structural Chemistry*, John Wiley & Sons, Ltd, Chichester.
D. Pettifor (1995), *Bonding and Structure of Molecules and Solids*, Oxford University Press, Oxford.
J. C. Phillips and G. Lucovsky (2009), *Bonds and Bands in Semiconductors*, 2nd edn, Momentum Press, New York.
A. L. G. Rees (1954), *Chemistry of the Defect Solid State*, Methuen, London.
M. Senechal (1995), *Quasicrystals and Geometry*, Cambridge University Press, Cambridge.
R. J. D. Tilley (1987), *Defect Crystal Chemistry and its Applications*, Blackie, Glasgow.
A. F. Wells (1984), *Structural Inorganic Chemistry*, Oxford University Press, Oxford.
M. J. Winter (1994) *Chemical Bonding*, Oxford University Press, Oxford.
R. W. G. Wyckoff (1971), *Crystal Structures*, Vols 1–6, John Wiley & Sons, Inc., New York.
E. Zolotoyabko (2011) *Basic Concepts of Crystallography*, Wiley-VCH Verlag GmbH, Weinheim.

Reviews and research publications

D. Arnold and F. Morrison (2009), B-cation effects in relaxor and ferroelectric tetragonal tungsten bronzes, *J. Mater. Chem.*, **19**, 6485.
T. P. Beales (1998), Properties and substitutional chemistry of layered lead cuprate superconductors, *J. Mater. Chem.*, **8**, 1.
L. Bindi, P. J. Steinhardt, N. Yao and P. J. Lu (2009), Natural quasicrystals, *Science*, **324**, 1306.
J. K. Burdett (1994), Some aspects of the metal–insulator transition, *Chem. Soc. Rev.*, **23**, 299.
A. Corma and A. Martinez (1995), Zeolites and zeotypes as catalysts, *Adv. Mater.*, **7**, 137.
J. Darriet and M. A. Subramanian (1995), Structural relationships between compounds based on the stacking of mixed layers related to hexagonal perovskite structures, *J. Mater. Chem.*, **5**, 543.
W. I. F. David, R. M. Ibberson, J. C. Matthewman, *et al.* (1991), Crystal structure and bonding of ordered C_{60}, *Nature*, **353**, 147.
R. C. Ewing, W. J. Weber and J. Lian (2004), Nuclear waste disposal – pyrochlore ($A_2B_2O_7$): nuclear waste form for the immobilisation of plutonium and 'minor' actinides, *J. Appl. Phys*, **95**, 5949.
L. Howes (2011), Quasicrystals scoop prize, *Chem. World*, November, 38.
M. Kawaguchi (1997), B/C/N semiconductors based on the graphite network, *Adv. Mater.*, **9**, 615.
H. W. Kroto, A. W. Allaf and S. P. Balm (1991), C_{60}: buckminsterfullerene, *Chem. Rev.*, **91**, 1213.
J. Nolting (1970), Disorder in solids, *Angew. Chem. Int. Ed. Engl.*, **9**, 989.
B. Raveau, C. Michel, M. Hervieu and A. Maignan (1995), Crystal chemistry of superconducting mercury-based cuprates and oxycarbonates, *J. Mater. Chem.*, **5**, 803.
M. J. Rosseinsky (1995), Fullerene intercalation chemistry, *J. Mater. Chem.*, **5**, 1497.
R. D. Shannon and C. T. Prewitt (1969–1976), Effective ionic radii in oxides and fluorides, *Acta Crystallogr.*, **1969**, B25, 925; Revised values of effective ionic radii, *Acta Crystallogr.*, 1970, **B26**, 1046; Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, *Acta Crystallogr.*, 1976, **A32**, 751.
M. A. Subramanian, G. Aravamudan and G. V. Subba Rao (1983), Oxide pyrochlores, *Prog. Solid State Chem.*, **15**, 55.
R. Tenne, Doped and heteroatom-containing fullerene-like structures and nanotubes (1995), *Adv. Mater.*, **7**, 965.

J. M. Thomas (1974), Topography and topology in solid state chemistry, *Philos. Trans. R. Soc. London, Ser. A*, **277**, 251.

Chapter 4

Books

R. M. Barrer (1982), *Hydrothermal Chemistry of Zeolites*, Academic Press, London.
L. C. Klein (ed.) (1988), *Sol–Gel Technology for Thin Films, Fibers, Preforms, Electronics and Speciality Shapes*, Noyes Publications, Park Ridge, NJ.
J. Rouxel, M. Tournoux and R. Brec (1994), *Soft Chemistry Routes to New Materials: Chimie Douce*, Materials Science Forum, Vols 152 and 153, Transtech Publications, Zürich.
U. Schubert and N. Hüsing (2012), *Synthesis of Inorganic Materials*, 3rd edn, Wiley-VCH Verlag GmbH, Weinheim.
D. H. Segal (1994), *Chemical Synthesis of Advanced Ceramic Materials*, Cambridge University Press, Cambridge.
R. A. Street (1991), *Hydrogenated Amorphous Silicon*, Cambridge University Press, Cambridge.
J. D. Wright and N. A. J. M. Sommerdijk (2001), *Sol–Gel Materials: Chemistry and Applications*, Gordon and Breach Science, New York.

Reviews and research publications

J.C. Angus and C. C. Hayman (1988), Low pressure growth of diamond, *Science*, **241**, 913.
S. T. Aruna and A. S. Mukasyan (2008), Combustion synthesis and nanomaterials, *Curr. Opin. Solid State Mater. Sci.*, **12**, 44.
M. N. R. Ashfold, P. W. May, C. A. Rego and N. M. Everitt (1994), Thin film diamond by chemical vapour deposition methods, *Chem. Soc. Rev.*, **23**, 21.
J. H. Bang, W. H. Suh and K. S. Suslick (2008), Quantum dots from chemical aerosol flow synthesis: preparation, characterisation and cellular imaging, *Chem. Mater.*, **20**, 4033.
A. R. Barron (1996), CVD of SiO_2 and related materials: an overview, *Adv. Mater. Opt. Electron.*, **6**, 101.
I. Bilecka and M. Niederberger (2010), Microwave chemistry for inorganic nanomaterials synthesis, *Nanoscale*, **2**, 1358.
J. Bill and F. Aldinger (1995), Precursor-derived covalent ceramics, *Adv. Mater.*, **7**, 775.
M. Binnewies, R. Glaum, M. Schmidt and P. Schmidt (2013), Chemical vapour transport reactions – a historical review, *Z. Anorg. Allg. Chem.*, **639**, 219.
J.-O. Carlsson and U. Jansson (1993), Progress in chemical vapour deposition, *Prog. Solid State Chem.*, **22**, 237.
L.-H. Chen, *et al.* (2012), Hierarchically structured zeolites synthesis, mass transport properties and applications, *J. Mater. Chem.*, **22**, 17381.
W. Choi, *et al.* (2010), Synthesis of graphene and its applications: A review, *Crit. Rev. Solid State and Mat. Sci.*, **35**, 52.
A. H. Cowley and R. A. Jones (1989), Single source III/V precursors, *Angew. Chem. Int. Ed. Engl.*, **28**, 1208.
P. B. Davies and P. M. Martineau (1992), Diagnostics and modelling of silane and methane plasma CVD processes, *Adv. Mater.*, **4**, 729.
M. K. Devaraju, D. Rangappa and I. Honma (2012), Controlled synthesis of nanocrystalline Li_2MnSiO_4 particles for high capacity cathode application in lithium-ion batteries, *Chem. Commun.*, **48**, 2698.

W. Du, X. Jiang and L. Zhu (2013), From graphite to graphene: direct liquid-phase exfoliation of graphite to produce single- and few-layered pristine graphene, *J. Mater. Chem.*, **Al**, 12695.
E. G. Gillan and R. B. Kaner (1996), Synthesis of refractory ceramics via metathesis reactions between solid state precursors, *Chem. Mater.*, **8**, 333.
C. Greaves and M. G. Francesconi (1998), Fluorine insertion in inorganic materials, *Curr. Opin. Solid State Mater. Sci.*, **3**, 144.
A. Gurav, T. Kodas, T. Pluym and Y. Xiong (1993), Aerosol processing of materials, *Aerosol Sci. Technol.*, **19**, 411.
J. F. Hamet and B. Mercey (1998), Laser ablation for the growth of materials, *Curr. Opin. Solid State Mater. Sci.*, **3**, 144.
M. J. Hampden-Smith and T. V. Kodas (1995), Chemical vapour deposition of metals, *Chem. Vap. Deposition*, **1**, 8–23; 39–48.
M. L. Hand, M. C. Stennett and N. C. Hyatt (2012), Rapid low temperature synthesis of a titanate pyrochlore by molten salt mediated reaction, *J. Eur. Ceram. Soc.*, **32**, 3211.
A. C. Jones (1997), Developments in metalorganic precursors for semiconductor growth from the vapour phase, *Chem. Soc. Rev.*, **26**, 101.
A. I. Khan and D. O'Hare (2002), Intercalation chemistry of layered double hydroxides: recent developments and applications, *J. Mater. Chem.*, **12**, 3191.
L. C. Klein (1985), Sol–gel processing of silicates, *Annu. Rev. Mater. Sci.*, **15**, 227.
F. F. Lange (1989), Powder processing science and technology for increased reliability, *J. Am. Ceram. Soc.*, **72**, 3.
H. Lange, G. Wötting and G. Winter (1991), Silicon nitride – from powder synthesis to ceramic materials, *Angew. Chem. Int. Ed. Engl.*, **30**, 1579.
H. Schäfer (1971), Preparative solid state chemistry: the present position. *Angew. Chem.*, **10**, 43.
M. Leskelä and M. Ritala (2003), Atomic layer deposition chemistry: recent developments and future challenges, *Angew. Chem. Int. Ed.*, **42**, 5548.
J. Livage (1996), Sol–gel processes, *Curr. Opin. Solid State Mater. Sci.*, **1**, 132.
J. Livage, M. Henry and C. Sanchez (1988), Sol–gel chemistry of transition metal oxides, *Prog. Solid State Chem.*, **18**, 259.
V. Miikkulainen, M. Leskelä, M. Ritala and R. L. Puurunen (2013), Crystallinity of inorganic films grown by ALD; overview and general trends, *J. Appl. Phys.*, **113**, 021301.
H. Morkoç (2001), III-nitride semiconductor growth by MBE: recent issues, *J. Mater. Sci. Mater. Electron.*, **12**, 677.
R. E. Morris and S. J. Weigel (1997), The synthesis of molecular sieves from non-aqueous solvents, *Chem. Soc. Rev.*, **26**, 309.
L. Niinstö, Atomic layer epitaxy (1998), *Curr. Opin. Solid State Mater. Sci.*, **3**, 147.
P. O'Brien and R. Nomura (1995), Single molecule precursors for deposition of semiconductors by MOCVD, *J. Mater. Chem.*, **5**, 1761.
K. Ohtsuka (1997), Preparation and properties of two-dimensional microporous pillared interlayered solids, *Chem. Mater.*, **9**, 2039.
R. T. Paine and C. K. Narula (1990), Synthetic routes to boron nitride, *Chem. Rev.*, **90**, 73.
I. P. Parkin (2002), Solvent-free reactions in the solid state: solid state metathesis, *Transition. Met. Chem.*, **27**, 569.
J. Prado-Gonjal, A. M. Arévalo-López and E. Moran (2011), Microwave-assisted synthesis: a fast and efficient route to produce $LaMO_3$: M = Al, Cr, Mn, Fe, Co perovskite, *Mater. Res. Bull.*, **46**, 222.
C. N. R. Rao (1993), Chemical synthesis of solid inorganic materials, *Mater. Sci. Eng.*, **B18**, 1–21

C. N. R. Rao and J. Rouxel (1996), Synthesis and reactivity of solids, *Curr. Opin. Solid State Mater. Sci.*, **1**, 225.
K. J. Rao, B. Vaidhyanathan, M. Ganguli and P. A. Ramakrishnan (1999), Synthesis of inorganic solids using microwaves, *Chem. Mater.*, **11**, 882.
J. Robertson (1991), Hard amorphous (diamond-like) carbons, *Prog. Solid State Chem.*, **21**, 199.
R. Roy, S. Komarneni and L. J. Yang (1985), Controlled microwave heating and melting of gels, *J. Am. Ceram. Soc.*, **68**, 392.
R. Safi and H. Shokrollahi (2012), Physics, chemistry and synthesis methods of nanostructured bismuth ferrite, $BiFeO_3$, as a ferroelectric-magnetic material, *Prog. Solid State Chem.*, **40**, 6.
A. B. Sawaska, M. Takamatsu and T. Akashi (1994), Shock compression synthesis of diamond, *Adv. Mater.*, **6**, 346.
R. Schöllhorn (1996), Intercalation systems as nanostructured functional materials, *Chem. Mater.*, **8**, 1747.
P. G. Schultz and X.-D, Xiang (1998), Combinatorial approaches to materials science, *Curr. Opin. Solid State Mater. Sci.*, **3**, 153.
D. Segal (1997), Chemical synthesis of ceramic materials, *J. Mater. Chem.*, **7**, 1297.
V. Šepelák, S. M. Becker, I. Bergmann, *et al.* (2012), Nonequilibrium structure of Zn_2SnO_4 spinel nanoparticles, *J. Mater. Chem.*, **22**, 3117.
J. Shim, *et al.* (2013), Atomic layer deposition of thin-film ceramic electrodes for high performance fuel cells, *J. Mater. Chem.*, **Al**, 12695.
M. K. Singh, Y. Yang and C. G. Takoudis (2009), Synthesis of multifunctional multiferroic materials from metal organics, *Coord. Chem. Rev.*, **253**, 2920.
F. T. J. Smith (1989), Low pressure organometallic epitaxy of the III–V compounds, *Prog. Solid State Chem.*, **19**, 111.
K. E. Spear (1989), Diamond – ceramic coating of the future, *J. Am. Ceram. Soc.*, **72**, 171.
R. Strobel and S. E. Pratsinis (2007), Flame aerosol synthesis of smart nanostructured materials, *J. Mater. Chem.* **17**, 4743.
K. S. Suslick and G. J. Price (1999), Applications of ultrasound to materials chemistry, *Annu. Rev. Mater. Sci.*, **29**, 295.
M. T. Weller and S. E. Dann (1998), Hydrothermal synthesis of zeolites, *Curr. Opin. Solid State Mater. Sci.*, **3**, 137.
H. C. Yi and J. J. Moore (1990), Self-propagating high-temperature (combustion) synthesis (SHS) of powder-compacted materials, *J. Mater. Sci.*, **25**, 1159.
F. Zaera (2008), The surface chemistry of thin film atomic layer deposition processes for electronic device manufacturing, *J. Mater. Chem.*, **18**, 3521.
V. V. Zyranov (2008), Mechanochemical synthesis of complex oxides, *Russ. Chem. Rev.*, **7**, 105.

Chapter 5

Books

G. E. Bacon (1975), *Neutron Diffraction*, Oxford University Press, Oxford.
P. Coppens (1992), *Synchrotron Radiation Crystallography*, Academic Press, New York.
B. D. Cullity (1978), *Elements of X-ray Diffraction*, 2nd edn, Addison Wesley Reading, MA.
W. I. F. David, K. Shankland, L. B. McCuster and C. Baerlocher (eds) (2003), *Structure Determination from Powder Diffraction Data*, IUCr Monographs on Crystallography 13, International Union of Crystallography, Oxford University Press, Oxford.

M. De Graef and M. E. McHenry (2012), *Structure of Materials: an Introduction to Crystallography, Diffraction and Symmetry*, 2nd edn, Cambridge University Press, Cambridge.
L. S. Dent Glasser (1977), *Crystallography and its Applications*, Van Nostrand Reinhold, New York.
R. E. Dinnebier and S. J. L. Billinge (eds) (2008), *Powder Diffraction, Theory and Practice*, RSC Publishing, Cambridge.
T. Hahn (ed.) (2005), *International Tables for X-ray Crystallography*, Vol. A, 5th edn, John Wiley & Sons, Ltd, Chichester.
C. Hammond (1997), *The Basics of Crystallography and Diffraction*, IUCr texts on Crystallography 3, International Union of Crystallography, Oxford University Press, Oxford.
A. Kelly and K. M. Knowles (2012), *Crystallography and Crystal Defects*, 2nd edn, John Wiley & Sons, Ltd, Chichester.
D. McKie and C. McKie (1986), *Essentials of Crystallography*, Blackwell, Oxford.
V. K. Pecharsky and P. Y. Zavalij (2003), *Fundamentals of Powder Diffraction and Structural Characterisation of Materials*, Kluwer, Dordrecht.
J. Wormald (1973), *Diffraction Methods*, Clarendon Press, Oxford.
R. A. Young (ed.) (1993), *The Rietveld Method*, IUCr Monographs on Crystallography 5, International Union of Crystallography, Oxford University Press, Oxford.
E. Zolotoyabko (2011), *Basic Concepts of Crystallography*, Wiley-VCH Verlag GmbH, Weinheim.

Chapter 6

Books

D. C. Apperley, R. K. Harris and P. Hodgkinson (2012), *Solid-State NMR: Basic Principles and Practice*, Momentum Press, New York.
G. M. Bancroft (1973), *Mössbauer Spectroscopy*, McGraw-Hill, New York.
A. Bianconi, L. Incoccia and S. Stipcich (eds) (1983), *EXAFS and Near Edge Structure*, Springer, Berlin.
R. W. Cahn (ed.) (2005), *Concise Encyclopedia of Materials Characterization*, 2nd edn, Elsevier, Amsterdam.
A. K. Cheetham and P. Day (eds) (1987), *Solid State Chemistry: Techniques*, Oxford University Press, Oxford.
T. Daniels (1973), *Thermal Analysis*, Kogan Page, London.
M. J. Duer (2004), *Introduction to Solid-State NMR Spectroscopy*, Blackwell, Oxford.
G. Engelhardt and D. Michel (1987), *High Resolution Solid State NMR of Silicates and Zeolites*, John Wiley & Sons, Ltd, Chichester.
J. García Solé, L. E. Bausá and D. Jaque (2005), *An Introduction to the Optical Spectroscopy of Inorganic Solids*, John Wiley & Sons, Ltd, Chichester.
J. T. Grant and D. Briggs (2003), *Surface Analysis by Auger and X-Ray Photoelectron Spectroscopy*, IM Publications, Chichester.
P. J. Grundy and G. A. Jones (1976), *Electron Microscopy in the Study of Materials*, Arnold, London.
N. H. Hartshorne and A. Stuart (1971), *Practical Optical Crystallography*, Arnold, London.
P. C. H. Mitchell, S. F. Parker, A. J. Ramirez-Cuesta and J. Tomkinson (2005), *Vibrational Spectroscopy with Neutrons*, World Scientific, Singapore.
R. E. Newnham and R. Roy (1975), Structural characterisation of solids, in *Treatise on Solid State Chemistry*, Vol. 2 (ed. N. B. Hannay), Plenum Press, New York, p. 437.
K. Oura, V. G. Lifshits, A. A. Saranin, *et al.* (2003), *Surface Science: an Introduction* Springer, Berlin.
K. Siegbahn, C. Nordling, A. Fahlman, *et al.* (1967), *ESCA: Atomic, Molecular and Solid State Structure Studied by Means of Electron Spectroscopy*, Almquist and Wicksells, Uppsala.

P. van der Heide (2012), *X-Ray Photoelectron Spectroscopy: an Introduction to Principles and Practice*, John Wiley & Sons, Inc., Hoboken, NJ.
W. W. Wendlandt (1974), *Thermal Methods of Analysis*, John Wiley & Sons, Inc., New York.
D. B. Williams and C. B. Carter (1996), *Transmission Electron Microscopy*, Plenum Press, New York.
A. H. Zewail and J. M. Thomas (2010), *4D Electron Microscopy*, Imperial College Press, London.
J. Zussman (ed.) (1977), *Physical Methods in Determinative Mineralogy*, Academic Press, New York.

Reviews and research publications

C. W. T. Bulle-Lieuwma, W. Coene and A. F. de Jong (1991), High resolution electron microscopy of semiconductors and metals, *Adv. Mater.*, **3**, 368.
L. Eyring (1980), The application of high-resolution electron microscopy to problems in solid state chemistry, *J. Chem. Educ.*, **57**, 565.
S. Hackwood and R. G. Linford (1981), Physical techniques for the study of solid electrolytes, *Chem. Rev.*, **81**, 327.
W. Jones and J. M. Thomas (1979), Applications of electron microscopy to organic solid state chemistry, *Prog. Solid State Chem.*, **12**, 101.
L. Kihlborg (1990), High resolution electron microscopy in solid state chemistry, *Prog. Solid State Chem.*, **20**, 101.
M. W. Roberts (1981), Photoelectron spectroscopy and surface chemistry, *Chem. Br.*, **510**.
M. Rühle (1997), Microscopy of structural ceramics, *Adv. Mater.*, **9**, 195.
B. G. Williams (1986–1987), Electron energy loss spectroscopy, *Prog. Solid State Chem.*, **17**, 87.

Chapter 7

Books

A. M. Alper (ed.) (1971), *High Temperature Oxides*, Vols 1–4, Academic Press, New York.
A. M. Alper (1976), *Phase Diagrams*, Vols 1–5, Academic Press, New York.
C. G. Bergeron and S. H. Risbud (1984), *Introduction to Phase Equilibria in Ceramics*, American Ceramic Society, Westerville, OH.
W. G. Ernst (1976), *Petrologic Phase Equilibria*, Freeman, San Francisco, CA.
A. Findlay (1951), *The Phase Rule and Its Applications*, 9th edn, Dover, New York.
P. Gordon (1968), *Principles of Phase Diagrams in Materials Systems*, McGraw-Hill, New York.
Phase Diagrams for Ceramists, 1964 edn, 1969 Suppl., 1975 Suppl., 1981 Suppl., American Ceramic Society, Columbus, OH: the standard work of reference for phase diagrams of non-metallic inorganic materials.

Reviews and research publications

S. V. Ushakov and A. Navrotsky (2012), Experimental approaches to the thermodynamics of ceramics above 1500 °C, *J. Am. Ceram. Soc.*, **95**, 1463.

Chapter 8

Books

V. S. Bagotsky (2012), *Fuel Cells: Problems and Solutions*, 2nd edn, John Wiley & Sons, Inc., Hoboken, NJ.
S. Blundell (2009), *Magnetism*, Oxford University Press, Oxford.

P. G. Bruce (ed.) (1995), *Solid State Electrochemistry*, Cambridge University Press, Cambridge.
J. C. Burfoot and G. W. Taylor (1979), *Polar Dielectrics and Their Applications*, University of California Press, Berkeley, CA.
G. Burns (1992), *High-Temperature Superconductivity: an Introduction*, Academic Press, San Diego, CA.
P. Colomban (ed.) (1992), *Proton Conductors: Solids, Membranes and Gels – Materials and Devices*, Cambridge University Press, Cambridge.
P. A. Cox (1987), *The Electronic Structure and Chemistry of Solids*, Oxford University Press, Oxford.
R. M. Dell and D. A J. Rand (2001), *Understanding Batteries*, Royal Society of Chemistry, Cambridge.
P. P. Edwards and C. N. R. Rao (eds) (1995), *Metal–Insulator Transitions Revisited*, Taylor and Francis, London.
D. A. Fraser (1986), *The Physics of Semiconductor Devices*, Oxford University Press, Oxford.
W. Gao and N. M. Sammes (2006), *An Introduction to Electronic and Ionic Materials*, World Scientific, Singapore.
P. J. Gellings and H. J. M. Bouwmeester (1997), *Solid State Electrochemistry*, CRC Press, Boca Raton, FL.
K. Huang and J. B. Goodenough (2009), *Solid Oxide Fuel Cell Technology*, Woodhead Publishing, Cambridge.
R. A. Huggins (2009), *Advanced Batteries*, Springer, New York.
D. Jiles (1994), *Electronic Properties of Materials*, Chapman and Hall, London.
A. K. Jonscher (1996), *Universal Relaxation Law*, Chelsea Dielectrics Press.
T. Kudo and K. Fueki (1990), *Solid State Ionics*, VCH Verlag GmbH, Weinheim.
P. Knauth and M. L. DiVona (eds) (2012) *Solid State Proton Conductors*, John Wiley & Sons, Ltd, Chichester.
J. Maier (2004) *Physical Chemistry of Ionic Materials: Ions and Electrons in Solids*, John Wiley & Sons, Ltd, Chichester.
M. D. McCluskey and E. E. Haller (2012), *Dopants and Defects in Semiconductors*, CRC Press, Boca Raton, FL.
N. B. McKeown (1998), *Phthalocyanine Materials: Synthesis, Structure and Function*, Cambridge University Press, Cambridge.
D. V. Morgan and K. Board (1983), *An Introduction to Semiconductor Microtechnology*, John Wiley & Sons, Ltd, Chichester.
A. J. Moulson and J. M. Herbert (2003), *Electroceramics: Materials, Properties, Applications*, 2nd edn, John Wiley & Sons, Ltd, Chichester.
G. Schopf and G. Kossmehl (1997), *Polythiophenes – Electrically Conductive Polymers*, Springer, Berlin.
R. B. Seymour (1981), *Conductive Polymers*, Plenum Press, New York.
L. Solymar and D. Walsh (2004), *Electrical Properties of Materials*, Oxford University Press, Oxford.
J. L. Sudworth and A. R. Tilley (1985), *The Sodium Sulphur Battery*, Chapman & Hall, London.
S. M. Sze (1985), *Semiconductor Devices*, John Wiley & Sons, Inc., New York.
G. Vidali (1993), *Superconductivity: the Next Revolution?*, Cambridge University Press, Cambridge.
C. A. Vincent (1984), *Modern Batteries: an Introduction to Electrochemical Power Sources*, Arnold, London.

Reviews and research publications

H. Asahi (1997), Self-organised quantum wires and dots in III–V semiconductors, *Adv. Mater.*, **9**, 1019.
A. Battacharya and A. De (1996), Conducting composites of polypyrrole and polyaniline, *Prog. Solid State Chem.*, **24**, 141.
K. Bechgard and D. Jerome (1982), Organic superconductors, *Sci. Am.*, **247**, 50.
D. Bloor (1982), Plastics that conduct electricity, *New Sci.*, **1981**, 577.
P. G. Bruce, S. A. Freunberger, L. J. Hardwick and J.-M. Tarascon (2012), Li-O_2 and Li-S batteries with high energy storage, *Nature Materials*, **11**, 19.
M. R. Bryce and L. C. Murphy (1984), Organic metals, *Nature*, **309**, 119.

R. M. Dell (ed.) (1996), Materials for electrochemical power systems: a discussion, *Philos. Trans. R. Soc. London, Ser. A*, **354**, 1513.
V. Etacheri, R. Marom, R. Elazari, G. Salitra and D. Aurbach (2011), Challenges in the development of advanced Li-ion batteries: a review, *Energy and Environ.*, **4**, 3243.
O. Gunnarsson (1997), Superconductivity in fullerides, *Rev. Mod. Phys.*, **69**, 575.
P. He, H. Yu, De Li and H. Zhou (2012), Layered lithium transition metal oxide cathodes towards high energy lithium-ion batteries, *J. Mater. Chem.*, **22**, 3680.
H. E. Katz (1997), Organic molecular solids as thin film transistor semiconductors, *J. Mater. Chem.*, **7**, 369.
V. V. Kharton, F. M. B. Marques and A. Atkinson (2004), Transport properties of solid oxide electrolyte ceramics: a brief review, *Solid State Ionics*, **174**, 135.
P. Knauth (2009), Inorganic solid Li ion conductors: An overview, *Solid State Ionics*, **180**, 911.
P. M. Levy and S. Zhang (1999), Spin dependent tunnelling, *Curr. Opin. Solid State Mater. Sci.*, **4**, 223.
N. Martin (2013), Tetrathiafulvalene: the advent of organic metals, *Chem. Commun.*, **49**, 7025.
P. J. Nigrey, D. McInnes, D. P. Nairs, *et al.* (1981), Lightweight rechargeable storage batteries using polyacetylene $(CH)_x$ as the cathode active material, *J. Electrochem. Soc.*, **128**, 1651.
S. C. O'Brien (1996), The chemistry of the semiconductor industry, *Chem. Soc. Rev.*, **25**, 393.
M. O'Keeffe and B. G. Hyde (1976), The solid electrolyte transition and melting in salts, *Philos. Mag.*, **B3**, 219.
J. R. Owen (1997), Rechargeable lithium batteries, *Chem. Soc. Rev.*, **26**, 259.
C. N. R. Rao and A. K. Ganguli (1995), Structure–property relations in superconducting cuprates, *Chem. Soc. Rev.*, **24**, 1.
H. Rickert (1978), Solid ionic conductors – principles and applications, *Angew. Chem. Int. Ed. Engl.*, **17**, 37–46.
P. R. Slater, J. E. H. Sansom and J. R. Tolchard (2004), Development of apatite-type oxide ion conductors, *Chem. Rec.*, **4**, 373.
K. Takeda (2013), Progress and prospective of solid-state lithium batteries, *Acta Mat.*, **61**, 759.
T. Takahashi (1973), Solid silver ion conductors, *J. Appl. Electrochem.*, **3**, 79.
K. Tanigaki and K. Prassides (1995), Conducting and superconducting properties of alkali metal C_{60} fullerides, *J. Mater. Chem.*, **5**, 1515.
J. Tirado (2003), Inorganic materials for the negative electrode of lithium-ion batteries: state-of-the-art and future prospects, *Mater. Sci. Eng.*, **R40**, 103.
A. E. Underhill (1992), Molecular metals and superconductors, *J. Mater. Chem.*, **2**, 1.
A. Vayrynen and J. Salminen (2012), Lithium ion battery production, *J. Chem. Thermodynamics*, **46**, 80.
A. R. West (1991), Solid electrolytes, *J. Mater. Chem.*, **1**, 157.
M. Winter, J. O. Besenhard, M. E. Spahr and P. Novák (1998), Insertion electrode materials for rechargeable lithium batteries, *Adv. Mater.*, **10**, 725.
F. Wudl, G. M. Smith and E. J. Hufnagel (1970), bis-1,3-dithiolium chloride: an unusually stable organic radical cation, *Chem. Commun.*, 1453.

Chapter 9

Books

S. Blundell (2012), *Magnetism*, Oxford University Press, Oxford.
D. J. Craik (ed.) (1975), *Magnetic Oxides, Parts 1 and 2*, John Wiley & Sons, Ltd, London.

J. Crangle (1991), *Solid State Magnetism*, Arnold, London.
J. B. Goodenough (1963), *Magnetism and the Chemical Bond*, John Wiley & Sons, Inc., New York.
D. Jiles (1991), *Introduction to Magnetism and Magnetic Materials*, Chapman and Hall, London.
K. J. Standley (1972), *Oxide Magnetic Materials*, Clarendon Press, Oxford.
R. S. Tebble and D. J. Craik (1979), *Magnetic Materials*, John Wiley & Sons, Inc., New York.
R. Valenzuela (1994), *Magnetic Ceramics*, Cambridge University Press, Cambridge.

Reviews and research publications

J. M. D. Coey, M. Viret and S. von Molnár (1999), Mixed-valence manganites, *Adv. Phys.*, **48**, 167.
D. Gatteschi (1994), Molecular magnetism, *Adv. Mater.*, **6**, 635.
H. Hibst (1982), Hexagonal ferrites from melts and aqueous solutions, magnetic recording materials, *Angew. Chem. Int. Ed. Engl.*, **21**, 270.
J. S. Moodera, J. Nassar and G. Mathon (1999), Spin-tunneling in ferromagnetic junctions, *Annu. Rev. Mater. Sci.*, **29**, 381.
C. N. R. Rao and A. K. Cheetham (1977), Giant magnetoresistance, charge ordering and related aspects of manganates and other oxide systems, *Adv. Mater.*, **9**, 1009.
M. Sugimoto (1999), The past, present and future of ferrites, *J. Am. Ceram. Soc.*, **82**, 269.
E. Y. Tsymbal and D. G. Pettifor (2001), Perspectives of giant magnetoresistance, *Solid State Phys.*, **48**, 113.
W. E. Wallace (1985), Rare earth–transition metal permanent magnet materials, *Prog. Solid State Chem.*, **16**, 127.
C.-H. Yang, *et al.* (2012), Doping $BiFeO_3$: approaches and enhanced functionality, *Phys. Chem. Chem. Phys.*, **14**, 15953.

Chapter 10

Books

G. Blasse and B. C. Grabmaier (1994), *Luminescent Materials*, Springer-Verlag.
H. L. Burrus (1972), *Lamp Phosphors*, Harlequin Mills and Boon Ltd.
T. R. Evans (1982), *Applications of Lasers to Chemical Problems*, John Wiley Interscience, New York.
A. Kitai (2011), *Principles of Solar Cells, LEDs and Diodes*, John Wiley & Sons, Ltd, Chichester.
S. W. S. McKeever (1988), *Thermoluminescence of Solids*. Cambridge University Press, Cambridge.
J. Watson (1989), *Optoelectronics*, van Nostrand Reinhold.

Reviews and research publications

G. Blasse (1988), Luminescence of inorganic solids, *Prog. Solid State Chem.*, **18**, 79.
J. A. DeLuca (1980), An introduction to luminescence in inorganic solids, *J. Chem. Educ.*, **57**, 541.
M. Gratzel (2004), Conversion of sunlight to electric power by nanocrystalline dye-sensitized solar cells, *J. Photochem. Photobiol. A: Chemistry*, **164**, 3.
T. H. Maiman (1960), Stimulated optical radiation in ruby, *Nature*, **187**, 493.
Y. Tokura (1998), Optical and magnetic properties of transition metal oxides, *Curr. Opin. Solid State Mater. Sci.*, **3**, 175.

Appendix A

Interplanar Spacings and Unit Cell Volumes

The value of d, the perpendicular distance between adjacent planes in the set (hkl), may be calculated using the following equations:

Cubic
$$\frac{1}{d^2} = \frac{h^2 + k^2 + l^2}{a^2}$$

Tetragonal
$$\frac{1}{d^2} = \frac{h^2 + k^2}{a^2} + \frac{l^2}{c^2}$$

Orthorhombic
$$\frac{1}{d^2} = \frac{h^2}{a^2} + \frac{h^2}{b^2} + \frac{l^2}{c^2}$$

Hexagonal
$$\frac{1}{d^2} = \frac{4}{3}\left(\frac{h^2 + hk + k^2}{a^2}\right) + \frac{l^2}{c^2}$$

Monoclinic
$$\frac{1}{d^2} = \frac{1}{\sin^2\beta}\left(\frac{h^2}{a^2} + \frac{k^2\sin^2\beta}{b^2} + \frac{l^2}{c^2} - \frac{2hl\cos\beta}{ac}\right)$$

Triclinic
$$\begin{aligned}\frac{1}{d^2} = \frac{1}{V^2}[&h^2b^2c^2\sin^2\alpha + k^2a^2c^2\sin^2\beta \\ &+ l^2a^2b^2\sin^2\gamma + 2hkabc^2(\cos\alpha\cos\beta - \cos\gamma) \\ &+ 2kla^2bc(\cos\beta\cos\gamma - \cos\alpha) \\ &+ 2hlab^2c(\cos\alpha\cos\gamma - \cos\beta)]\end{aligned}$$

where V is the cell volume.

The unit cell volumes are given by

Cubic	$V = a^3$
Tetragonal	$V = a^2c$
Orthorhombic	$V = abc$
Hexagonal	$V = \left(\sqrt{3}a^2c\right)/2 = 0.866a^2c$
Monoclinic	$V = abc\sin\beta$
Triclinic	$V = abc\left(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma\right)^{\frac{1}{2}}$

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

Appendix B

Model Building

The construction of two types of model, based on spheres and polyhedra, is outlined.

Equipment Needed

Polystyrene spheres (100–200), any diameter but 30 mm is convenient. Alternatively, small plastic spheres.
Solvent (5–10 ml) in a bottle with dropper or paint brush; chloroform is suitable.
Card paper (2–3 m^2), perhaps sheets of different colour.
Glue stick.

Sphere Packing Arrangements

In order to join polystyrene spheres, glue may be used but is often messy. An alternative is to use a solvent in which polystyrene dissolves. A small amount of solvent is spotted or brushed onto a sphere and a second sphere is placed in contact with the first in the region of the solvent. If too much solvent is used it may run and spoil the appearance of the sphere. The spheres usually stick immediately on contact but the bond hardens only after leaving for a few hours. It is therefore often better to construct larger three-dimensional models in stages.

It should be relatively easy to stick together spheres to form *cp* layers and then build up layers in *hcp* and *ccp* stacking sequences. Some useful exercises are as follows:

To show the relation between a *ccp* structure and an *fcc* unit cell

Stick together six spheres to form that part of a *cp* layer shown in Fig. A.1(a). Place a single sphere in the layer above, as shown in (b). Repeat steps (a) and (b) so that two identical arrangements, (b), are obtained. Allow the models to harden. Orient one model as in (c) and place the second model over the first as in (d). The resulting structure should be a face centred cubic unit cell with spheres in corner and face centre positions. The *fcc* cell therefore contains *cp* layers parallel to the {111} planes of the unit cell. On standing the model on one corner so that the *cp* layers are horizontal, it should be apparent that the stacking sequence is ABC.

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

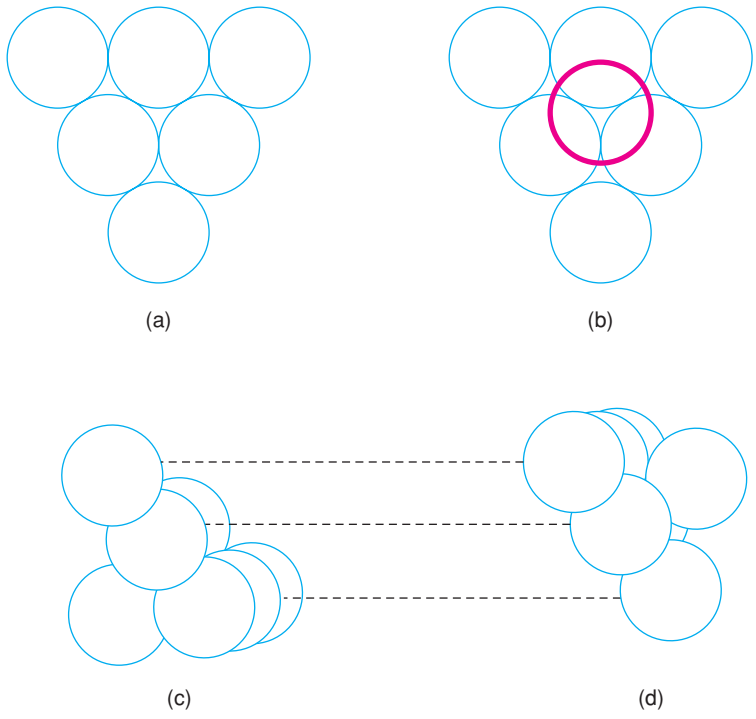


Figure A.1 ***To show the relation between* ccp *and the* fcc *unit cell.***

To show the four orientations of the *cp* layers in an *fcc* unit cell

In a *ccp* array of spheres, *cp* layers occur in four orientations. This may be shown by constructing the pyramid shown in Fig. A.2. First, construct a triangular base of six spheres, as in Fig. A.1(a); then add three spheres in the layer above and one sphere to form the apex to give the pyramid shown in Fig. A.2. Check that the packing sequence is ABC, i.e. *ccp*. The pyramid may now be reoriented so that any of its other three faces forms the base and an identical structure results. The four equivalent orientations of the pyramid correspond, therefore, to the four orientations of *cp* layers in a *ccp* structure, i.e. the *cp* layers are the $\{111\}$ planes in the *fcc* unit cell.

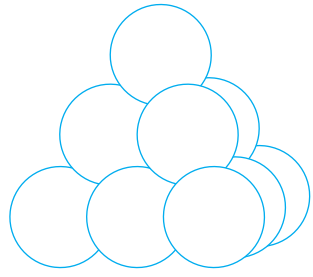


Figure A.2 ***The* cp *layers in a* ccp *structure.***

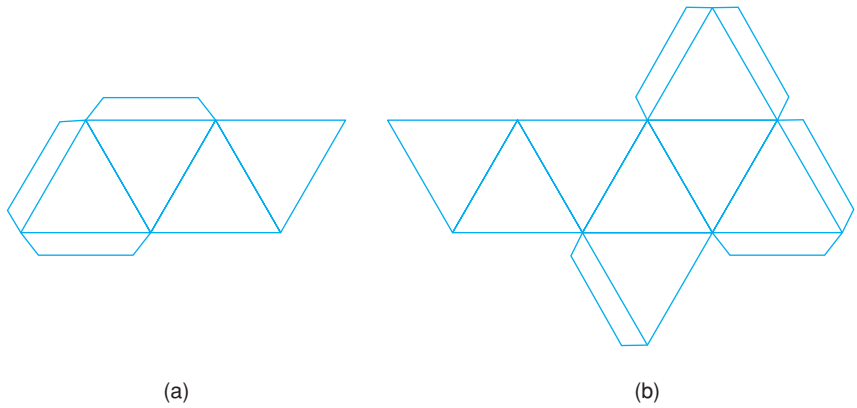


Figure A.3 ***Templates for making tetrahedra and octahedra.***

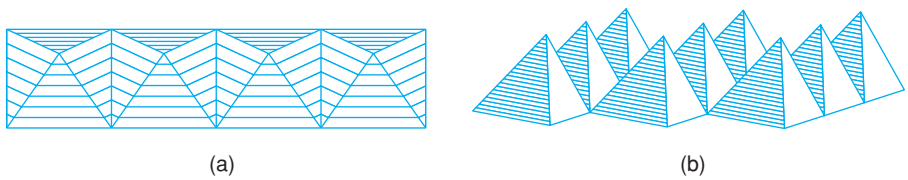


Figure A.4 ***Some polyhedral linkages.***

To show the *cp* layers and unit cell of an *hcp* stacking sequence

One major difference between *ccp* and *hcp* is that in an *hcp* array of spheres, only one set of *cp* layers occurs, parallel to the basal plane of the hexagonal unit cell. To see this, construct separately three triangular bases of six spheres, as in Fig. A.1(a). Place the three layers on top of each other so that the third layer is directly above the first layer, giving the ABA sequence. It should be clear that *cp* layers are present in only one orientation, parallel to the base.

To construct the *hcp* unit cell, see Fig. 1.21(a, b); you need eight spheres to form the diamond-shaped top and bottom faces of the unit cell (A layers), which are separated by one sphere in the B layer.

Polyhedral structures

Tetrahedra, octahedra and any other polyhedra may be made from card paper. Using templates such as in Fig. A.3, the polyhedra may be copied, cut out, folded and glued at the tabs. A convenient size for making fairly rigid polyhedra is to make the polyhedron edge ~5 cm. The polyhedra may be linked by glueing corners, edges or faces; glueing polyhedra by their corners only is a little tricky. Some polyhedral linkages are given in Fig. A.4; others are shown as 3D crystal structures throughout Chapter 1.

Appendix C

Geometrical Considerations in Crystal Chemistry

Notes on the Geometry of Tetrahedra and Octahedra

Relation of a tetrahedron to a cube

In Fig. A.5 is shown a cube with a tetrahedron inside; the centre of the tetrahedron, M, is at the cube body centre and the corners of the tetrahedron are at four alternate corners of the cube, X_1 to X_4. Using this relation, it is relatively easy to make calculations on the geometry of tetrahedra.

Relation between distances M–X and X–X in a tetrahedron

Let the cube edge have length a. From Pythagoras, the distance X–X is the face diagonal of the cube, i.e. $X - X = \sqrt{2}a$. The distance M–X is half the body diagonal of the cube, i.e. $M - X = \sqrt{3}a/2$. Hence, the ratio $XX/MX = \sqrt{2}a/(\sqrt{3}a/2) = \sqrt{8/3} = 1.633$. In silica structures, the basic building blocks are SiO_4 tetrahedra. In these, usually Si–O ≈ 1.62 Å. Therefore, the tetrahedron edges are O–O = 1.633(Si–O) = 2.65 Å.

Angle XMX of a tetrahedron

Using the above result and the cosine equation [Fig. A.5(a) and (b)]:

$$(X_1X_2)^2 = (MX_1)^2 + (MX_2)^2 - 2\,(MX_1)\,(MX_2)\cos \angle X_1MX_2$$

Therefore,

$$(1.633MX_1)^2 = 2(MX_1)^2 - 2(MX_1)^2 \cos \angle X_1MX_2$$

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

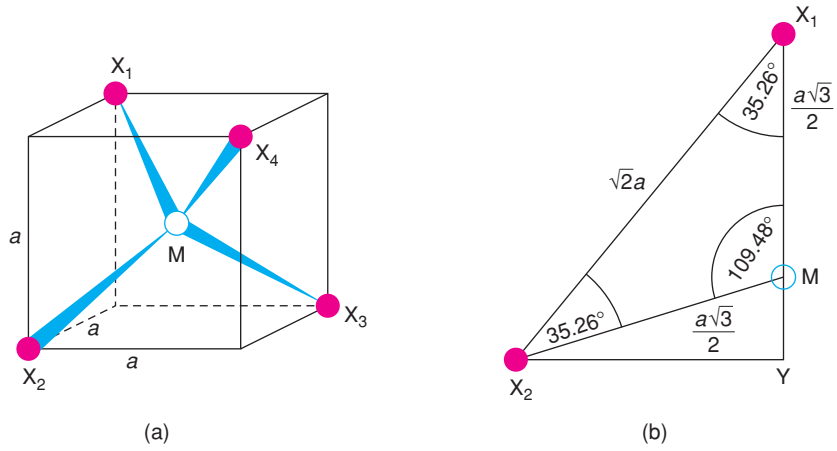


Figure A.5 ***Relation of a tetrahedron to a cube.***

and so

$$\angle \mathrm{XMX} = 109.48^\circ$$

Symmetry of a tetrahedron

A tetrahedron has a threefold rotation axis along each of the M–X directions, i.e. along each cube body diagonal. Both the tetrahedron and the cube therefore possess four threefold axes. The cube also possesses three fourfold rotation axes, passing through each pair of opposite cube faces. In the tetrahedron, these axes reduce to twofold rotation axes, but retain fourfold character by becoming fourfold inversion axes, $\bar{4}$ (i.e. rotation by 90° followed by inversion through the centre of the tetrahedron).

Centre of gravity of a tetrahedron

For this, we need to know the vertical height of position M above the triangular base of the tetrahedron. Consider the triangular section X_1X_2M [Fig. A.5(b)]. Point Y, the extension of the line X_1–M, lies in the centre of the base X_2, X_3, X_4 [Fig. A.5(a), not shown]. Hence, distance M–Y gives the height of the centre of gravity above the base. Since

$$\angle \mathrm{X_1MX_2} = 109.48^\circ$$

then

$$\angle \mathrm{MX_1X_2} = \angle \mathrm{MX_2X_1} = 35.26^\circ$$

and so

$$\cos 35.26^\circ = \frac{\mathrm{X_1Y}}{\mathrm{X_1X_2}} = \frac{\mathrm{X_1Y}}{\sqrt{2}a}$$

i.e.

$$X_1Y = 1.155a$$

Therefore,

$$YM = X_1Y - X_1M = 0.289a$$

and

$$\frac{YM}{YX_1} = \frac{0.289a}{1.155a} = 0.25$$

i.e. the centre of gravity of a tetrahedron, given by position M, is at one-quarter of the vertical height above the (or any) base and in the direction of the apex.

Relation of an octahedron to a cube

In Fig. A.6 is shown an octahedron centred at the body centre of a cube with the corners of the octahedron at the face centres of the cube. If the cube has edge a the distance MX in the octahedron is $a/2$. From Pythagoras, distance XX is $a/\sqrt{2}$.

The octahedron has fourfold rotation axes (3) parallel to each XMX straight line. These axes coincide with the fourfold axes of the cube that pass through opposite cube faces. The cube also has threefold axes (4) which pass through opposite cube corners, along the body diagonal of the cube. In the octahedron, these threefold axes pass through pairs of opposite faces.

Hexagonal unit cell: proof that the axial ratio, *c/a*, ideally equals 1.633

An *hcp* array of atoms contains atoms at the corners and inside the unit cell at $\frac{1}{3}\ \frac{2}{3}\ \frac{1}{2}$ (Fig. 1.21). Along the a and b edges, atoms are in contact; hence a is equal to the atom diameter. For the purpose of calculating c/a, 'tetrahedral arrays' of atoms may be identified, as in Fig. A.7, in which two tetrahedra share a common

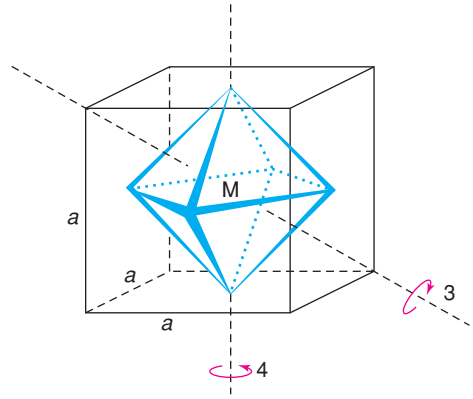


Figure A.6 ***Relation of an octahedron to a cube.***

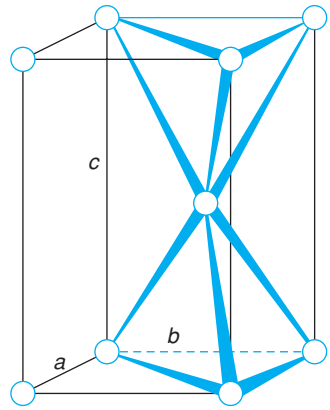


Figure A.7 ***Hexagonal unit cell; c/a = 1.633.***

corner (i.e. the atom at $\frac{1}{3}\ \frac{2}{3}\ \frac{1}{2}$). Hence, the c dimension equals twice the vertical height of such a tetrahedron. Therefore:

$$\frac{c}{a} = \frac{2\mathrm{X_1Y}}{\mathrm{X_1X_2}} = \frac{2 \times 1.155a}{\sqrt{2}a} = 1.633$$

Appendix D

How to Recognise Close Packed (Eutactic) Structures

On encountering a new or unfamiliar structure, a common questions is: Is it close packed? These notes are intended to provide guidelines for helping to decide whether or not a structure is *cp*, although it is difficult to give hard and fast rules. We focus mainly on non-molecular, non-metallic ionic structures such as oxides, halides and sulfides, but similar criteria apply to metallic structures and, in some cases, molecular structures. The requirements may be summarised as follows:

(a) For non-metallic, close packed (or eutactic) oxide structures of general formula $A_xB_yO_z$ we are interested mainly in whether the oxide ions form a *cp* array. Sometimes the A or B atoms themselves may form a *cp* array or indeed, mixed anion-cation arrangements may form a *cp* array, as in perovskites.
(b) Coordination numbers, given by the number of anions surrounding a cation are ideally, either four or six.

Point (a) must hold in a eutactic *cp* structure but the difficulty is often to visualise a sufficiently large part of the structure (often more than one unit cell is needed) that the packing arrangements of like atoms can be seen. The best test, I find, is to look for one close packed layer, i.e. to choose a central atom and look for six equidistant neighbours of the same type such that a coplanar, hexagonal arrangement exists, as in Fig. 7.1. If such a layer can be identified, it is usually relatively easy to find six other atoms in the layers on either side that complete the *cp* requirement of 12 neighbour atoms. Often, different layers may be seen in a structure, some of which are *cp* and others non-*cp*. Quite often, layers may be identified that contain both anions and cations but are not *cp*. Thus, the unit cell faces in the rock salt structure may be regarded as layers of atoms but they do not constitute *cp* layers because each Na^+ (or Cl^-) has only four Na^+ (or Cl^-) neighbours in the layer. By contrast, the layers parallel to the $\{111\}$ planes of the rock salt structure are *cp* because (i) each Cl^- (or Na^+) has six like neighbours and (ii) layers of ions of one kind may be identified which alternate with layers of the other kind.

Point (b) refers to the fact that in a *cp* array of anions, tetrahedral and octahedral interstitial sites exist for the cations. The cation coordination number is therefore four or six. The coordination number of cations surrounding the anions is more variable, as shown by the following examples:

(i) Na_2O (antifluorite) has *ccp* oxide ions with tetrahedral coordination for Na^+. The coordination number of oxygen (by Na^+) is eight.

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

(ii) NiAs has *hcp* As^{2-} ions with octahedrally coordinated Ni^{2+} ions. The coordination number of arsenic (by nickel) is also six but is trigonal prismatic.
(iii) Cu_2O (cuprite) has *ccp* Cu^{+} ions with tetrahedral O^{2-} ions. The coordination number of Cu^{+} (by O^{2-}) is two (linear).

As stated above, in structures containing large cations such as perovskites, the large cations may take the place of anions in a *cp* array, in which case the cation coordination number is ideally 12 but can be lower, such as 10 or 8 in distorted *cp* structures. In summary, in order fully to understand and classify a structure as belonging to the *cp* family of structures, it is necessary to examine, in particular, the anion, packing arrangement and the cation coordination numbers. It is also worthwhile to consider the cation packing arrangements, in isolation, and also the combined packing arrangements of anions and large cations.

Appendix E

Positive and Negative Atomic Coordinates

In assigning fractional coordinates to the positions of atoms in a unit cell, it is customary to include in the cell those atoms whose coordinates lie between 0 and 0.999. Atoms with negative coordinates or with coordinates ≥ 1 then lie in adjacent unit cells. This is illustrated in Fig. A.8, in which four unit cells are drawn. Let us consider the top right cell; its origin is given as the solid circle and positive directions of x, y and z are indicated by arrows. The diagram shows the coordinates of one pair of opposite face centre positions in each of the four cells. Relative to the origin of our chosen cell, all the face centre positions shown at the right hand side have $x = {}^1/_2$ and those at the left have $x = -{}^1/_2$. All positions in the diagram have a positive y value of ${}^1/_2$. The z values are in three sets, with values of 1, 0 and –1. In considering the contents of our unit cell, the position at ${}^1/_2{}^1/_20$ is regarded as belonging to the cell, but all other positions shown belong to neighbouring cells.

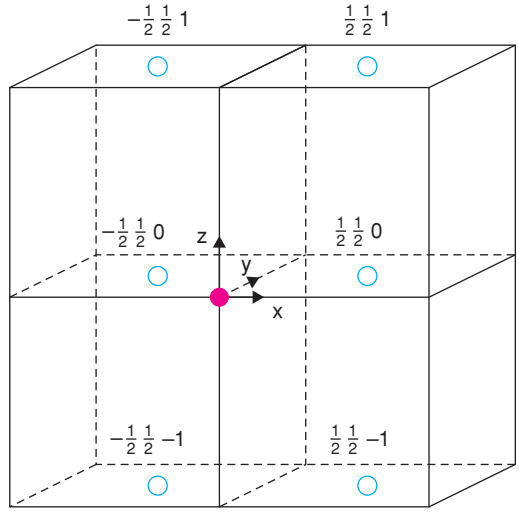


Figure A.8 ***Positive and negative atomic coordinates.***

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

Appendix F

The Elements and Some of Their Properties

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

Element	Symbol	Atomic weight[a]	Electronic configuration[b]	Atomic number	Main oxidation state(s)[c]	Some typical bond lengths–oxygen (Å) and coordination numbers[d]	Melting point (°C)	Crystallographic data: unit cell or structure; lattice parameters (Å); temperature[e]
Actinium	Ac	(227)	(Rn)$6d^1 7s^2$	89	3		1050	*fcc*; 5.311 (RT)
Aluminium	Al	26.98	(Ne)$3s^2 3p^1$	13	3	Al–O(4) 1.79; Al–O(6) 1.93	660	*fcc*; 4.0495 (25 °C)
Americium	Am	(243)	(Rn) $5f^7 7s^2$	95	3, 4	Am(III)–O(6) 2.40; Am(IV)–O(8) 2.35	850	hex, ABAC; 3.642, 11.76 (RT)
Antimony	Sb	121.75	(Kr)$4d^{10} 5s^2 5p^3$	51	3, 5		630	*R*3*m*; 4.5067, 57° 6.5′ (25 °C)
Argon	Ar	39.95	(Ne)$3s^3 3p^6$	18			−189	*fcc*; 5.42 (−233 °C)
Arsenic	As	74.92	(Ar)$3d^{10} 4s^2 4p^3$	33	3, 5	As(V)–O(4) 1.74; As(V)–O(6) 1.90	814	*R*3*m*; 4.131, 54° 10′ (25 °C)
Astatine	At	(210)	(Xe)$4f^{14} 5d^{10} 6s^2 6p^5$	85	1			
Barium	Ba	137.34	(Xe)$6s^2$	56	2	Ba–O(6) 2.76; Ba–O(12) 3.00	710	*bcc*; 5.019 (RT)
Berkelium	Bk	(247)	(Rn)$5f^8 6d^1 7s^2$	97	3(4)	Bk(III)–O(6) 2.36; Bk(IV)–O(8) 2.33		
Beryllium	Be	9.01	(He)$2s^2$	4	2	Be–O(3) 1.57; Be–O(4) 1.67	1280	*hcp*; 2.2856, 3.5832 (25 °C)
Bismuth	Bi	208.98	(Xe)$4f^{14} 5d^{10} 6s^2 6p^3$	83	1, 3(5)	Bi(III)–O(6) 2.42; Bi(III)–O(8) 2.51	271	*R*3*m*; 4.7457, 57° 14.2′ (31 °C)
Boron	B	10.81	(He)$2s^2 2p^1$	5	3	B–O(3) 1.42; B–O(4) 1.52	2300	tet; 8.73, 5.03 (RT)
Bromine	Br	79.91	(Ar)$3d^{10} 4s^2 4p^5$	35	1(3, 5, 7)	Br(VII)–O(4) 1.66	−7	orth; 4.48, 6.67, 8.72 (−150 °C)
Cadmium	Cd	112.40	(Kr) $4d^{10} 5s^2$	48	2	Cd–O(4) 2.19; Cd–O(6) 2.35	321	*hcp*; 2.9788 5.6167 (21 °C)
Caesium	Cs	132.91	(Xe) $6s^1$	55	1	Cs–O(8) 3.22; Cs–O(12) 3.30	29	*bcc*; 6.14 (−10 °C)
Calcium	Ca	40.08	(Ar)$4s^2$	20	2	Ca–O(6) 2.40; Ca–O(8) 2.47	850	*fcc*; 5.582 (18 °C)
Californium	Cf	(249)	(Rn)$5f^{10} 7s^2$	98	3	Cf–O(6) 2.35		
Carbon	C	12.01	(He)$2s^2 2p^2$	6	4	C–O(3) 1.32	3500	hex, graphite; 2.4612, 6.7079 (RT)
Cerium	Ce	140.12	(Xe)$4f^2 6s^2$	58	3, 4	Ce(III)–O(9) 2.55; Ce(IV)–O(8) 2.33	804	*fcc*, 5.1604 (20 °C)
Chlorine	Cl	35.46	(Ne)$3s^2 3p^5$	17	1(3, 5, 7)	Cl(V)–O(3) 1.52; Cl(VII)–O(4) 1.60	−101	tet; 8.56, 6.12 (−185 °C)
Chromium	Cr	52.00	(Ar)$3d^5 4s^1$	24	3, 6	Cr(III)–O(6) 2.02; Cr(VI)–O(4) 1.70	1900	*bcc*; 2.8846 (20 °C)
Cobalt	Co	58.93	(Ar)$3d^7 4s^2$	27	2, 3	Co(II)–O(6) 2.05–2.14; Co(III)–O(6) 1.93–2 01	1492	*hcp*; 2.507, 4.069 (RT)

Copper	Cu	63.54	(Ar) $3d^{10}4s^{1}$	29	1, 2	Cu(I)–O(2) 1.86; Cu(II)–O(6) 1.97–2.66	1083	*fcc*; 3.6147 (20 °C)
Curium	Cm	(247)	(Rn)$5f^{7}6d^{1}7s^{2}$	96	3(4)	Cm(III)–O(6) 2.38; Cm(IV)–O(8) 2.35		
Dysprosium	Dy	162.50	(Xe)$4f^{10}6s^{2}$	66	3(4)	Dy(III)–O(6) 2.31	1500	*hcp*; 3.5923, 5.6545 (20 °C)
Einsteinium	Es	(254)	(Rn)$5f^{11}7s^{2}$	99	3			
Erbium	Er	167.26	(Xe)$4f^{12}6s^{2}$	68	3	Er(III)–O(6) 2.29	1525	*hcp*; 3.559, 5.592 (20 °C)
Europium	Eu	151.96	(Xe)$4f^{7}6s^{2}$	63	2, 3	Eu(III)–O(6) 2.35; Eu(III)–O(7) 2.43	900	*bcc*; 4.578 (20 °C)
Fermium	Fm	(253)	(Rn)$5f^{12}7s^{2}$	100	3			
Fluorine	F	19.00	(He)$2s^{2}2p^{5}$	9	1		−220	
Francium	Fr	223	(Rn)$7s^{1}$	87	1			
Gadolinium	Gd	157.25	(Xe)$4f^{7}5d^{1}6s^{2}$	64	3	Gd–O(7) 2.44	1320	*hcp*; 3.6315, 5.777 (20 °C)
Gallium	Ga	69.72	(Ar)$3d^{10}4s^{2}4p^{1}$	31	(1)3	Ga–O(4) 1.87; Ga–O(6) 2.00	30	orth; 4.520, 7.661, 4.526 (20 °C)
Germanium	Ge	72.59	(Ar)$3d^{10}4s^{2}4p^{2}$	32	(2)4	Ge–O(4) 1.79; Ge–O (6) 1.94	958	diam; 5.6575 (20 °C)
Gold	Au	196.97	(Xe)$4f^{14}5d^{10}6s^{1}$	79	(1)(3)	Au(III)–O(4) 2.10	1063	*fcc*; 4.0783 (25 °C)
Hafnium	Hf	178.49	(Xe)$4f^{14}5d^{2}6s^{2}$	72	4	Hf–O(6) 2.11; Hf–O(8) 2.23	2000	*hcp*; 3.1946, 5.0511 (24 °C)
Helium	He	4.00	$1s^{2}$	2	—		−270	
Holmium	Ho	164.93	(Xe)$4f^{11}6s^{2}$	67	3	Ho–O(6) 2.30; Ho–O(8) 2.42	1500	*hcp*; 3.5761, 5.6174 (20 °C)
Hydrogen	H	1.01	$1s^{1}$	1	1	H–O 1.02–1.22	−259	hex; 3.75, 6.12 (−271 °C)
Indium	In	114.82	(Kr)$4d^{10}5s^{2}5p^{1}$	49	(1)3	In–O(6) 2.18; In–O(8) 2.32	156	tet; 3.2512, 4.9467 (20 °C)
Iodine	I	126.90	(Kr)$4d^{10}5s^{2}55p^{5}$	53	1(3, 5, 7)	I(V)–O(3) 1.83	114	orth; 4.792, 7.271, 9.773 (RT)
Iridium	Ir	192.22	(Xe)$4f^{14}5d^{7}6s^{2}$	77	3, 4(6)	Ir(III)–O(6) 2.13; Ir(IV)–O(6) 2.03	2443	*fcc*; 3.8389 (RT)
Iron	Fe	55.85	(Ar)$3d^{6}4s^{2}$	26	2, 3(4, 6)	Fe(II)–O(6) 2.18; Fe(III)–O(4) 1.86	1539	*bcc*; 2.8664 (20 °C)
Krypton	Kr	83.80	(Ar)$3d^{10}4s^{2}4p^{6}$	36			−157	*fcc* 5.68 (−191 °C)
Lanthanum	La	138.91	(Xe)$5d^{1}6s^{2}$	57	3	La–O(6) 2.46; La–O(10) 2.68	920	hex; ABAC; 3.770, 12.131 (20 °C)
Lawrencium	Lr	(257)	(Rn)$5f^{14}6d^{1}7s^{2}$	103	3			
Lead	Pb	207.19	(Xe)$4f^{14}5d^{10}6s^{2}6p^{2}$	82	2, 4	Pb(II)–O(4) 2.34; Pb(IV)–O(6) 2.18	327	*fcc*; 4.9502 (25 °C)
Lithium	Li	6.94	(He)$2s^{1}$	3	1	Li–O(4) 1.99; Li–O(6) 2.14	180	*bcc*; 3.5092 (20 °C)
Lutetium	Lu	174.97	(Xe)$4f^{14}5d^{1}6s^{2}$	71	3	Lu–O(6) 2.25; Lu–O(S) 2.37	1700	*hcp*; 3.5050, 5.5486 (20 °C)

(*Continued*)

Element	Symbol	Atomic weight[a]	Electronic configuration[b]	Atomic number	Main oxidation state(s)[c]	Some typical bond lengths–oxygen (Å) and coordination numbers[d]	Melting point (°C)	Crystallographic data: unit cell or structure; lattice parameters (Å); temperature[e]
Magnesium	Mg	24.31	(Ne)$3s^2$	12	2	Mg–O(4) 1.89; Mg–O(6) 2.12	650	*hcp*; 3.2094, 5.2105 (25 °C)
Manganese	Mn	54.94	(Ar)$3d^5 4s^2$	25	2, 3, 4, 7	Mn(II)–O(6) ~2.10; Mn(VII)–O(4) 1.66	1250	cub; 8.914 (25 °C)
Mendelevium	Md	(256)	(Rn)$5f^{13} 7s^2$	101	3			
Mercury	Hg	200.59	(Xe)$4f^{14} 5d^{10} 6s^2$	80	1, 2	Hg(I)–O(3) 2.37; Hg(II)–O(2) 2.09	−39	$R\bar{3}m$; 3.005, 70° 32′ (−46 °C)
Molybdenum	Mo	95.94	(Kr)$4d^5 5s^1$	42	(3, 4, 5)6	Mo(VI)–O(4) 1.82; Mo(VI)–O(6) 2.00	2620	*bcc*; 3.1469 (20 °C)
Neodymium	Nd	144.24	(Xe)$4f^4 6s^2$	60	3(4)	Nd–O(6) 2.40; Nd–O(8) 2.52	1024	hex; ABAC; 3.6582, 11.802 (20 °C)
Neon	Ne	20.18	(He)$2s^2 2p^6$	10	—		−249	*fcc*; 4.52 (−268 °C)
Neptunium	Np	(237)	(Rn)$5f^5 7s^2$	93	(2, 3)4(6, 7)	Np(II)–O(6) 2.60; Np(IV)–O(8) 2.38	640	orth; 4.723, 4.887, 6.663 (20 °C)
Nickel	Ni	58.71	(Ar)$3d^8 4s^2$	28	2(3)	Ni(II)–O(6) 2.10; Ni(III)–O(6) 1.98	1453	*fcc*; 3.524 (18 °C)
Niobium	Nb	92.91	(Kr)$4d^4 5s^1$	41	(4)5	Nb(V)–O(6) 2.04; Nb(V)–O(7) 2.06	2420	*bcc*; 3.006 (20 °C)
Nitrogen	N	14.01	(He)$2s^2 2p^3$	7	(2, 3, 4)5	N(V)–O(3) 1.28	−210	hex; 4.03,6.59 (−234 °C)
Nobelium	No	(253)	(Rn)$5f^{14} 7s^2$	102	3			
Osmium	Os	190.20	(Xe)$4f^{14} 5d^6 6s^2$	76	4(6)8	Os(IV)–O(6) 2.03	2700	*hcp*; 2.7353, 4.3191 (20 °C)
Oxygen	O	16.00	(He)$2s^2 2p^4$	8	(1)2		−219	cub; 6.83 (−225 °C)
Palladium	Pd	106.40	(Kr)$4d^{10}$	46	2(4)	Pd(II)–O(4) 2.04; Pd(IV)–O(6) 2.02	1552	*fcc*; 3.8907 (22 °C)
Phosphorus	P	30.97	(Ne)$3s^2 3p^3$	15	3, 5	P(V)–O(4) 1.57	44	orth; 3.32. 10.52, 4.39 (black) (RT)
Platinum	Pt	195.09	(Xe)$4f^{14} 5d^9 6s^1$	78	4(6)	Pt(IV)–O(6) 2.03	1769	*fcc*; 3.9239 (20 °C)
Plutonium	Pu	(242)	(Rn)$5f^6 7s^2$	94	3, 4, 6	Pu(III)–O(6) 2.40; Pu(IV)–O(8) 2.36		monocl; 6.18. 4.82. 10.97, 101.81° (21 °C)
Polonium	Po	(210)	(Xe)$4f^{14} 5d^{10} 6s^2 6p^4$	84	2, 4	Po(IV)–O(8) 2.50	254	cub; 3.345 (10 °C)
Potassium	K	39.10	(Ar)$4s^1$	19	1	K–O(6) 2.78; K–O(12) 3.00	63	*bcc*; 5.32 (20 °C)
Praseodymium	Pr	140.91	(Xe)$4f^3 6s^2$	59	3(4)	Pr(III)–O(6) 2.41; Pr(IV)–O(8) 2.39	935	hex, ABAC; 3.6702, 11.828 (20 °C)
Promethium	Pm	(147)	(Xe)$4f^5 6s^2$	61	3	Pm(III)–O(6) 2.38		

Protactinium	Pa	(231)	(Rn)$5f^2 6d^1 7s^2$	91	4, 5	Pa(IV)–O(8) 2.41; Pa(V)–O(9) 2.35	3000	tet; 3.935, 3.238 (RT)
Radium	Ra	(226)	(Rn)$7s^2$	88	2		700	
Radon	Rn	(222)	(Xe)$4f^{14}5d^{10}6s^2 6p^6$	86	—		−71	
Rhenium	Re	186.23	(Xe)$4f^{14}5d^5 6s^2$	75	3, 4, 5, 7	Re(IV)–O(6) 2.03; Re(VII)–O(6) 1.97	3170	*hcp*; 2.760, 4.458 (RT)
Rhodium	Rh	102.91	(Kr)$4d^8 5s^1$	45	3(4, 6)	Rh(III)–O(6) 2.07; Rh(IV)–O(6) 2.02	1960	*fcc*; 3.8044 (20 °C)
Rubidium	Rb	85.47	(Kr)$5s^1$	37	1	Rb–O(6) 2.89; Rb–O(12) 3.13	39	*bcc*; 5.70 (20 °C)
Ruthenium	Ru	101.07	(Kr)$4d^7 5s^1$	44	4(6)8	Ru(III)–O(6) 2.08; Ru(IV)–O(6) 2.02	2400	*hcp*; 2.7058, 4.2816 (25 °C)
Samarium	Sm	150.35	(Xe)$4f^6 6s^2$	62	(2)3	Sm(III)–O(6) 2.36; Sm(III)–O(8) 2.49	1052	R; 8.996, 23° 13′ (20 °C)
Scandium	Sc	44.96	(Ar)$3d^1 4s^2$	21	3	Sc–O(6) 2.13; Sc–O(8) 2.27	1400	*hcp*; 3.3080, 5.2653 (20 °C)
Selenium	Se	78.96	(Ar)$3d^{10}4s^2 4p^4$	34	(2) 4, 6	Se(VI)–O(4) 1.69	217	hex; 4.3656, 4.9590 (25 °C)
Silicon	Si	28.09	(Ne)$3s^2 3p^2$	14	4	Si–O(4) 1.66; Si–O(6) 1.80	1410	diam; 5.4305 (RT)
Silver	Ag	107.87	(Kr)$4d^{10}5s^1$	47	1(2)	Ag(I)–O(2) 2.07; Ag(I)–O(8) 2.70	961	*fcc*; 4.0857 (20 °C)
Sodium	Na	22.99	(Ne)$3s^1$	11	1	Na–O(4) 2.39; Na–O(9) 2.72	98	*bcc*; 4.2906 (20 °C)
Strontium	Sr	87.62	(Kr)$5s^2$	38	2	Sr–O(6) 2.56; Sr–O(12) 2.84	770	*fcc*; 6.0849 (25 °C)
Sulfur	S	32.06	(Ne)$3s^2 3p^4$	16	2, 4, 6	S(VI)–O(4) 1.52	119	orth; 10.414, 10.845, 24.369 (RT)
Tantalum	Ta	180.95	(Xe)$4f^{14}5d^3 6s^2$	73	(3, 4)5	Ta(III)–O(6) 2.07; Ta(V)–O(6) 2.04	3000	*bcc*; 3.3026 (20 °C)
Technetium	Tc	98.91	(Kr)$4d^6 5s^1$	43	(3)4(5, 6)7	Tc(IV)–O(6) 2.04	2700	*hcp*; 2.735, 4.388 (RT)
Tellurium	Te	127.60	(Kr)$4d^{10}5s^2 5p^4$	52	(2)4(6)	Te(IV)–O(3) 1.92	450	hex; 4.45666, 5.9268 (25 °C)
Terbium	Tb	158.93	(Xe)$4f^9 6s^2$	65	3, 4	Tb(III)–O(6) 2.32; Tb(IV)–O(8) 2.28	1450	*hcp*; 3.599, 5.696 (20 °C)
Thallium	Tl	204.37	(Xe)$4f^{14}5d^{10}6s^2 6p^1$	81	1(3)	Tl(I)–O(6) 2.90; Tl(I)–0(12) 3.16	304	*hcp*; 3.4566, 5.5248 (18 °C)
Thorium	Th	232.04	(Rn)$6d^2 7s^2$	90	(3)4	Th–O(6) 2.40; Th–O(9) 2.49	1700	*fcc*; 5.0843 (RT)
Thulium	Tm	168.93	(Xe)$4f^{13}6s^2$	69	(2)3	Tm(III)–O(6) 2.27; Tm(III)–O(8) 2.39	1600	*hcp*; 3.5372, 5.5619 (20 °C)

(*Continued*)

Element	Symbol	Atomic weight[a]	Electronic configuration[b]	Atomic number	Main oxidation state(s)[c]	Some typical bond lengths–oxygen (Å) and coordination numbers[d]	Melting point (°C)	Crystallographic data: unit cell or structure; lattice parameters (Å); temperature[e]
Tin	Sn	118.69	(Kr)$4d^{10}5s^25p^2$	50	2, 4	Sn(II)–O(8) 2.62; Sn(IV)–O(6) 2.09	232	tet; 5.8315, 3.1814 (25 °C)
Titanium	Ti	47.90	(Ar)$3d^24s^2$	22	(2)4	Ti(II)–O(6) 2.26; Ti(IV)–O(6) 2.01	1680	*hcp*; 2.9506, 4.6788 (25 °C)
Tungsten	W	183.85	(Xe)$4f^{14}5d^46s^2$	74	(4)6	W(VI)–O(4) 1.81; W(VI)–O(6) 1.98	3380	*bcc*; 3.1650 (25 °C)
Uranium	U	238.03	(Xe)$4f^{14}5d^46s^2$	92	(3)4(5)6	U(IV)–O(9) 2.45; U(VI)–O(4) 1.88	1133	orth; 2.854, 5.869, 4.955 (27 °C)
Vanadium	V	50.94	(Ar)$3d^34s^2$	23	2–5	V(II)–O(6) 2.19; V(V)–O(4) 1.76	1920	*bcc*; 3.028 (30 °C)
Xenon	Xe	131.30	(Kr)$4d^{10}5s^25p^6$	54	2, 4, 6		−112	*fcc*; 6.24 (−185 °C)
Ytterbium	Yb	173.04	(Xe)$4f^{14}6s^2$	70	2, 3	Yb(III)–O(6) 2.26; Yb(III)–O(9) 2.38	824	*fcc*; 5.481 (20 °C)
Yttrium	Y	88.91	(Kr)$4d^15s^2$	39	3	Y–O(6) 2.29; Y–O(9) 2.50	1500	*hcp*; 3.6451, 5.7305 (20 °C)
Zinc	Zn	65.37	(Ar)$3d^{10}4s^2$	30	2	Zn–O(4) 2.00; Zn–O(6) 2.15	419	*hcp*; 2.6649, 4.9468 (25 °C)
Zirconium	Zr	91.22	(Kr)$4d^25s^2$	40	4	Zr–O(6) 2.12; Zr–O(8) 2.24	1850	*hcp*; 3.2312, 5.1477 (25 °C)

[a]Based on the relative atomic mass of $^{12}C = 12.00$. These atomic weights are for natural isotopic abundances except for those given in parentheses, which refer to short-lived elements without a natural abundance. For these, the mass of the isotope with longest half-life is given.

[b]These are the ground state configurations of the elements. In some cases with transition elements, lanthanides and actinides, the difference in energy between the ground state and the first excited state is small.

[c]Most elements show a variety of unstable oxidation states in addition to the stable one(s). In compiling this table, principal note is taken of the oxidation state(s) found in crystalline oxides and halides.

[d]Values are taken from R. D. Shannon and C. T. Prewitt, *Acta Crystallogr*, **B25**, 925 (1969); **B26**, 1046 (1970). Note that bond distances generally increase with increase in cation coordination number and decrease with increase in cation oxidation state. Bond distances to fluorine are usually 0.05–0.10 Å shorter than distances to oxygen; bond distances to chlorine are usually 0.20–0.40 Å larger.

[e]Data taken mainly from *International Tables for X-ray Crystallography*, International Union of Crystallography, Chester, Vol. III, p. 278. *fcc* = face centred cubic (cubic close packed); hex = hexagonal; *hcp* = hexagonal close packed; $R\bar{3}m$, $R3m$ = rhombohedral; *bcc* = body centred cubic; tet = tetragonal; orth = orthorhombic; cub = cubic; monocl = monoclinic; diam = diamond structure; RT = room temperature.

Questions

Questions

1.1 What symmetry elements do the following tetrahedral-shaped molecules possess: (a) CH_3Cl, (b) CH_2C1_2, (c) CH_2ClBr, (d) CH_4?

1.2 What symmetry element do the following everyday objects have in common: (a) a teapot, (b) a pair of trousers, (c) a tricycle?

1.3 What symmetry elements are present in a cube (use a kiddies' building block for a model)?

1.4 If one pair of opposite cube faces in 1.3 is of different colour to the rest, what symmetries are present?

1.5 What symmetries does this book have (assume all pages and covers are blank)?

1.6 The hypothetical octahedral molecule MX_2Y_4 has two geometrical isomers. One has a centre of symmetry, the other does not. Draw the two isomers and show which one has a centre of symmetry. What other symmetries do the molecules possess?

1.7 How many lattice points are there in (a) a primitive lattice, (b) a body centred lattice, (c) a C-centred lattice, (d) a face centred lattice?

1.8 What is the Bravais lattice of the calcium carbide structure, Fig. 1.10a?

1.9 In the table of crystal systems, Table 1.1, only certain unit cell–lattice type combinations (i.e. Bravais lattices) are possible. Suggest reasons why the following are not included: (a) C-centred cubic, (b) F-centred tetragonal and (c) C-centred tetragonal

1.10 What are the indices for (a) planes that pass through opposite *ab* faces of a cubic unit cell, (b) planes parallel to *b* and *c* which cut *a* at 0, 1/2, 1, 3/2, etc., (c) planes perpendicular to a unit cell body diagonal, that cut *a*, *b* and *c* at 1/2 1/2 1/2, 111, 3/2 3/2 3/2, etc?

1.11 What are the indices for crystallographic directions that are parallel to the unit cell *b* edge (a) in the positive *y* direction and (b) in the negative *y* direction?

1.12 Using the *d*-spacing formula for a cubic substance, equation (1.2), what are the *hkl* values that correspond to the five largest *d*-spacings? Calculate these *d* values if $a = 5.00$ Å.

1.13 KCl has the rock salt structure with $a = 6.2931$ Å. Calculate its density.

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent

1.14 TlBr has a cubic unit cell, $a = 3.97$ Å, $D = 7.458$ g cm^{-3}. How many formula units, Z, are in the unit cell? Suggest a possible structure for TlBr.

1.15 In an *fcc*/*ccp* metal structure, (a) what are the indices of the *cp* layers of metal atoms, (b) what are the indices of the *cp* directions of metal atoms (i.e. in which directions are atoms in contact)?; (c) evaluate the atom radius, r, in terms of the unit cell edge, a.

1.16 Metallic Au and Pt have *fcc* unit cells, $a = 4.08$ and 3.91 Å. Calculate the metal atom radii.

1.17 In a *bcc* metal structure, (a) what is the atom coordination number, (b) is the structure close packed, (c) what are the indices of the *cp* directions?; (d) evaluate the atom radius in terms of the unit cell edge, a.

1.18 At 20 °C, α-Fe is *bcc*, $a = 2.866$ Å. At 950 °C, γ-Fe is *fcc*, $a = 3.656$ Å. At 1425 °C, δ-Fe is again *bcc*, $a = 2.940$ Å. At each temperature, calculate (a) the density and (b) the atomic radius of Fe.

1.19 Show that the packing density of a *bcc* arrangement of spheres is 0.6802.

1.20 Similarly, show that the packing density in *fcc* is 0.740.

1.21 Starting with a *ccp* array of anions, what structure types are generated by (a) filling all tetrahedral sites with cations, (b) filling half of the tetrahedral sites, e.g. T_+, with cations, (c) filling all octahedral sites with cations and (d) filling alternate layers of octahedral sites with cations?

1.22 Repeat the above question but with an *hcp* array of anions. Comment on the absence of any known structure type in one of these four categories.

1.23 Identify the following cubic structure types from the information on atomic coordinates:

(i) MX: M: 1/2 00, 0 1/2 0, 00 1/2, 1/2 1/2 1/2;
X: 000, 1/2 1/2 0, 1/2 0 1/2, 0 1/2 1/2

(ii) MX: M: 1/4 1/4 1/4, 3/4 3/4 1/4, 1/4 3/4 3/4, 3/4 1/4 3/4;
X: 000, 1/2 1/2 0, 1/2 0 1/2, 0 1/2 1/2

(iii) MX: M: 1/2 1/2 1/2;
X: 000

(iv) MX_2: M: 000, 1/2 1/2 0, 1/2 0 1/2, 0 1/2 1/2;
X: 1/4 1/4 1/4, 1/4 1/4 3/4, 1/4 3/4 1/4, 3/4 1/4 1/4, 1/4 3/4 3/4, 3/4 1/4 3/4, 3/4 3/4 1/4, 3/4 3/4 3/4

(v) MX_3: M: 000;
X: 1/2 00, 0 1/2 0, 00 1/2

(vi) AMX_3: A: 1/2 1/2 1/2;
M: 000;
X: 1/2 00, 0 1/2 0, 00 1/2

1.24 Starting from a rock salt structure, what structures are generated by the following imaginary steps:

(i) removal of all atoms or ions of one type;
(ii) removal of alternate layers of cations;
(iii) replacement of all cations in O sites by an equal number of cations in one set of T sites?

1.25 Explain why the NiAs structure is commonly found with metallic compounds but not with ionic compounds.

1.26 MgO has the rock salt structure, $a = 4.213$ Å. Calculate the Mg–O bond length. Assuming the oxide ion radius is 1.26 Å, what is the radius of Mg^{2+}? Are the oxide ions in contact?

1.27 Zinc blende, ZnS, has $a = 5.406$ Å. Calculate the Zn–S bond length.

1.28 Li_2O has the antifluorite structure, $a = 4.611$ Å. Calculate the (a) Li–O distance, (b) O–O distance, (c) Li–Li distance. (d) Are the oxide ions in contact ($r_0 = 1.26$ Å)?

1.29 $SrTiO_3$ has the cubic perovskite structure, $a = 3.91$ Å, with atomic coordinates:
Sr: 1/2 1/2 1/2, Ti: 0 0 0, O: 1/2 0 0, 0 1/2 0, 0 0 1/2.
(a) Draw a projection of the structure on the x–z plane.
(b) What is the coordination environment of (i) Sr, (ii) Ti, (iii) O.
(c) Calculate the Sr–O and Ti–O bond lengths. Calculate the density of $SrTiO_3$.
(d) Is the structure close packed? If so, describe it. What is the lattice type?
(e) What compositional modifications may be made to $SrTiO_3$ in an attempt to induce (i) ferroelectricity, (ii) superconductivity, (iii) ionic conductivity.

1.30 Si has the diamond structure, $a = 5.4307$ Å. Calculate the radius of Si.

1.31 Silver oxide, Ag_2O, has a cubic unit cell, $Z = 2$, $a = 4.726$ Å, with atomic coordinates:
Ag: 1/4 1/4 1/4, 3/4 3/4 1/4, 3/4 1/4 3/4, 1/4 3/4 3/4;
O: 000, 1/2 1/2 1/2.
What are the atomic coordinates if the unit cell is displaced so that an Ag atom is at the origin? Draw a projection of the structure onto the a–b face, using the new set of atomic coordinates. What is the lattice type? What are the coordination numbers of Ag and O? Calculate the Ag–O distance? Does the structure possess a centre of symmetry?

1.32 In an ideal rock salt structure, MX, in which the anions are in contact, calculate the *radius ratio*, r_M/r_x.

1.33 Repeat question 1.32 for an ideal zinc blende structure.

1.34 Repeat question 1.32 for an ideal CsCl structure.

1.35 By consideration of the *radius ratio rules*, derived in questions 1.32–1.34, and assuming that the calculated radius ratios represented the lowest values possible for a given structure, what structure type would you expect for an MX compound with:
(a) $r_M/r_x = 0.3$, (b) $r_M/r_x = 0.6$, (c) $r_M/r_x = 0.8$?

1.36 Compare the packing density of the NaCl and CsCl structures for which, in both cases, anion–anion and anion–cation direct contacts occur.

1.37 What kind of complex anion do you expect in the following:
(a) Ca_2SiO_4; (b) $NaAlSiO_4$ (tetrahedral Al); (c) $BaTiSi_3O_9$; (d) melilite, $Ca_2MgSi_2O_7$; (e) diopside, $CaMgSi_2O_6$; (f) tremolite, $Ca_2Mg_5Si_8O_{22}(OH, F)_2$; in this amphibole, the OH, F are not bonded to Si; (*g*) margarite, $CaAl_2(OH)_2(Si_2Al_2)O_{10}$; in this mica, two A1 are tetrahedral and two are octahedral; (*h*) kaolinite, $Al_2(OH)_4Si_2O_5$; Al is octahedral and OH is not bonded to Si?

1.38 Fullerene, C_{60}, forms compounds in which inert gas atoms enter interstitial sites in the *fcc* array of C_{60} molecules; $a = 14.17$ Å. Calculate (a) the van der Waals radius of the C_{60} molecule and (b) the radii of octahedral and tetrahedral sites. (Assume that the C_{60} molecules are spherical and are 'touching' in the close packed [110] direction.). Inert gas atoms have van der Waals radii Ar 1.91 Å, Kr 1.98 Å and Xe 2.05 Å. Comment on the probable formulae and structures of inert gas fullerides and on the relative ease of formation of such compounds for Ar, Kr and Xe.

1.39 The perovskite structure is found in a wide range of materials, both stoichiometric and non-stoichiometric. Depending on composition, a variety of electrical, magnetic and optical properties are exhibited by perovskites. Describe (a) the perovskite structure, (b) typical compounds and their formulae that have this structure, (c) how non-stoichiometry can be accommodated, (d) three properties of your choice exhibited by perovskites, focusing on the links between stoichiometry, structure and properties.

2.1 Explain why crystalline solids are generally more defective as a result of increasing temperature.

2.2 Calculate the enthalpy of formation of Frenkel defects in AgCl using the data given in Fig. 2.4.

2.3 What kind of defects would you expect to predominate in crystals of the following: (a) NaCl doped with $MnCl_2$; (b) ZrO_2 doped with Y_2O_3; (c) CaF_2 doped with YF_3; (d) Si doped with As; (e) a piece of aluminium that has been hammered into a thin sheet; (f) WO_3 after heating in a reducing atmosphere?

2.4 Explain why copper is a much softer metal than tungsten.

2.5 What effect would you expect the ordering of Cu, Zn atoms in β'-brass, Fig. 2.11, to have on the X-ray powder pattern, i.e. how do you expect the X-ray powder patterns of β- and β'-brass to differ?

2.6 The law of mass action can be used to analyse defect equilibria in systems of low defect concentrations. What difficulties are likely to be encountered if this method is applied to systems with larger defect concentrations?

2.7 What kind of dislocations are characterised by the following: (a) the Burger's vector is parallel to the direction of shear and perpendicular to the line of the dislocation; (b) the Burger's vector is perpendicular to the direction of shear and parallel to the line of the dislocation?

2.8 In which directions would you expect slip to occur most readily in (a) Zn, (b) Cu, (c) α-Fe, (d) NaCl?

2.9 Assuming that the enthalpy of creation of Schottky defects in NaCl is 2.3 eV and that the ratio of vacancies to occupied sites at 750 °C is 10^{-5}, estimate the equilibrium concentration of Schottky defects in NaCl at (a) 300 °C and (b) 25 °C.

2.10 For the solid solution of YF_3 in CaF_2, calculate the density as a function of composition for (a) a cation vacancy model and (b) an interstitial F^- model. Unit cell data for CaF_2 are given in Chapter 1. Assume that the unit cell volume is independent of solid solution composition.

3.1 The mineral grossular, $Ca_3Al_2Si_3O_{12}$, has the garnet structure with eight-coordinate Ca, octahedral Al and tetrahedral Si. Each oxygen is coordinated to one Si, one Al and two Ca. Show that the structure obeys Pauling's electrostatic valency rule.

3.2 Explain, using Pauling's electrostatic valency rule, why silicate structures never contain more than two SiO_4 tetrahedra sharing a common corner.

3.3 BeF_2 has the same structure as SiO_2, MgF_2 is the same as rutile and CaF_2 has the fluorite structure. Does this seem reasonable from size considerations?

3.4 In Table 1.8 are given unit cell constants for some oxides MO with the rock salt structure. Assuming that either (i) $r_{O^{2-}} = 1.26$ Å or (ii) $r_{O^{2-}} = 1.40$ Å, calculate for each (a) two values of the cation radius, $r_{M^{2+}}$, (b) two values for the radius ratio, $r_{M^{2+}}/r_{O^{2-}}$. Assess the usefulness of the radius ratio rules in predicting octahedral coordination for M^{2+} for these oxides. Repeat the calculations for two oxides with the wurtzite structure, Table 1.12, and for some oxides with the fluorite and antifluorite structures, Table 1.10.

3.5 Calculate lattice energy values for the alkaline earth oxides using Kapustinskii's equation and the data given in Table 1.8. Compare your results with those given in Table 3.6. Estimate the enthalpies of formation of these oxides.

3.6 Account for the observation that whereas CuF_2 and CuI are stable compounds, CuF and CuI_2 are not stable.

3.7 The most stable oxide of lithium is Li_2O but for rubidium and caesium, the peroxides M_2O_2 and superoxides MO_2 are more stable than the simple oxides M_2O. Comment.

3.8 Using Sanderson's methods, estimate the partial ionic character of the sodium halides. Hence calculate the radii of the atoms involved and the unit cell a value (all have the rock salt structure). Compare your answers with the data given in Table 1.8.

3.9 The oxides MnO, FeO, CoO and NiO all have the cubic rock salt structure with octahedral coordination of the cations. The structure of CuO is different and contains grossly distorted CuO_6 octahedra. Explain.

3.10 Although the *d* electrons in many transition metal compounds may not be involved directly in bond formation, they nevertheless exert a considerable influence on structure. Explain.

3.11 Show schematically the band structure of elemental magnesium and account for its metallic conductivity.

3.12 TiO and NiO both have the rock salt structure but whereas TiO is metallic, pure NiO is an electrical insulator. Explain.

3.13 Pure WO_3 (ReO_3 structure) is an electrical insulator whereas tungsten bronzes such as $Na_{0.5}WO_3$ are metallic. Sketch the probable band structure of these and indicate which of the *d* orbitals on W are responsible for the electrical properties.

3.14 The rock salt structure is shown by substances as diverse as LiF, NiO, TiO, LaN and TiC. Discuss the type(s) of bonding likely to be present in these substances.

3.15 Using the Kapustinskii equation and the Born–Haber cycle with the following data for MgO (in kJ mol^{-1}), estimate the electron affinity for oxygen (O/O^{2-}): $S_{Mg} = 148$; $IP_{Mg/Mg^{2+}} = 2188$; $D_{O_2} = 498$; $\Delta H_{f\,MgO} = -602$; $r_{Mg^{2+}} = 0.86$ Å; $r_{O^{2-}} = 1.26$ Å.

4.1 Discuss the use of 'chimie douce' methods for the synthesis of inorganic solids. What are the advantages and disadvantages compared with traditional 'shake 'n bake' methods?

4.2 How could 'chimie douce' methods be used to prepare the following: $LiTiS_2$, $Li_{0.5}CoO_2$, $BaTiO_3$, $CoFe_2O_4$?

4.3 Solid state reactions of inorganic materials generally require long periods of heating at high temperatures. Explain the reasons for this and the steps that may be taken to speed up reactions.

4.4 Synthesis of crystalline materials is greatly facilitated if the reaction mixture is homogeneous on an atomic scale.

(a) Describe some (at least three) of the routes that could be used to synthesise a sample of $BaTiO_3$ powder and compare the advantages/disadvantages of each.

(b) Describe the strategies that can be adopted to synthesise thin films of semiconductor materials such as CdS or GaAs.

4.5 Synthesis of diamonds traditionally requires very high temperatures and pressures, yet it is now possible to synthesise diamonds from the gas phase at sub-ambient pressures. Account for these observations and discuss the thermodynamic/kinetic factors involved in diamond synthesis.

4.6 Compare the response of crystalline and amorphous silicon to doping with (a) a trivalent and (b) a pentavalent element.

4.7 Explain why, and how, TiS_2 could, in principle, be used as a cathode in lithium cells. Could NiS be used in a similar way?

4.8 What methods could be used to prepare a film of $BaTiO_3$ on a glass substrate?

4.9 What is the role of the template in zeolite synthesis?

5.1 Using the $K\alpha_1$ data in Table 5.2, verify graphically Moseley's law. What wavelength do you expect for Co $K\alpha_1$ radiation?

5.2 What is the probable lattice type of crystalline substances that give the following observed reflections?
(a) 110, 200, 103, 202, 211
(b) 111, 200, 113, 220, 222
(c) 100, 110, 111, 200, 210
(d) 001, 110, 200, 111, 201.

5.3 Calculate the d and 2θ values for the 111 and 200 lines in the X-ray powder pattern, Cu $K\alpha$ radiation, of a cubic substance with $a = 5.0$ Å.

5.4 The value of n in Bragg's law is always set equal to 1. What happens to the higher order diffraction peaks?

5.5 A cubic alkali metal halide has its first six lines with d-spacing 4.08, 3.53, 2.50, 2.13, 2.04 and 1.77 Å. Assign Miller indices to the lines and calculate the value of the unit cell dimension. The alkali metal halide has density 3.126 g cm^{-3}. Identify the alkali metal halide.

5.6 An imaginary orthorhombic crystal has two atoms of the same kind per unit cell located at 000 and $^1/_2{}^1/_2 0$. Derive a simplified structure factor equation for this. Hence, show that, for a C-centred lattice, the condition for reflection is hkl: $h + k = 2n$.

5.7 Derive a simplified structure factor master equation for the perovskite structure of $SrTiO_3$ with the origin at a centre of symmetry. Atomic coordinates are Sr $^1/_2{}^1/_2{}^1/_2$, Ti 000 and O $^1/_2 00$, $0{}^1/_2 0$, $00{}^1/_2$.

5.8 The 111 reflection in the powder pattern of KCl has zero intensity but in the powder pattern of KF it is fairly strong. They have the same crystal structure. Explain.

5.9 The gold–copper alloy has a face centred cubic unit cell at high temperatures in which the Au and Cu atoms are distributed at random over the available corner and face centre sites. At lower temperatures, ordering occurs: Cu atoms are located preferentially on the corner sites and one pair of face centre sites; Au atoms are located on the other two pairs of face centre sites. What effects would you expect this ordering process to have on the X-ray powder pattern?

5.10 The X-ray powder pattern of orthorhombic Li_2PdO_2 includes the following lines: 4.68 Å (002), 3.47 Å (101), 2.084 Å (112). Calculate the values of the unit cell parameters. The density is 4.87 g cm^{-3}; what are the cell contents?

5.11 An ammonium halide, NH_4X, has the CsCl structure at room temperature, $a = 4.059$ Å, and transforms to the NaCl structure at 138 °C, $a = 6.867$ Å.
(a) The density of the room temperature polymorph is 2.431 g cm^{-3}. Identify the substance.
(b) Calculate the d-spacings of the first four lines in the powder pattern of each polymorph.
(c) Calculate the percentage difference in molar volume between the two polymorphs, ignoring thermal expansion effects.
(d) Assuming an effective radius of 1.50 Å for the spherical NH_4^+ ion and that anions and cations are in contact, calculate the radius of the anion in each structure. Are the anions in contact in the two structures?

5.12 'Each crystalline solid gives a characteristic X-ray powder diffraction pattern which may be used as a fingerprint for its identification.' Discuss the reasons for the validity of this statement and indicate why two solids with similar structures, e.g. NaCl and NaF, may be distinguished by their powder patterns.

5.13 What qualitative differences would you expect between powder X-ray and neutron diffraction patterns, obtained with monochromatic radiation of similar wavelength, of (a) NaCl, (b) metallic Fe and (c) antiferromagnetic NiO?

5.14 Show by means of qualitative sketches the essential differences between the X-ray powder diffraction patterns of (a) a 1:1 mechanical mixture of powders of NaCl and AgCl and (b) a sample of (a) that has been heated to produce a homogeneous solid solution.

5.15 A sample of aluminium hydroxide was shown by chemical analysis to contain a few percent of Fe^{3+} ions as impurity. What effect, if any, would the Fe^{3+} ions have on the powder pattern if it was present (a) as a separate iron hydroxide phase and (b) substituting for Al^{3+} in the crystal structure of $Al(OH)_3$.

6.1 Under what conditions may (i) optical microscopy and (ii) powder X-ray diffraction, be used to determine the purity of a solid material?

6.2 Compare the methods that may be used to identify and characterise crystalline samples of (a) iodoform, CHI_3, and (b) CdI_2.

6.3 How would lightly crushed samples of the following differ when viewed in a polarising microscope: (a) SiO_2 glass, (b) quartz crystal, (c) NaCl, (d) Cu metal?

6.4 Calculate the wavelength of the electrons accelerated through 50 kV in an electron microscope.

6.5 Explain the following abbreviations: EXAFS, XANES, EELS, XRF, ESCA, XPS, UPS, MAS NMR, SEM, TEM, AEM, EDX, EPMA.

6.6 What technique(s) might you use to determine the following: (a) the oxidation state and coordination number of iron in brown/green bottle glass; (b) the coordination number of Cr^{3+} in a crystal of ruby; (c) the nature of a surface layer on a piece of aluminium metal; (d) the nature of some microscopic, crystalline inclusions in a polycrystalline, silicon carbide ceramic piece; (e) the Ca:Si ratio in particles of hydrated cement; (f) the location of the hydrogen atoms in a powdered sample of palladium hydride; (g) whether or not MnO possesses a magnetically ordered superstructure.

6.7 What are the approximate energy values, in units of (a) eV, (b) kJ mol^{-1}, (c) cm^{-1}, (d) Hz, for (i) changes of electron spin orientation, (ii) *d–d* transitions in NiO, (iii) the K absorption edge in Cu metal, (iv) lattice vibrations in KCl?

6.8 Which microscopic technique(s) would you use to study (a) the texture of a piece of metal, (b) the homogeneity of a sample of powdered glass, (c) defects such as dislocations, stacking faults and twinning, (d) the possible contamination of a sample of salt by washing soda?

6.9 Explain why it is generally much easier to characterise organic materials than non-molecular inorganic materials.

6.10 Suppose that you are trying to synthesise a new perovskite and that you end up with a white solid. How would you determine if it was a new perovskite, its structure, composition and purity?

6.11 What kind of DTA and TG traces would you expect to obtain on heating samples of the following until they became liquid: (a) beach sand, (b) window glass, (c) salt, (d) washing soda, (e) Epsom salts, (f) metallic Ni, (g) ferroelectric $BaTiO_3$, (h) a clay mineral?

6.12 Which of the following would you expect to give a reversible DTA effect, with or without hysteresis: (a) melting of salt, (b) decomposition of $CaCO_3$, (c) melting of beach sand, (d) oxidation of metallic Mg, (e) decomposition of $Ca(OH)_2$?

6.13 Salt is often added to icy roads in winter. Could DTA be used to quantify the effects of salt on ice? What results would you expect?

6.14 Sketch the kind of DTA trace to be expected on heating (a) a sample of pure Fe and (b) an Fe–C alloy containing 1 wt% C.

6.15 Spectroscopic techniques utilise a very wide range of the electromagnetic spectrum from radiofrequency to microwave, infrared, visible, ultraviolet, X-rays and γ-rays. Outline the principal techniques that operate in each of these regions and give examples of their possible use in solid state chemistry for (a) elemental analysis, (b) phase analysis and (c) structure determination.

7.1 What are (i) the mole percent and (ii) the weight percent of (a) Al_2O_3 in mullite, $Al_6Si_2O_{13}$; (b) Na_2O in devitrite, $Na_2Ca_3Si_6O_{16}$; (c) Y_2O_3 in yttrium iron garnet, $Y_3Fe_5O_{12}$?

7.2 Phase diagrams are drawn with the compositions represented usually as either weight percent or mole percent. Derive expressions for converting from one to the other for a binary AB system.

7.3 The system $CaO–SiO_2$ contains the following phases: Ca_3SiO_5, Ca_2SiO_4, $Ca_3Si_2O_7$, $CaSiO_3$. For each, determine the composition, as mole percent SiO_2, and hence construct the composition axis, in mole percent SiO_2, for the $CaO–SiO_2$ phase diagram.

7.4 Repeat question 7.3 for the $Na_2O–P_2O_5$ phase diagram which contains the phases Na_3PO_4, $Na_4P_2O_7$, $Na_5P_3O_{10}$ and $NaPO_3$.

7.5 Sketch a phase diagram for a system A–B that has the following features. Three binary compounds are present A_2B, AB and AB_2. Both A_2B and AB_2 melt congruently. AB melts incongruently to give A_2B and liquid. AB also has a lower limit of stability.

7.6 Sketch the phase diagram for the system $Al_2O_3–SiO_2$ using the following information. Al_2O_3 and SiO_2 melt at 2060 and 1720 °C. One congruently melting compound, $Al_6Si_2O_{13}$, forms between Al_2O_3 and SiO_2 with a melting point of 1850 °C. Eutectics occur at ~5 mole % Al_2O_3, 1595 °C, and ~67 mole % Al_2O_3, 1840 °C. Label all regions.

7.7 Sketch, using graph paper, the phase diagram for $Na_2O–Nb_2O_5$ using the following information:
There are four binary phases with melting points as follows:
Na_3NbO_4 congruent, 992 °C
$NaNbO_3$ congruent, 1412 °C
$Na_2Nb_8O_{21}$ incongruent, 1265 °C (to $Na_2Nb_{20}O_{51}$ + liquid)
$Na_2Nb_{20}O_{51}$ incongruent, 1290 °C (to Nb_2O_5 + liquid)
Na_2O melts at ~1200 °C
Nb_2O_5 melts at 1485 °C
Eutectics at: 10% Nb_2O_5, 830 °C;
31% Nb_2O_5, 975 °C;
68% Nb_2O_5, 1220 °C
$NaNbO_3$ undergoes polymorphic transitions at 368 and 640 °C.
Label all regions on the diagram.

7.8 Pure iron undergoes the $\alpha \rightarrow \gamma$ transformation at 910 °C. The effect of added carbon is to reduce the transformation temperature from 910 to 723 °C. Sketch the general appearance of the Fe-rich end of the Fe–C phase diagram using this information.

7.9 The system A–B is a simple eutectic system in which the two end-member phases form limited ranges of solid solution. Sketch a probable phase diagram for this system. Compare your phase diagram with that for the $Mg_2SiO_4–Zn_2SiO_4$ system, Fig. 7.14. How would you determine experimentally: (a) the compositions of the solid solution limits; (b) the mechanism of solid solution formation in each case; (c) the eutectic temperature?

7.10 Explain, with the aid of examples, the difference between a phase and a component. Under what conditions can a component be a phase?

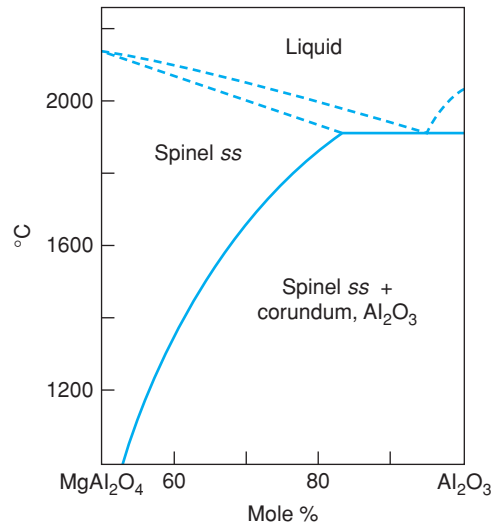


Figure Q.1 ***Phase diagram for the $MgAl_2O_4$–Al_2O_3 system.***

7.11 For the $MgAl_2O_4$–Al_2O_3 phase diagram (Fig. Q.1), describe the reactions that would be expected to occur, under equilibrium conditions, on cooling a liquid of composition 40 mole % MgO, 60 mole % Al_2O_3. Using rapid cooling rates, how might the product(s) differ?

8.1 Explain why polyacetylene is an electrical conductor whereas polyethylene is not.

8.2 What kind of materials may be added to polypyrrole to make it (a) *n*-type, (b) *p*-type?

8.3 Using the atomic coordinates given in Fig. 8.8(a), draw the unit cell contents of $YBa_2Cu_3O_7$ as a projection on (a) the *ac* plane and (b) the *bc* plane. Calculate the Cu–O, Ba–O and Y–O distances using the cell dimensions given in Fig. 8.8.

8.4 Sketch the energy band structure of silicon with a band gap of 1.1 eV. What elements would you add to silicon to make it *p*-type? Sketch the resulting energy level structure for acceptor levels that are located 0.01 eV above the top of the valence band. What fraction of the acceptor levels would be occupied at room temperature? Assume that the probability of excitation is proportional to $\exp(-E/kT)$. If the impurity concentration is 10^{-4} at.%, what is the carrier density due to the impurities? What would be the intrinsic carrier concentration at room temperature, in the absence of impurities? At what temperature is the intrinsic density equal to that caused by the impurities?

8.5 Repeat the above problem but for germanium with a band gap of 0.7 eV.

8.6 For use in practical semiconductor devices, why is it desirable to have materials with a large band gap between valence and conduction band but a small gap between valence/conduction band and impurity levels?

8.7 (a) A piece of Si is doped with 5×10^{18} B atoms cm^{-3}. What is the concentration of (i) electrons and (ii) holes? Assume the intrinsic defect concentration is 1×10^{10} cm^{-3}. (b) Calculate the conductivity of the sample. Assume the hole mobility, μ_h, is 0.05 m^2 V^{-1} s^{-1}.

8.8 (a) What effect, if any, do you expect small amounts of the following impurities to have on the conductivity of NaCl crystals: (i) KCl; (ii) NaBr; (iii) $CaCl_2$; (iv) AgCl; (v) Na_2O? (b)What effect,

if any, do you expect small amounts of the following impurities to have on the conductivity of AgCl crystals: (i) AgBr; (ii) $ZnCl_2$; (iii) Ag_2O?

8.9 Using the conductivity data for NaCl given in Fig. 8.18, estimate (a) the enthalpy of cation vacancy migration, (b) the enthalpy of Schottky defect formation. Your results should be similar to those given in Table 8.5.

8.10 A particular solid has a conductivity of $10^{-5}\ \Omega^{-1}\ cm^{-1}$ at room temperature. Suggest ways for determining the charge carrier responsible for conduction. How could you distinguish conduction by (a) electrons, (b) Na^+ ions, (c) O^{2-} ions?

8.11 Using the conductivity data for different ion exchanged β-aluminas given in Fig. 8.26, calculate the pre-exponential factor A and the activation energy E for each.

8.12 With the aid of Fig. 7.24, explain how the Na/S cell would differ in its mode of operation or usefulness if operated at (a) 200 °C, (b) 300 °C.

8.13 The Na/S cell is being developed for possible applications in electrically powered buses and cars. For such applications, what advantages and disadvantages do you think the Na/S cell would have in comparison with conventional Pb/acid batteries?

8.14 Suggest a design for a device containing a solid electrolyte that would be sensitive to fluorine gas.

8.15 What are the differences between paraelectric (i.e. dielectric), ferroelectric, ferrielectric and antiferroelectric materials?

8.16 If the following substances were placed between the plates of a capacitor, approximately what values would you expect for their apparent dielectric constant measured at 1 kHz: (a) Ar gas; (b) water; (c) ice; (d) pure single-crystal silicon; (e) pure single-crystal KBr; (f) single crystal of KBr doped with $CaBr_2$; (g) Na β-alumina; (h) $BaTiO_3$?

8.17 Which, if any, of the following crystals might you expect to exhibit piezoelectricity: (a) NaCl; (b) CaF_2; (c) CsCl; (d) ZnS, wurtzite; (e) NiAs; (f) TiO_2, rutile?

8.18 Comment on the validity of the following statement: 'Pyroelectric substances are those that develop a net spontaneous polarisation on heating.'

8.19 What are dielectric losses, what are they caused by and how might they be minimised in materials that are to be used as electrical insulators?

8.20 Describe, with the use of examples, the structural characteristics that are a prerequisite for high ionic conductivity in solid electrolytes.

8.21 Discuss the types of defects that may arise on doping (a) LiF with MgF_2, (b) NiO with Li_2O. What would be the consequences of such doping on their electrical properties?

8.22 Account for the following observations:
(a) LiF is an electronic insulator;
(b) pure NiO is a pale green insulator but, on heating in air, it turns black and becomes a semiconductor;
(c) TiO exhibits metallic conduction whereas MnO is a semiconductor and pure NiO is an insulator.
All have the same rock salt crystal structure.

8.23 The main defect in stoichiometric CsI is the Schottky defect. The fraction of lattice sites that exhibit Schottky defects, as a function of temperature, is as follows: 1.08×10^{-16} at 300 K, 1.06×10^{-12} at 400 K, 2.63×10^{-10} at 500 K, 1.04×10^{-8} at 600 K, 4.76×10^{-6} at 900 K.

Show whether the formation of Schottky defects is given by an Arrhenius (or Boltzmann) expression and, if so, evaluate the enthalpy of formation of (a) one Schottky defect, (b) 1 mol of Schottky defects.

8.24 Describe briefly the formulae and structures of the high-T_c cuprate superconductors. Indicate their relation to the perovskite structure and show how the coordination environment of Cu may be taken as an indication of its oxidation state.

8.25 Compare critically the methods that may be used to determine the oxygen content, δ, in a material such as $LaBa_2Cu_3O_\delta$.

8.26 The emf of the solid electrolyte cell

Pt: Ni, Cu, NiO/YSZ/Ni, NiO: Pt

is +0.2 V at 1500 K. If the partial pressure over the mixture of Ni, NiO is 10^{-6} atm, calculate the partial pressure of oxygen over the Ni, Cu, NiO mixture.

8.27 Explain why oxygen sensors based on YSZ cannot be used in very low ($<10^{-20}$ atm) or very high ($>10^{10}$ atm) oxygen partial pressure environments.

9.1 Indicate how you would distinguish between paramagnetic, ferromagnetic and antiferromagnetic behaviour using a Gouy balance.

9.2 Vanadium monoxide is paramagnetic and a good conductor of electricity whereas nickel oxide is paramagnetic/antiferromagnetic and a poor conductor of electricity. Account for these observations.

9.3 Account for the magnetic behaviour of the following spinels: (a) $ZnFe_2O_4$ is antiferromagnetic; (b) $MgFe_2O_4$ is ferrimagnetic and its magnetic moment increases with increase in temperature; (c) $MnFe_2O_4$ is ferrimagnetic and its magnetic moment is independent of temperature.

9.4 Explain why magnetic materials that are to be used in information storage should have either square- or rectangular-shaped hysteresis loops.

9.5 Why cannot pure iron metal be used in transformer cores?

9.6 Show that the following magnetic susceptibility data for Ni fit the Curie–Weiss law:

T (K)	800	900	1000	1100	1200
$\chi \times 10^{-5}$	3.3	2.1	1.55	1.2	1.0

Evaluate T_c or θ and C.

9.7 Explain why the substitution of Zn for Ni in $Ni_{1-x}Zn_xFe_2O_4$ spinel causes the saturation magnetisation to increase from 2.3 BM at $x = 0$ to 5.2 BM at $x = 0.5$.

10.1 The potassium halides are all transparent to visible light. Calculate the wavelengths at which they become opaque. Band gap data are given in Table 3.19.

10.2 What are the general requirements for a solid material that is to be used as a laser source?

10.3 Anti-Stokes phosphors emit light of shorter wavelength than that used for excitation. Explain why the law of conservation of energy is not violated.

A set of worked answers has been prepared and can be found on the Companion Website, http://www.wiley.com/go/west/solidstatechemistrystudent.

Index

Note: page references in **bold** refer to main sections.

Solid State Chemistry and its Applications, Second Edition, Student Edition. Anthony R. West.

Companion Website: http://www.wiley.com/go/west/solidstatechemistrystudent