Physics of matter

Lecture notes by Paolo Fornasini

Department of Physics, University of Trento

Revised:

March 30, 2023

Contents

1 Introduction **1**
1.1 Macroscopic approach . . . 1
1.2 Microscopic approach . . . 5
1.3 Key concepts . . . 7
1.4 Main experimental techniques . . . 9
1.5 Bibliography of Chapter 1 . . . 10

2 Real-space structure **13**
2.1 Introduction . . . 13
2.2 Crystal structure: basics . . . 14
2.3 Crystal structures: examples . . . 18
2.4 Formal description of crystal structures . . . 21
2.5 Deviations from the infinite perfect crystal model . . . 24
2.6 Radial distributions . . . 29
2.7 Complements and problems . . . 33
2.8 Bibliography of Chapter 2 . . . 36

3 Reciprocal space **37**
3.1 Reciprocal space in one dimension . . . 37
3.2 Reciprocal space in three dimensions . . . 42
3.3 Relation between direct and reciprocal lattices . . . 43
3.4 Physical interpretation of the reciprocal space . . . 45
3.5 Complements and problems . . . 49
3.6 Bibliography of Chapter 3 . . . 52

4 Symmetry in molecules and crystals **53**
4.1 Introduction . . . 53
4.2 Group theory, basic elements . . . 56
4.3 Discrete point symmetries of molecules . . . 60
4.4 Point symmetry groups, examples . . . 63
4.5 Point symmetry groups, classification . . . 67
4.6 Symmetry in crystals . . . 70
4.7 Symmetry breaking . . . 74
4.8 Complements on group symbols . . . 76
4.9 Bibliography of Chapter 4 . . . 79

5 Representations of symmetry groups **81**
5.1 General properties of group representations . . . 81
5.2 Representation in the three-dimensional Euclidean space . . . 82
5.3 Irreducible representations: theory . . . 87
5.4 Irreducible representations: examples . . . 92
5.5 Group representations in the function space . . . 100
5.6 Group representations and Quantum Mechanics . . . 103

5.7 Molecular Orbitals and degeneracy removal . . . 108
5.8 Crystal field splitting . . . 112
5.9 Evaluation of integrals . . . 117
5.10 Matrix elements and selection rules . . . 121
5.11 Complements . . . 126
5.12 Bibliography of Chapter 5 . . . 127

6 The origin of structures 129
6.1 Born-Oppenheimer (adiabatic) approximation . . . 129
6.2 Atomic aggregates . . . 134
6.3 Van der Waals bonds and molecular crystals . . . 138
6.4 Ionic crystals . . . 144
6.5 Covalent crystals . . . 147
6.6 Metals . . . 151
6.7 Hydrogen-bonded solids . . . 151
6.8 Non-crystalline solids . . . 153
6.9 An introduction to Density Functional Theory (DFT) . . . 158
6.10 Bibliography of Chapter 6 . . . 161

7 Vibrational dynamics 163
7.1 Two-atomic system (one degree of freedom) . . . 163
7.2 Many-atomic systems . . . 171
7.3 Vibrational dynamics of crystals . . . 177
7.4 Energy of normal modes. Phonons . . . 182
7.5 First Brillouin Zone and dispersion relations . . . 187
7.6 Complements and demonstrations . . . 190
7.7 Bibliography of Chapter 7 . . . 196

8 Vibrational thermal properties 197
8.1 Vibrational heat capacity . . . 197
8.2 Thermal expansion . . . 202
8.3 Thermal conductivity . . . 209
8.4 Complements and demonstrations . . . 217
8.5 Bibliography of Chapter 8 . . . 219

9 Ground-state electronic properties of crystals 221
9.1 Non-interacting electron gas: ground state . . . 221
9.2 Independent electrons in a periodic potential . . . 224
9.3 Nearly-free electron model . . . 231
9.4 Filling-up of electron bands . . . 234
9.5 Bibliography of Chapter 9 . . . 237

10 Electron thermal and transport properties 239
10.1 Electronic contribution to the specific heat . . . 239
10.2 Electric conductivity in metals . . . 243
10.3 Complements and demonstrations . . . 245
10.4 Bibliography of Chapter 10 . . . 247

11 Structural probes 249
11.1 Introduction . . . 249
11.2 Scattering cross-section . . . 251
11.3 X-ray elastic scattering . . . 254
11.4 X-ray scattering from atoms . . . 261
11.5 X-rays, electrons and thermal neutrons . . . 266
11.6 Elastic scattering from atomic aggregates . . . 270

11.7 Temperature effects on diffraction peaks 279
11.8 Extended X-ray Absorption Fine Structure (EXAFS) 281
11.9 Inelastic scattering 287
11.10 Complements and demonstrations 291
11.11 Bibliography of Chapter 11 295

A Physical constants and conversion factors **297**

Chapter 1

Introduction

The name "Physics of Matter" is very comprehensive. It includes a number of different research fields, distinguished by the size of the studied objects and characterised by different experimental methods and theoretical approaches: from elementary particles to matter at the cosmological scale.

In these lecture notes, we focus our attention on *atomic aggregates of condensed matter*, mainly large aggregates where the number of atoms of the order of 10^{23}; in some cases we consider molecules and nanoclusters too.

It is convenient to distinguish two types of physical properties of condensed matter,

1. *equilibrium properties*, such as structure, specific heat, thermal expansion, magnetisation, electric polarisation,
2. *non equilibrium properties*, including transport properties such as thermal conductivity and electrical conductivity

and two different approaches to their treatment,

a) macroscopic approach
b) microscopic (atomistic) approach

Note: "Macroscopic" and "microscopic" are conventional terms: microscopic means here at the atomic or subatomic level. In other situations, the term microscopic refers more exactly to objects of μm (micrometer) size, and the term nanoscopic to objects of nm (nanometer) size.

1.1 Macroscopic approach

The macroscopic approach is based on the measurement of macroscopic quantities, without any attempt at detecting and measuring the properties at the microscopic (atomistic) scale.

1.1.1 Equilibrium properties

A powerful approach to the equilibrium properties of systems composed of a very large number of elementary entities (e.g. atoms) is represented by *Thermodynamics*.
The *equilibrium state* of a thermodynamic system is described in terms of a little number of macroscopic coordinates (volume, temperature, magnetisation, electrical polarisation, and so on). The *physical properties* of the system are described by a number of phenomenological relationships among the thermodynamic coordinates (e.g the equations of state). The Laws of Thermodynamics impose constraints between macroscopic quantities, establish evolution criteria for physical and chemical processes, allow one to find unifying criteria for phenomena of very different nature.

Example 1: The constant-volume heat capacity C_v, the constant-pressure heat capacity C_p, the isothermal compressibility χ_T and the coefficient of volume thermal expansion β are different

for different substances. Their values and their temperature dependencies cannot be determined from thermodynamic considerations, and have to be experimentally measured for each substance. As a consequence of the law of thermodynamics, one can however show that the four quantities are connected by the general relation $C_p - C_v = TV\beta^2/\chi_T$, valid for every substance.

Example 2: Phase transitions can be classified according to their thermodynamic behaviour (first-order transitions and continuous transitions). Phase transitions of completely different nature (critical transition of fluids, ferromagnetic transition, ferroelectric transition, superfluid transition and so on) can be treated within the unifying framework of the thermodynamic Landau theory.

Thermodynamic functions

The thermodynamic properties of a system are generally expressed by a suitable function of a set of suitable coordinates. Different choices of thermodynamic functions are possible, according to the nature of the problem.
The internal energy

$$U = U(S, V, \{n_i\}, \{X_j\}) , \tag{1.1}$$

is a function of a set of extensive coordinates: S is the entropy, V is the volume, $\{n_i\}$ are the molar quantities of the components, $\{X_j\}$ are additional coordinates describing for example magnetisation, electrical polarisation, etc. For a system with no additional $\{X_j\}$ coordinates,

$$U = U(S, V, \{n_i\}) , \tag{1.2}$$

whose differential form expresses the conservation of energy for a pure substance:

$$dU = T\,dS - p\,dV + \sum\nolimits_i \mu_i\,dn_i , \tag{1.3}$$

where $\mu_i = \partial U/\partial n_i$ are the chemical potentials.
For different choices of the independent coordinates, the information content is maintained by new thermodynamic functions, obtained from $U(S, V, n)$ through Legendre transforms:

- the enthalpy

$$H(S, p, \{n_i\}) = U + pV , \tag{1.4}$$

- the Helmholtz function (or Helmoltz free energy)

$$F(T, V, n) = U - TS , \tag{1.5}$$

- the Gibbs function (or Gibbs free energy)

$$G = G(T, p, n) = H - TS = U + pV - TS . \tag{1.6}$$

Phase diagrams

Temperature–pressure diagrams ($T - p$) are generally employed for pure substances.
The T-p plane is partitioned into regions, corresponding to homogeneous phases, each one characterised by a smooth dependence on T and p of the macroscopic properties: typically molar entropy and molar volume, but in some cases even viscosity, magnetic susceptibility, etc. The phase regions are separated by coexistence curves, where the macroscopic properties undergo discontinuous variations. Three coexistence curves meet at a triple point.
The phase diagrams can be obtained by macroscopic experimental techniques, such as differential scanning calorimetry (DSC).

The simplest $T - p$ diagrams are used to distinguish the three states of aggregation of matter: gas, liquid, solid (we neglect here the high-temperature ionized gas, or plasma state). Notice however that:

- The liquid and gas states can be considered as different modifications of one single phase, the fluid phase (the liquid-gas coexistence curve terminates at the critical point).
- A number of different solid crystalline phases are possible for the same substance at different T, p values (allotropy or polymorphism).
- The phase diagram of helium is characterised by two different liquid phases, the normal one and the superfluid one.

Example: As an example of allotropy, let us consider tin (Sn) at ambient pressure. Depending on temperature, two allotropic forms are possible, which can be distinguished in terms of macroscopic properties (e.g. electrical conductivity) as well as in terms of atomic structure:
- below $T_c \simeq 286\,\mathrm{K}$, phase α (grey tin); diamond structure; semiconductor;
- above $T_c \simeq 286\,\mathrm{K}$, phase β (white tin); body-centered tetragonal; metal.

The equilibrium states are characterised by extremum properties of a suitable thermodynamic function. For example, if T and P are held constant, the stable equilibrium phase of the system is characterised by the minimum value of the Gibbs free energy.

Equations of state

The equation of state of a pure substance is a function connecting the thermodynamic variables, $f(p, V, T) = 0$.
The equation of state of the ideal gas is well known. Different equations of state have been proposed for real gases, the simplest one being the Van der Waals equation.
At present, there is no single equation of state that can describe the properties of all substances under all conditions.
A number of equations of state have been proposed for solids, which in general try to describe the relation between volume and pressure, neglecting the generally weaker temperature effects.

Statistical Thermodynamics

The macroscopic approach of classical Thermodynamics is connected to the microscopic description by statistical thermodynamics. In statistical thermodynamics, the thermodynamical properties are expressed by the partition function

$$Z \;=\; \sum\nolimits_i g_i \exp(-E_i/k_B T) \tag{1.7}$$

where the sum is over all the possible energy levels of the entire system. In (1.7), E_i and g_i are the energy value and the degeneracy of the i-th level, respectively, and k_B is the Boltzmann constant. The most direct connection between macroscopic and statistical thermodynamics is represented by the relation between the Helmholtz F function and the partition function Z:

$$F \;=\; -k_B\, T\, \ln Z\,. \tag{1.8}$$

The relation between internal energy and partition function is

$$U \;=\; k_B\, T^2 \left(\frac{\partial \ln Z}{\partial T}\right)_V . \tag{1.9}$$

Finally, the relation between Gibbs function and partition function is

$$G \;=\; k_B\, TV \left(\frac{\partial \ln Z}{\partial V}\right)_T - k_B\, T\, \ln Z\,. \tag{1.10}$$

All thermodynamic properties can be derived from the partition function Z. To calculate Z it is necessary to know the energy levels E_i and their degeneracy.

Note: If the energy levels E_i of (1.7) include a contribution of potential energy due to external fields, the connection between Z and the thermodynamic functions is less simple. For example,

for a magnetic system the total energy includes the potential energy $-\mathcal{M}\mathcal{H}$, which doesn't contribute to the internal energy U. The total energy corresponds to the magnetic enthalpy $H^* = U - \mathcal{M}\mathcal{H}$ and the partition function is connected to the magnetic Gibbs function:

$$G^* \;=\; H^* - TS \;=\; -k_B\,T\,\ln Z$$

.

1.1.2 Non-equilibrium properties

Only for thermodynamic equilibrium states are all thermodynamic quantities defined. A large number of systems and phenomena cannot however be described by equilibrium thermodynamics.

Static non-equilibrium states

A number of systems are characterised by constant values of the macroscopic properties (and of the microscopic structural properties) even in the absence of thermodynamical equilibrium. One can distinguish

a) metastable equilibrium states, corresponding to a partial minimum of the suitable thermodynamic function, for example the Gibbs function for (T, p) systems;
b) non-equilibrium states, for which the suitable thermodynamic function is not at a minimum value (be it partial or total).

The existence of metastable states and of non-equilibrium states depends on the kinetics at the atomic level. Let us give two significant examples.

Example 1: Carbon (6–C) has two crystalline phases, graphite and diamond, which are thermodynamically stable in different regions of the (T, p) plane, and have very different physical properties (e.g. graphite is a good electrical conductor. diamond is an insulator). Below about 10 kbar, the stable phase is graphite. Diamond is the stable phase above about 10 kbar. At ambient pressure, the diamond phase is metastable. Diamond can be formed only at very high pressures. At ambient pressure and room temperature the kinetic of transformation from diamond to graphite (the stable phase) would require an exceedingly high amount of energy to break the bonds of the diamond structure.

Example 2: Silica SiO_2 can be found as a crystalline solid (characterised by different thermodynamically stable phases at different T, p values) or as a non-crystalline, glassy solid (thermodynamically not in equilibrium). The formation of a crystalline or a glassy phase depends on the cooling rate of the melt. A rapid cooling of the melt prevents the relatively slow formation of the ordered crystal structure and freezes the liquid-like structure into the non-equlibrium glassy state.

Transport properties

A system where some transport phenomena are under way (e.g. transport of matter, of heat, of electric charge) is not in thermodynamic equilibrium.

Example: One cannot define a unique temperature for a system where heat is flowing. The system is not in thermal equilibrium, and hence it is not in thermodynamic equilibrium.

A state of non equilibrium can be stationary (steady state) or non-stationary.

Example 1: A metal block initially at temperature T_1 is immersed in a water bath at temperature T_2, and exchanges heat until thermodynamic equilibrium is reached. During the exchange of heat, the system is in a non-stationary state of non-equilibrium.

Example 2: The two extremes of a metal bar are maintained in contact with two reservoirs at temperatures T_1 and T_2, respectively. Heat flows continuously within the bar from the hotter to the colder reservoir. The bar is in a stationary state of non-equilibrium.

The behaviour of systems where transport phenomena are under way is described by phenomenological equations derived from experimental measurements of macroscopic quantities.

Example: For heat transport, the heat flux $\vec{J}_Q$ is connected to the temperature gradient through

$$\vec{J}_Q \;=\; -\,K_{\rm th}\,\vec{\nabla}T\,, \tag{1.11}$$

where $K_{\rm th}$ is the thermal conductivity.

The Thermodynamics of irreversible processes allows one to treat non-equilibrium stationary processes, establishing evolution and stability criteria based on the local production of entropy. It also sets up a macroscopic theoretical framework for coupled irreversible processes, such as thermo-electrical or thermo-mechanical effects.

1.2 Microscopic approach

The basic goal of the microscopic approach is to determine the properties of matter at the atomic level and to connect them to the macroscopic properties. To this aim it is necessary to determine how the atoms are linked together to form the macroscopic structure, taking into account the dynamics of ions and the role of electrons.
It can be convenient to distinguish three main subjects, which are anyway strictly connected:

- the geometrical structure, formed by the mutual positions of atoms,
- the vibrational structure,
- the electronic structure.

1.2.1 Geometric structure

Experimental information on the geometrical structure at the atomic level is generally obtained by the joint phenomena of scattering + interference (globally referred to as "diffraction") of suitable probes, typically X-rays, neutrons or electrons.

To understand the origin of the different arrangement of atoms in different systems, it is important to take into account

a) the interaction between atoms, which can often be calculated with good accuracy from first principles but can even be fruitfully modelled in terms of different types of chemical bond (ionic, covalent, Van der Waals, metallic, hydrogen bond),
b) the kinetics at the atomic level, both in the final phase and in the formation process (e.g. cooling of melt), which is responsible for the formation of stable, metastable or unstable phases.

Gases

The interaction between atoms or molecules is very weak, and has very little influence on atomic kinetics. Weak short-range correlations between atomic positions can anyway exist and give origin to faint diffuse halos in diffraction patterns.

Liquids

The interactions between atoms are strong enough to guarantee a very weak dependence of volume on pressure and temperature, not enough however to prevent a fast atomic kinetics (weak viscosity). Short-range structures are frequently present, for example H_2O molecules in liquid water, tetrahedral coordination of Si to four O atoms in liquid silica, icosahedral coordination in some metallic liquids. Short-range correlations (nearest-neighbour distances) and sometimes medium-range correlations (distances of some ångström) give rise to well defined halos in diffraction patterns.

Solids

In solids, the strong interactions between atoms or molecules give rise to frozen structures.

A) In *crystalline solids*, the atomic arrangement is characterised by the correlation of atomic positions over the full size of single crystals, which corresponds to translational and rotational symmetry properties. The presence of long-range order gives origin to intense Bragg peaks in diffraction patterns, which facilitates the determination of the structure. Different crystal structures are possible for the same chemical composition (allotropy or polymorphism).

B) In *non-crystalline solids*, such as glasses, only short-range correlations (nearest-neighbour distances) and sometimes medium-range correlations (distances of some ångström) are present, giving rise to well defined halos in diffraction patterns; there is no possibility of univocal structural determination from diffraction patterns. The situation is similar to that of liquids. Non crystalline solids are in a thermodynamic state of non-equilibrium.
The difference between crystalline and non-crystalline solids is connected to the type of chemical bond and to the kinetics of formation from melt cooling.

C) *Nanoparticles* are atomic aggregates (crystalline or non-crystalline) whose size is between 1 and 100 nm. In case of narrow size distributions, one speaks of nanoclusters. In contrast to molecules, nanoclusters do not have a fixed size or composition. Nanocluster properties are often qualitatively different from those of their constituent parts and from those of macroscopic aggregates of the same composition. Nanoparticle properties (electronic, optical, vibrational) can vary dramatically with size. These peculiarities give rise to many technological applications, in the biomedical, optical and electronic fields.

1.2.2 Vibrational structure

The vibrations of atoms are responsible for a number of equilibrium as well as transport properties of matter. The specific heat and the thermal conductivity of solids are examples of such equilibrium and transport properties, respectively.
The study of the vibrational properties of molecules and solids is a typical many-body problem, and is based on a number of approximations.

1. The *adiabatic approximation* allows the separation of the motion of ions (nuclei + core electrons) from the motion of (valence) electrons; the total energy of the valence electrons system is then considered as a contribution to the potential energy of the ion system.
2. In the *Born-von Karman model*, the potential energy of the ion system is expanded as a power series of the ionic displacements with respect to the equilibrium positions.
3. In the *harmonic approximation*, only the second-order terms of the Born-von Karman expansion are retained. The dynamics of N strongly interacting ions can be transformed into the dynamics of $3N$ independent and delocalised normal modes, each one behaving as a harmonic oscillator. Attention is then focussed on the quantised energy content of normal modes. In the ground state of the system, at zero kelvin, only the zero-point energy is present. When temperature is increased, energy quanta (*phonons*) are progressively stored in the normal modes, according to the Bose-Einstein statistics; phonons can thus be created and annihilated. Phonons are an example of collective excitations. The harmonic approximation is sufficient to account for the main properties of the vibrational contribution to the specific heats.
4. *Anharmonic contributions* to the crystal potential are taken into account by higher-order terms in the Born-von Karman expansion, and are necessary to account for the thermal expansion and for the resistivity to heat conduction. Anharmonicity is generally treated by perturbative approaches with respect to the harmonic Hamiltonian; normal modes can exchange energy by annihilation and creation of phonons, the number of phonons involved corresponding to the order of the perturbation term.
   To account for the resistivity to thermal conduction, phonons are considered as moving particles (represented by suitable wave packets) and the exchanges of energy between normal modes are considered as collisions between particles.

The approach to vibrational dynamics in crystals (sometimes referred to as lattice dynamics) is facilitated by symmetry considerations. More complex is the treatment of vibrational dynamics of non-crystalline systems.

1.2.3 Electronic structure

Also the electronic structure is responsible for both equilibrium and non-equilibrium properties of solids, such as the arrangement of atoms (geometric structure), the electronic contribution to specific heat, the thermal and electric conductivities.
Again, the many-body problem can be faced by different approaches and degrees of approximation.

1. The simplest approximation considers the valence electrons as a gas of free particles, neglecting the interactions between electrons and ions and between electrons. The eigenfunctions of the Schrödinger equation are plane waves labeled by the wave-vectors $\vec{k}$; the energy is purely kinetic . The ground state is obtained by filling up the possible states according to the Fermi-Dirac statistics.
   Although quite crude, the free electron model is sufficient to account for the main properties of the electronic specific heat.
2. The interaction of the electrons with the periodic array of ions can be accounted for within a self-consistent field approximation, where single electrons are affected by a periodic potential due to the ions and to all the other electrons. As a consequence of the translational symmetry of the crystal, one can establish some general properties of the electron wavefunctions (Bloch functions) and of the dispersion relations $\epsilon(\vec{k})$ connecting energy to wavevector, leading to the electronic band structure and to the explanation of the different conduction properties of metals, semiconductors and insulators.
3. Different techniques have been developed to calculate the band structure for different systems. One has again to distinguish between ground state and excited states. The collective excitations of the electron many-body system are described in terms of quasi-particles. The interaction among electrons and among electrons and phonons is necessary to account for thermal conductivity in metals, as well as for important properties such as magnetism and superconductivity.

1.3 Key concepts

A number of basic concepts underpin and simplify the interpretation of the basic properties of matter.

1.3.1 Direct space and reciprocal space

The possibility of the alternative description of a given phenomenon in two conjugate spaces connected by the Fourier transform is a powerful tool in many branches of Physics (e.g. the coordinate and momentum representations of quantum mechanical states) as well as for many technological applications (e.g. the time and frequency domains in the analysis of electrical signals). In Solid State Physics the two Fourier-conjugated spaces are called direct (or real or coordinate) space and reciprocal (or wave-vector or momentum) space. The two spaces are connected to the complementary descriptions in terms of particle or wave properties.
In these lectures, we consider the following main applications:

1. *Geometric structure.* The Fourier transform is a powerful tool for the interpretation of diffraction patterns: as a matter of fact, diffraction techniques give a representation of the structure of matter in the reciprocal space. For crystalline solids, to the Bravais lattice in direct space it corresponds the reciprocal lattice in reciprocal space (discrete Fourier transform). To a non periodic structure in real space it corresponds a non periodic structure in reciprocal space (continuous Fourier transform).

2. *Vibrational structure.* Normal modes of vibration are described by wave-vectors in the reciprocal space. For each wave-vector, different values of frequencies are possible (different dispersion relations). A limited region of the reciprocal space (the first Brillouin zone) contains all the information on the lattice dynamics of a crystal.
3. *Electronic structure.* The interplay between real and reciprocal space is relevant for understanding the ground state and the transport properties of the electron system. The Bloch wave-functions of electrons moving in a periodic lattice potential are characterised by wavevectors in the reciprocal space. As for phonons, a limited region of the reciprocal space (the first Brillouin zone) contains all the information on the electron dynamics of a crystal.

1.3.2 Symmetry

The first scientific applications of symmetry were connected with the study of the macroscopic morphology of natural crystals. The idea of symmetry was progressively clarified and connected to the concepts of transformation and invariance of the objects (or of their properties) under symmetry transformations. The group theory allows an elegant and powerful treatment of symmetry properties. According to the Von-Neumann principle, every physical property must be invariant under the symmetry operations of the system. These invariance properties represent a powerful tool in many fields of Physics.
In these lectures, we focus our attention on a number of applications of discrete symmetries to geometric, vibrational and electronic structure of solids.

a) Geometrical symmetries affect molecules (point symmetries) and crystals (point and translational symmetries). Symmetry operations form groups. Point and space groups are key concepts for the classification of molecules and crystals as well as for the interpretation of diffraction spectra.
b) Symmetry groups can be represented in mathematical form by matrices which represent transformations in suitable vectorial spaces. Different representations are possible for the same symmetry group, depending on the vectorial space and on the basis vectors that are chosen. Of the possible representations, particularly relevant are the so called irreducible representations.
c) The study of irreducible representations can give important qualitative and semi-quantitative information on a number of physical and chemical phenomena, such as chemical bond formation, radiative transitions, degeneracies of energy levels, Bloch functions of electrons in crystals.

1.3.3 Complexity and emergence

Quoting the American physicist and Nobel laureate P.W. Anderson (Science 177, 393-396, 1972):
«The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe. (...) The constructionist hypothesis breaks down when confronted with the twin difficulties of scale and complexity. The behaviour of large and complex aggregates of elementary particles is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear. (...) Psychology is not applied biology, nor is biology applied chemistry. (...) We can now see that the whole becomes not merely more, but very different from the sum of its parts.»
«In principle, the structure and behaviour of condensed matter can be reduced to the few laws of fundamental interactions, in particular, for our purposes, the laws of electromagnetism (reductionist hypothesis). By converse, reconstructing the structure and behaviour of condensed matter from the basic laws is far from trivial: this "constructionist" hypothesis breaks down when confronted with the complexity of large systems. At each level of complexity entirely new properties appear: the whole becomes not merely more, but very different from the sum of its parts.»
The appearance of new levels of complexity is sometimes called emergence. It is the case, for example, of the onset of collective properties of electrons, such as magnetic ordering in ferromagnets or more generally of the breaking of symmetries.

1.3.4 Approximations and models

In principle, the laws governing the behaviour of atomic aggregates are well known. Basically, the properties depend on the electromagnetic interaction.
Exact solutions are however generally impossible. Different approximate approaches are developed for different situations. Sometimes the same phenomenon can be described by different alternative and approaches. Sometimes different approaches of different complexity are possible.

Example 1: The Born-von Karman power expansion of the crystal potential is an approximation. The further harmonic approximation can however quite accurately account for some phenomena, such as specific heats, but completely fails to explain other phenomena, such as thermal expansion.

Example 2: The interaction of electromagnetic radiation with matter can be interpreted at different levels of approximation. A purely classical treatment can be sufficient to account for the basics of X-ray diffraction. A quantum treatment is necessary to account for X-ray absorption.

Phenomenological models are frequently used.

Example 1: Chemical bonds are phenomenological models useful to describe different properties of the interaction between atoms.

Example 2: The free electron gas model is an approximate approach that can quite satisfactorily account for the electronic contribution to the specific heats of metals.

1.4 Main experimental techniques

1.4.1 Production techniques

At the beginning of the 20th Century, the scientific research was limited to solids (crystals and glasses) that could be found in nature.
Since then, a number of different techniques have been progressively developed to produce condensed-matter systems in the laboratory. The availability of different techniques has in turn led to the continuous development of new materials, whose properties are often technologically interesting and scientifically challenging.
Let us give here some examples.

a) Polycrystalline powders or glasses can be obtained by relatively fast cooling of melts.
b) Single-crystals can be grown by different methods
   - Czochralski : crystals are slowly pulled on from a melt, starting with a small seed
   - Bridgman: powders are melted in a hollow cylinder
   - supersaturated solutions can precipitate giving birth to single crystals
c) Deposition of thin films on a substrate; different methods are available to bring atoms to the surface, broadly distinguished as:
   - chemical deposition methods, where a fluid precursor undergoes a chemical change at a solid surface, leaving a solid layer;
   - physical deposition methods, where mechanical, electromechanical or thermodynamic means are used to produce a thin film of solid.
d) Sol-gel techniques: from a chemical solution (sol) that acts as the precursor for an integrated network (gel) of either discrete particles or network polymers

1.4.2 Characterisation techniques

Along with the production techniques, also the characterisation techniques are in continuous evolution.

Macroscopic techniques

Macroscopic techniques include

a) Techniques based on equilibrium thermodynamic properties: one measures the response of macroscopic variables (energy, volume, etc.) to variations of the temperature, pressure, etc.
b) Techniques based on transport properties: on sets up an electric potential gradient $\nabla\phi$ or a thermal gradient ∇T and one measures the electrical current or the heat flow. The gradients can be held constant or constained to oscillate at a well defined frequency.

Atomic-level techniques

Atomic-level techniques are based on the interaction of suitable probes (electromagnetic radiation, electrons, neutrons,...) with matter.

a) *Elastic scattering*. Elastic scattering and interference of scattered waves (for short "diffraction") is the basic technique for studying the structure of matter at the atomic level. The most powerful probe are X-rays, thanks to the easiness of production and of interpretation of diffraction spectra. Frequently used are also neutrons and electrons
b) *Inelastic scattering* is used to study the excitations of matter, by measuring the energy and momentum exchanged between probe and matter. Excitations can be single-particle (e.g. Compton effect) or collective (e.g. plasmons, phonons). Different information can be obtained from inelastic scattering of neutrons, X-rays, optical radiation (Brillouin and Raman scattering).
c) *Spectroscopy*. Absorption and emission of electromagnetic radiation gives information on the electronic structure of matter. Some spectroscopic techniques can also give information on the local geometrical structure of matter, e.g. NMR (Nuclear Magnetic Resonance) or EXAFS (Extended X-ray Absorption Fine Structure).

1.4.3 Large-scale facilities

The progressive development and refinement of experimental characterisation techniques has been accompanied by the development of large national and international facilities:

a) Synchrotron Radiation facilities for the generation of well collimated and intense electromagnetic radiation in a broad spectrum extending from infrared to X-rays. They offer unique possibilities for X-rays scattering, spectroscopy and imaging.
b) Free electron lasers (FEL) for the production of very high intensity, short pulsed and coherent electromagnetic radiation in the spectral regions of UV, soft X-rays and now also hard X-rays.
c) Neutron facilities, where neutrons produced by reactors or spallation sources are used for spectroscopy or scattering (elastic and inelastic) experiments.

1.5 Bibliography of Chapter 1

General References for the whole course:

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (First edition 1976, various reprints available). A very good classical textbook on Solid State Physics, unfortunately never updated.
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. A classical textbook on Solid State Physics, progressively updated.
- W. A. Harrison: *Solid State Theory*, Dover (1979). Reprint of a classical 1970 textbook.
- M. L. Cohen and S. G. Louie: *F*undamentals of Condensed Matter Physics, Cambridge University Press (2016). An advanced and updated textbook for graduate students.

Introductory macroscopic and statistical Thermodynamics, non-equilibrium Thermodynamics:

- Mark W. Zemansky: *Heat and Thermodynamics*, MacGraw Hill (last edition 1996). A classical textbook.
- H.B. Callen: *Thermodynamics and an Introduction to Thermostatistics*, MacGraw Hill (Wiley, 2nd. ed., 1985). A textbook based on an axiomatic introduction of the laws of Thermodynamics.
- F. Reif: *Statistical and Thermal Physics*, McGraw-Hill (1988), ISBN 0-07-Y85615-X. A classical textbook.
- L. E. Reichl: *A Modern Course in Statistical Physics*, Edward Arnold Ltd (1980), ISBN 0-7131-35174.
- I. Prigogine: *Introduction to Thermodynamics of Irreversible Processes*, John Wiley & Sons (1968)
- S. R. de Groot, P. Mazur: *Non-equilibrium thermodynamics*, Dover Publications (1984)

On complexity and emergence:

- Philip W. Anderson: *More is different*, Science, vol. 177, page. 393-396, (4 August 1972). A frequently cited seminal paper.

Chapter 2

Real-space structure

In this chapter, the basic elements are given for a description of the atomic arrangements of condensed matter in the real space (or direct space).
The attention is mainly focused on perfect crystals, where the presence of a regular geometrical arrangement allows a description in terms of three-dimensional periodic functions (§ 2.2–2.4.3). In § 2.5, an introductory account is given of the defects that can affect real crystals. The structural description based on the Radial distribution function (RDF), necessary for non-crystalline systems (liquids and glasses), is introduced in § 2.6.

2.1 Introduction

Atoms and crystals

For several centuries, mineralogists have known the existence of macroscopic crystals, with regular geometrical shapes. External observations led to recognise some regularities (such as the constancy of angles between planes), which suggested R.J. Haüy (1743-1822) that crystals could be regular packings of microscopic structural units.

The success of the atomic hypothesis among chemists of the XIX Century led to the hypothesis that the basic structural units of crystals could be the atoms. Calculations based on the known atomic masses suggested that the interatomic distances could be of the order of magnitude of the typical X-ray wavelengths (some ångströms). In 1912 M. von Laue suggested that atoms in a crystalline solid could form a 3-dimensional diffraction grating for an X-ray beam and shortly thereafter the first X-ray diffraction pattern from a crystal was obtained. In 1913, W. H. Bragg and W. L. Bragg gave a simple quantitative interpretation of X-ray diffraction from crystals, thus opening the possibility of its use for studying the atomic structure of crystalline solids.

Macro and micro-crystals

The crystalline structure is not necessarily evident at the macroscopic scale.

Many materials, typically metals, are made up by small mono-crystalline regions randomly oriented (called grains), whose typical size, of the order of some micrometer, is small on the macroscopic scale, but very large when compared with the X-rays wavelengths (Fig. 2.1, left panel).

Also the structure of macroscopically crystalline systems (monocrystals) is generally affected by mosaicity, say by the presence of regions of micrometric size, slightly misoriented by angles that can range between 0.01° and 0.1° (Fig. 2.1, right panel), as well as by other kinds of defects, such as point defects and dislocations (see below). Monocrystals of relatively large dimensions (of the order of many cm) free from mosaicity can be obtained for some substance, for example silicon, by appropriate growth methods.

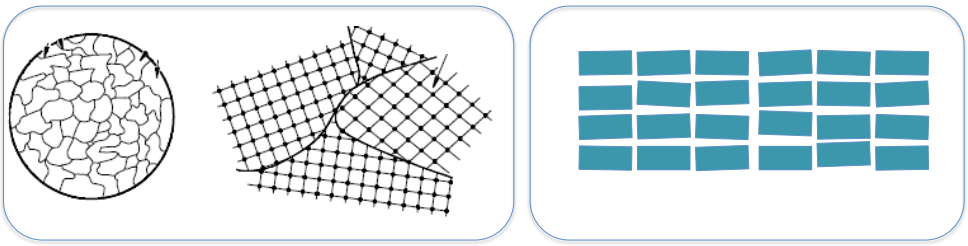

Figure 2.1: Left panel: microscopic and atomic-scale picture of a grain structure. Right panel: mosaic structure.

Deviations from the perfect crystal model

Crystals with a perfectly regular geometrical structure don't actually exist. For every crystal, thermal motion destroys the regularity of instantaneous atomic configurations even at the lowest attainable temperatures (zero-point motion). The intensity of thermal motion increases with temperature. Besides, the presence of defects (vacancies, interstitials, dislocations, stoichiometric defects, etc) is unavoidable, if only for statistical reasons. Defects play a fundamental role to explain some physical properties of crystals, such as heat conduction or plastic deformation.
In § 2.2–2.4.3, we consider the ideal case perfect crystals of infinite extension. Deviations from this model are considered in § 2.5

Non-crystalline materials

Non-crystalline materials are currently found in nature, in both liquid and solid state (e.g. volcanic lava) and have been artificially produced for many centuries (glasses). The present interest for non-crystalline materials is connected with their technological applications (for example amorphous semiconductors or glassy metallic alloys). Non-crystalline materials lack the long-range order of crystals, and the description of their structural and physical properties is less complete (§ 2.6).

2.2 Crystal structure: basics

The microscopic structure of perfect crystals can be conceived as made up by (Fig. 2.2)

a) a geometrical lattice (the Bravais lattice)
b) a structural unit (the basis or motif) regularly repeated at each point of the Bravais lattice.

The basis (motif) can be one atom (e.g. in copper), or a small group of atoms (2 atoms in Ge and in NaCl, 4 atoms in AgI), up to a molecule containing thousands of atoms (e.g. biological molecules, such as proteins, in artificially produced crystals).

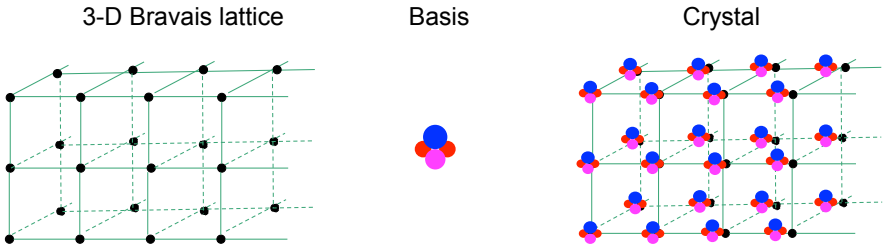

Figure 2.2: A perfect crystals can be conceived as an infinite Bravais lattice and a basis associated to each point of the lattice.

The decomposition *(lattice) + (basis)* can be considered for one, two or three-dimensional crystals. Real crystals are in general three-dimensional. Drawing two-dimensional structures is however simpler; as a consequence, two-dimensional structures are often considered for didactical purposes. Real linear and planar structures are anyway becoming more and more familiar in science and technology. The graphene foil is a good example of a truly two-dimensional crystal.

2.2.1 Bravais lattices

Definitions of Bravais lattice

A Bravais lattice can be defined in two equivalent ways:

1. A Bravais lattice is an infinite array of discrete points with an arrangement and orientation that appear exactly the same, from whichever of the points the array is viewed.
2. A (three-dimensional) Bravais lattice is the set of all points whose position can be expressed by a translation vector $\vec{T}$ defined as

$$\vec{T} = u\vec{a} + v\vec{b} + w\vec{c}, \tag{2.1}$$

where $\vec{a}, \vec{b}, \vec{c}$ are non co-planar vectors and u, v, w are integer numbers.

The first definition enlightens the symmetry properties of the lattice. The second definition is more suited for analytical descriptions.

Non-Bravais lattices

In a Bravais lattice each point has "equal surroundings". Not all conceivable arrays of points are Bravais lattices. Fig. 2.3 shows a simple example: within the array on the left of the figure there are points with different surroundings. If, however, we suitably group the points three by three, and substitute each group of three original points by only one point, we obtain a Bravais lattice (centre). We can again describe the original non-Bravais array as a structure composed by a Bravais lattice plus a basis of three points (right).

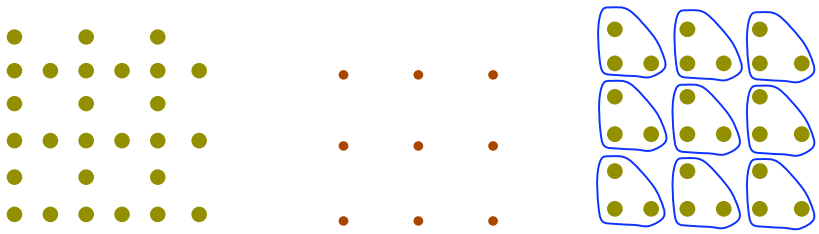

Figure 2.3: The array of points of the left figure doesn't correspond to a Bravais lattice. A possible Bravais lattice is shown in the centre figure. The basis connected to each lattice point contains three of the original points (right).

Translation vectors and primitive vectors

Once an origin is arbitrarily chosen at a Bravais lattice point, every other point of the Bravais lattice is described by a translation vector $\vec{T}$.
A two-dimensional Bravais lattice can be spanned with the aid of two primitive non-parallel vectors $\vec{a}$ and $\vec{b}$, so that :

$$\vec{T}_{uv} \;=\; u\vec{a} + v\vec{b} \qquad (u, v \text{ integers})\,. \tag{2.2}$$

Similarly, a three-dimensional Bravais lattice can be spanned with the aid of three primitive vectors $\vec{a}, \vec{b}, \vec{c}$, linearly independent, so that

$$\vec{T}_{uvw} \;=\; u\,\vec{a} + v\,\vec{b} + w\,\vec{c} \qquad (u, v, w \text{ integers})\,. \tag{2.3}$$

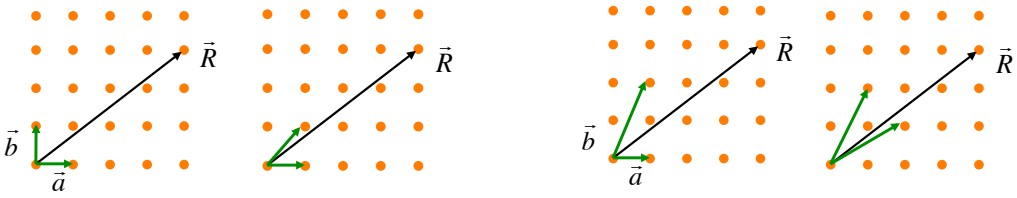

Figure 2.4: Two examples of primitive vectors (left) and non-primitive vectors (right) for a 2D Bravais lattice.

Note 1: The choice of the primitive vectors for a given lattice is not unique. Actually, whichever three vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ are primitive if each translation vector $\vec{T}$ can be expressed as a linear combination of $\vec{a}$, $\vec{b}$ and $\vec{c}$, with integer coefficients. (See Fig. 2.4, left, for a 2D example) Viceversa, not all vectors joining the origin with three whatever lattice points are primitive vectors. (See Fig. 2.4, right, for a 2D example)

Note 2: Equation (2.3) conforms with the standard nomenclature of elementary crystallography. A frequently used alternative form is

$$\vec{T} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 \tag{2.4}$$

2.2.2 Unit cells

Primitive unit cells

A parallelogram (in 2D) or a parallelepiped (in 3D) built over any set of primitive vectors represents a *primitive cell* (Fig. 2.5, top). The plane (in 2D) or the space (in 3D) can be completely filled up by an infinite repetition of one primitive cell. For each choice of primitive vectors, the shape of the primitive cell is determined, but infinite choices of the position of the primitive cell are possible (Fig. 2.5, bottom). A primitive cell corresponds to one Bravais lattice point.
Primitive cells don't necessarily exhibit the full symmetry of the lattice. (Crystal symmetries are discussed in Chapter 4). To evidence the symmetry properties of the lattice, two different approaches are widely used:

1) choose as primitive cell a Wigner-Seitz cell, whose peculiar definition intrinsically preserves the crystal symmetry
2) choose a non-primitive cell, say a *conventional* unit cell (or crystallographic cell)

Wigner-Seitz primitive cell

Definition. The Wigner-Seitz cell surrounding a given Bravais lattice point is the region of space that is closer to that point than to any other point of the lattice. The Wigner-Seitz cell maintains the full symmetry of the Bravais lattice. Its shape is however generally not simple.
Construction. Draw lines connecting the given lattice point with all other lattice points; bisect each line with a plane; the Wigner-Seitz cell is the smallest polyhedron containing the point and bounded by these planes.
Generalisation. The procedure depicted above for constructing the Wigner-Seitz cell can be applied to whichever set of points, not necessarily a Bravais lattice (for example, the atomic positions in an amorphous solid). The generalisation of Wigner-Seitz cells to generic sets of points is represented by the so-called Voronoi polyhedra.
Reciprocal space. The concept of Wigner-Seitz cell plays a particularly relevant role in the reciprocal space, where the Wigner-Seitz cell centred on the origin of the reciprocal space is called "First Brillouin zone" (Chapter 3).

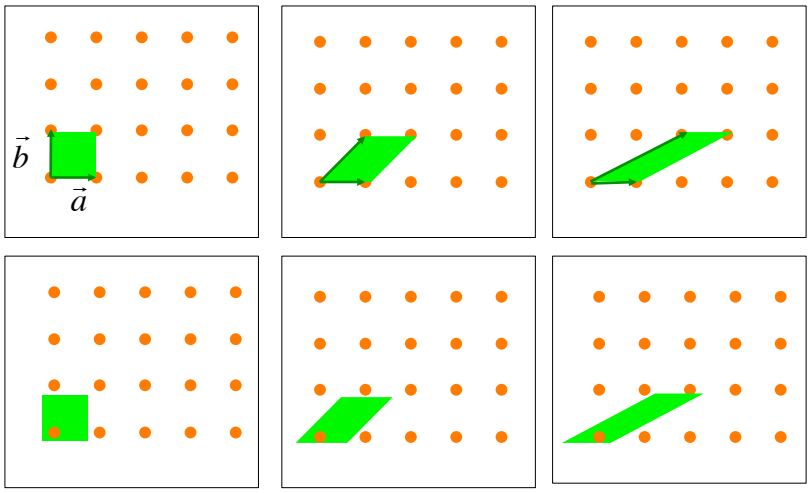

Figure 2.5: Top: three different choices of primitive cell for the same 2D Bravais lattice. Bottom: the position of the primitive cells can be shifted.

Conventional unit cells

One can fill up the plane (2D) or the space (3D) by an infinite repetition of non-primitive cells, say cells containing more than one Bravais lattice point.
Non-primitive cells are called *conventional cells*. Conventional cells are in many cases preferred to primitive cells, typically in structural studies, when they maintain the full symmetry of the lattice.

Problem: Consider a 2D rectangular centered Bravais lattice. Draw a set of possible primitive cells. Draw the Wigner-Seitz cell. Draw a conventional cell that preserves the symmetry of the lattice. Individuate the symmetry operations of the lattice and verify that they are shared by the Wigner-Seitz cell and by the conventional cell.

2.2.3 Characterization of unit cells

Cell parameters

A geometrical lattice can be fully characterized by two alternative choices:

1. The parameters of a primitive cell, say the lengths of its primitive vectors and the angles in between (Fig, 2.6, left).
2. The type (simple cubic, body centred cubic, face centred cubic, etc) and the parameters of the conventional unit cell, say the edges of the cell and the angles in between.

In two dimensions, there are at most two different lengths of primitive vectors, a and b, and one angle γ. In three dimensions, there are at most three different lengths of primitive vectors, a, b, c and three different angles in between, α, β, γ.

Internal coordinates

If more than one atom is contained in the unit cell, be it primitive or conventional, we need to describe the position of each atom within the cell (Fig, 2.6, right).
To this purpose, one again takes the unit vectors (primitive or conventional) as reference, and the position of each atom is identified by a set of fractional coordinates (two or three coordinates in two or three dimensions, respectively).

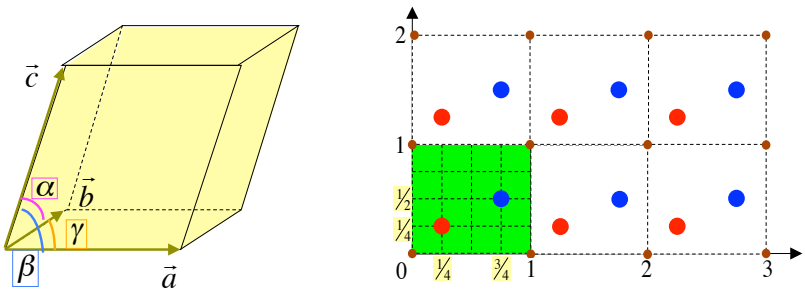

Figure 2.6: Left: parameters of a primitive cell. Right: internal coordinates of atoms inside a cell.

2.3 Crystal structures: examples

An exhaustive classification of the Bravais lattices and of the crystal structures is made according to the symmetry properties. We consider this subject in some detail in Chapter 4. For now, let us stress that in three dimensions there are only 14 different Bravais lattices, and all possible crystal structures can be classified within a set of 230 groups of symmetry.
In this § 2.3 some particularly relevant examples are given in order to familiarise with the concepts of Bravais lattice and of primitive and conventional unit cells.

2.3.1 The planar lattice of graphene

The two-dimensional structure of graphene is made of an hexagonal honeycomb of carbon atoms, at a distance $r = 0.142$ nm (Fig. 2.7, left). The positions of carbon atoms don't form a Bravais lattice: one can easily check that there are two non equivalent positions. The primitive cell contains two atoms: the graphene structure is made by a Bravais lattice + a two-atom basis.
One rhombic primitive cell is shown in Fig. 2.7 (left), with $a = \sqrt{3}\, r$ and the two atoms at positions (0, 0) and (2/3, 2/3). In the same figure one can see that the Wigner-Seitz cell is hexagonal
The 2D graphene sheets can be obtained in laboratory. They are also the basis of important structures (Fig. 2.7), such as:

- Graphite (the crystallographic form of carbon thermodynamically stable at ambient conditions) made up by parallel graphene sheets spaced by 0.34 nm.
- Fullerene buckyballs, whose most common is the Buckminster fullerene C_{60}
- Carbon nanotubes

2.3.2 Cubic Bravais lattices

As examples of 3D Bravais lattices, let us consider here the three Bravais lattices sharing the cubic simmetry, which is the highest possible symmetry for a crystal (Fig. 2.8). For cubic lattices, the conventional unit cell is characterised by one lattice parameter, the length of the cell edge a.

a) Simple cubic lattice (sc) (Fig. 2.8, left). The conventional unit cell is cubic and is *primitive* (one lattice point per cell). Each lattice point is surrounded by 6 nearest neighbors points.
b) Body centred cubic lattice (bcc) (Fig. 2.8, centre). The conventional unit cell is cubic and is *not primitive*: each conventional cell contains two lattice points (one at a corner and one at the centre). Each lattice point is surrounded by 8 nearest neighbors points.
c) Face centred cubic lattice (fcc) (Fig. 2.8, right). The conventional unit cell is cubic and is *not primitive*: each conventional cell contains four lattice points (one at a corner and three at face centres). Each lattice point is surrounded by 12 nearest neighbors points.

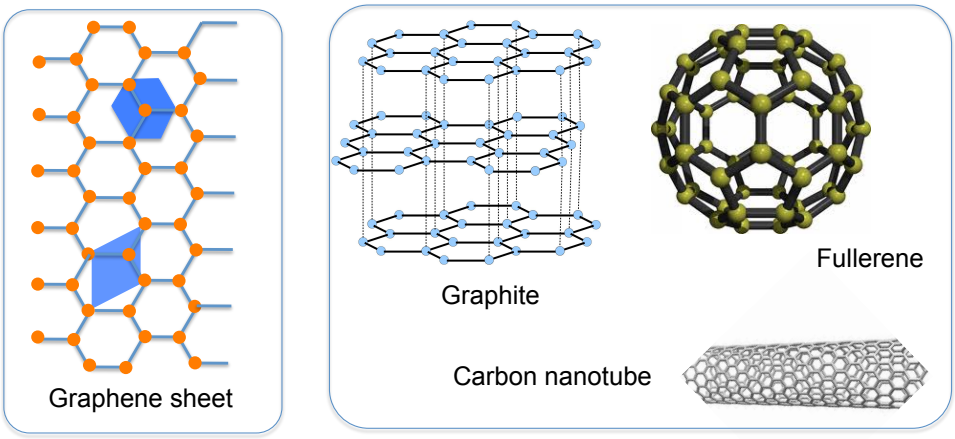

Figure 2.7: Left panel: a graphene sheet: in blue a rhombic primitive cell and the hexagonal Wigner-Seitz cell. Righ panel: graphite structure, fullerene C_{60} and carbon nanotube.

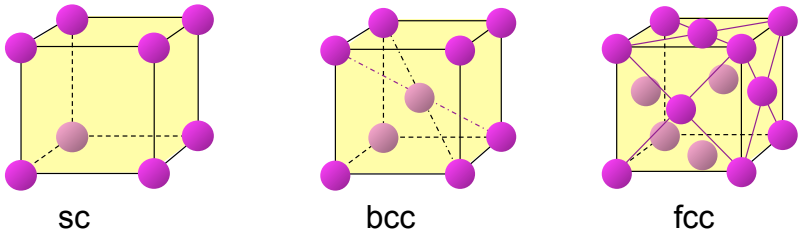

Figure 2.8: The three cubic Bravais lattices. From left: simple cubic (sc), body centered cubic (bcc) and face centered cubic (fcc). The lattice points are represented as small spheres for graphical purposes. The conventional cell is primitive only for the simple cubic case.

Problem: Choose a set of primitive vectors for the bcc structure and draw the corresponding primitive cell (see § 2.7.1). Draw the Wigner-Seitz primitive cell (Fig. 2.9, left).

Problem: Choose a set of primitive vectors for the fcc structure and draw the corresponding primitive cell (see § 2.7.1). Draw the Wigner-Seitz primitive cell (Fig. 2.9, right).

To each Bravais lattice many different crystal structures can correspond, depending on the structure of the basis. To exemplify, let us consider some simple structures which are frequently found in Physics.

A) fcc Bravais lattice - the fcc structure

The fcc structure (Fig. 2.10, left) is made up by an fcc Bravais lattice and a one-atom basis: there is one atom per primitive cell. Each atom is surrounded by 12 nearest-neighbour atoms, say the coordination number is 12. This value corresponds to the so called close-packing.
The close-packed fcc structure is shared by a large number of metals, such as Ni, Cu, Rh, Pd, Ag, Ir, Pt, Au.

Problem: Verify that the nearest neighbour distance r_1 is connected to the lattice parameter a (length of the primitive vector, corresponding to the edge of the cubic cell) by the relation $r = a/\sqrt{2}$. Find out also the distance of the 2nd and 3rd nearest neighbours.

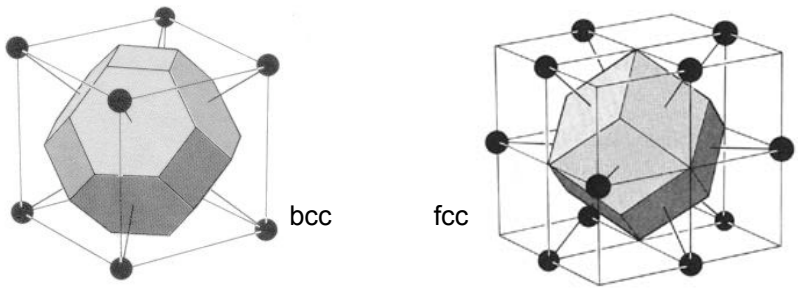

Figure 2.9: Wigner-Seitz (WS) primitive cells for the bcc and fcc Bravais lattices. For the bcc lattice, the WS cell is a truncated octahedron, with 8 hexagonal faces and 6 square faces (left). For the fcc lattice, the WS cell is a rhombic dodecaedron with 12 faces (right).

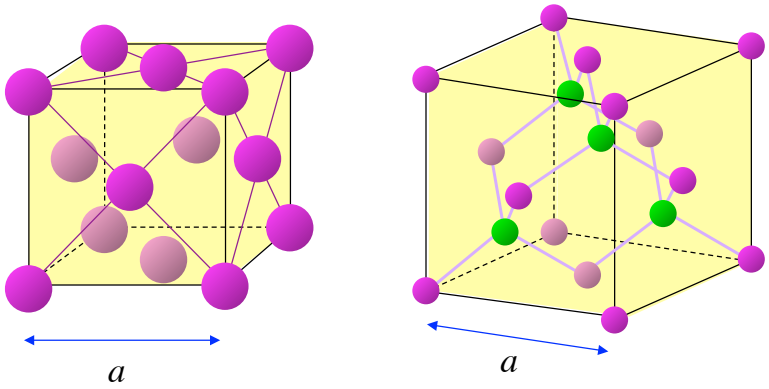

Figure 2.10: The fcc structure (left) and the zincblende structure (right) share the fcc Bravais lattice.

B) fcc Bravais lattice - the diamond and zincblende structures

The diamond structure and the zincblende (or sphalerite) structures (Fig. 2.10, right) are made up by an fcc Bravais lattice and a two-atom basis: there are two atoms per primitive cell. The positions of the two atoms are identified by their fractional coordinates (0,0,0) and (1/4, 1/4, 1/4) with respect to the conventional cell. Each atom is surrounded by 4 nearest neighbour atoms (coordination number = 4). The low-coordinate structure is typical of crystals where the interatomic bond has a strong covalent character.
In the diamond structure the two atoms of the primitive cell are equal. The structure is shared by diamond (carbon), silicon and germanium.
In the zincblende (or sphalerite) structure the two atoms of the primitive cell are different. The structure is shared by many binary compounds, such as ZnS (zincblende or sphalerite), GaAs, CdTe.

Problem: Find out the relation between the nearest-neighbour distance r_1 and the lattice parameter a for the diamond structure.

C) fcc Bravais lattice -the rocksalt structure

Also the rocksalt (NaCl) structure (Fig. 2.11, left) is made up by an fcc Bravais lattice and a two-atom basis. The positions of the two atoms are however different with respect to those of the diamond and zincblende structures; their fractional coordinates are (0,0,0) and (1/2, 1/2, 1/2) with respect to the conventional cell.
In the rocksalt structure, each atom of a given species is surrounded by 6 nearest neighbour atoms

Table 2.1: Some structural properties for selected crystals.

	Cu	Ge	NaCl
Lattice parameter a (Å)	3.61	5.66	5.64
Atoms/primitive cell	1.	2.	2.
Atoms/conventional cell	4.	8.	8.
Volume per atom (Å^3)	11.76	22.66	22.42

of the other species (coordination number 6). The rocksalt structure is typical of crystals where the interatomic bond has a prominent ionic character.

Problem: Find out the relation between the nearest-neighbour distance r_1 and the lattice parameter a for the rocksalt structure.

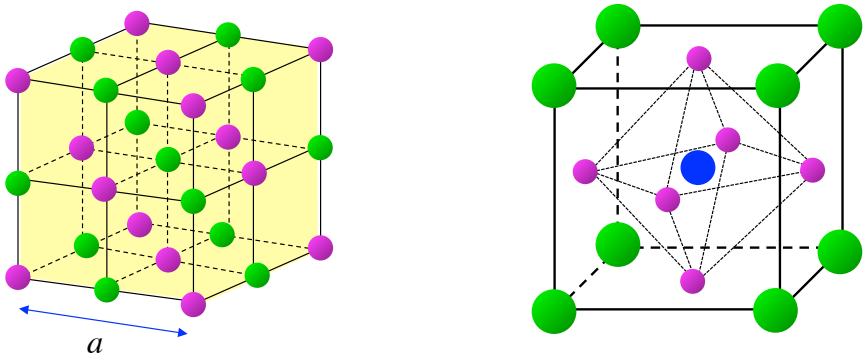

Figure 2.11: The rocksalt structure (left) and the perovskite structure (right).

D) sc Bravais lattice - the perovskite structure

The perovskite ($CaTiO_3$) structure (Fig. 2.11, right) is made by a simple cubic Bravais lattice with five atoms per primitive cell. The Ti positive ion is surrounded by 6 O negative ions forming a regular tetrahedron. Some crystals with the perovskite structure, such as $BaTiO_3$, are characterised by the ferroelectric phase transition: below a certain temperature, the lattice is distorted, loosing its cubic symmetry, and the crystal acquires a spontaneous electric polarisation.

Problem: Starting from the lattice parameters a listed in Table 2.1, evaluate the volume per atom for Cu, Ge and NaCl.

2.4 Formal description of crystal structures

The structure and the physical properties of an infinite perfect crystal (Bravais lattice + basis) can be analytically described according to two slightly different approaches, which are separately presented in the following § 2.4.1 and § 2.4.2.

2.4.1 First approach: periodic function

A given property of a crystal is described by a three-dimensional periodic function, whose period is the primitive cell. Let us give some examples.

1. The periodic electron density $\rho_e(\vec{r})$ of a crystal can be described by

$$\rho_e(\vec{r}+\vec{T}_n) \;=\; \rho_e(\vec{r})\,, \tag{2.5}$$

where $\vec{T}_n = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$ is any translation vector of the Bravais lattice.
X-rays are scattered by electrons, and in principle an X-ray diffraction experiment samples the electron density of the crystal.

2. The potential energy $\Phi(\vec{r})$ of the interaction between a conduction electron and the lattice of ions is expressed as

$$\Phi(\vec{r}+\vec{T}_n) \;=\; \Phi(\vec{r})\,. \tag{2.6}$$

3. The geometric structure of a crystal is often formally described by a function

$$F(\vec{r}) \;=\; \sum\nolimits_n \delta(\vec{r}-\vec{R}_n)\,, \tag{2.7}$$

where the sum is over all the atoms of the crystal, and R_n is the equilibrium position of atom n. The periodicity condition is again

$$F(\vec{r}+\vec{T}_n) \;=\; F(\vec{r})\,. \tag{2.8}$$

Neutrons are scattered by the atomic nuclei; in principle, a neutron diffraction experiment samples the $F(\vec{r})$ function. The function $F(\vec{r})$ defined in (2.7) doesn't take into account the spread of atomic positions due to zero-point and thermal vibrations.

The periodicity of the functions (2.5), (2.6) and (2.8) is generally expressed in terms of their Fourier coefficients (see Chapter 3).

2.4.2 Second approach: periodic Bravais lattice and basis

The periodic Bravais lattice and the non-periodic basis function are separated, taking advantage of the concept of convolution.
This approach is particularly suitable for classifying crystals according to symmetry properties, for interpreting diffraction patterns, as well as for treating atomic vibrations .

Lattice function

A Bravais lattice can be formally represented, making use of delta functions, by a *lattice function*.

In one dimension, the lattice function is

$$L(x) \;=\; \sum_n \delta(x-T_n) \;=\; \sum_n \delta(x-na)\,, \tag{2.9}$$

where the sum is over all the one-dimensional translations $T = na$ of the lattice.

In three dimensions, the lattice function is

$$L(\vec{r}) \;=\; \sum_n \delta(\vec{r}-\vec{T}_n) = \sum_{n_1 n_2 n_3} \delta\left[\vec{r}-(n_1\vec{a}_1+n_2\vec{a}_2+n_3\vec{a}_3)\right]\,, \tag{2.10}$$

where the sum is over all the translation vectors $\vec{T}_n$ of the Bravais lattice.

Basis function

The structure of the basis can be described by a continuous function of position $f(\vec{r})$. For example, $f(\vec{r})$ can be the continuous electronic density inside the primitive cell. This approach is sometimes used for the interpretation of X-ray diffraction patterns.
Alternatively, the basis is sometimes described by a discrete function

$$F(\vec{r}) \;=\; \sum\nolimits_{\kappa} \delta(\vec{r} - \vec{R}_{\kappa})\,, \tag{2.11}$$

where the sum is over the limited number of atomic positions within the primitive cell. This approach is used for the interpretation of both X-ray and neutron diffraction patterns.

Note: In (2.11) the index κ is the greek letter "kappa", which is used here and in the following to label the atomic positions within the primitive cell. It shouldn't be confused with the latin letter k, generally used to label the wavevectors.

Crystal structure

The structure of a perfect infinite crystal can be described by the convolution of the two three-dimensional functions representing the lattice and the basis, respectively:

$$\boxed{S(\vec{r}) \;=\; L(\vec{r}) * f(\vec{r}) \;=\; \sum\nolimits_{n} f(\vec{r} - \vec{T}_n)} \tag{2.12}$$

where again the sum is over all the translation vectors $\vec{T}_n$ of the infinite Bravais lattice ($-\infty < n_,n_2, n_3 < \infty$).
The physical meaning of $S(\vec{r})$ in (2.12) depends on the nature of the basis $f(\vec{r})$: it can represent an electronic density, a potential energy, a set of discrete atomic positions.
The convolution operation (2.12) corresponds to the periodic repetition of the basis at the lattice points. (For an introduction to convolution, see § 2.7.2).

The approach based on the convolution (2.12) is particularly useful in structural studies based on diffraction of X-rays, neutrons or electrons. Diffraction experiments give a picture in the reciprocal space (Chapter 3), say in the space Fourier transformed of the real space. The convolution theorem transforms the relatively complex convolution of two functions in the real space into the simple product of two functions in the reciprocal space.

2.4.3 Crystal planes

The points of a three-dimensional lattice can be grouped into sets of parallel planes in infinite different modes. The interpretation of a number of different phenomena (diffraction patterns, lattice dynamics, motion of electrons in crystals) is greatly facilitated by the organization of the lattice points into families of planes.

Miller indices A family of parallel planes is labeled by three integer numbers, the Miller indices h, k, l, one for each independent direction. A Miller index counts how many times the planes of a family intersect the corresponding edge of the unit cell. A zero value means that the planes are parallel to the corresponding direction. As an example, the simplest and most common families of planes of the fcc structure are shown in Fig. 2.12.

Interplanar distance To each family of planes it corresponds a given inter-planar distance, which depends on the lattice parameters. For *cubic crystals* the relation between lattice parameter a, Miller indices h, k, l and interplanar distance d_{hkl} is quite simple (see § 2.7.3 for a demonstration):

$$d_{hkl}^2 = \frac{a^2}{h^2 + k^2 + l^2}\,. \tag{2.13}$$

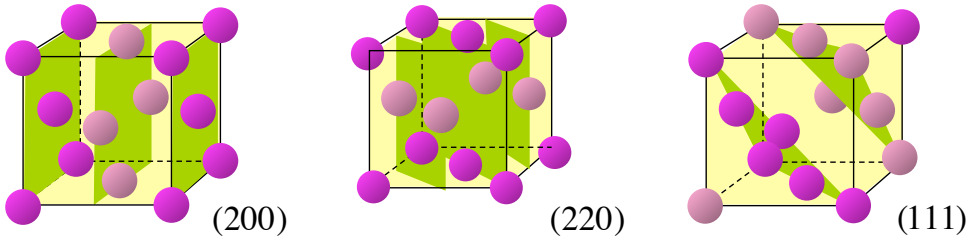

Figure 2.12: The simplest families of planes of the fcc structure.

Planes and directions To each family of planes, labeled by 3 Miller indices hkl, it corresponds a direction perpendicular to the planes, labeled by the same indices.
Different families of planes can be equivalent for symmetry reasons, and share the same interplanar distance. They thus give the same contribution to diffraction patterns (one speaks of multiplicity).
The following convention has been established:

(hkl) indicates a family of planes
$[hkl]$ indicates a direction
$\{hkl\}$ indicates a set of equivalent families of planes
$\langle hkl \rangle$ indicates a set of equivalent directions

Vector representation

A) To each family of lattice planes (hkl) there corresponds a perpendicular direction and an interatomic distance d_{hkl}. The information on direction and interplanar distance can be condensed into a vector $\vec{K}_{hkl}$, whose direction is perpendicular to the planes, and whose magnitude is inversely proportional to the distance, $K = 2\pi/d_{hkl}$.
B) The different vectors $\vec{K}_{hkl}$ correspond to points in the 3-dimensional reciprocal space, which belong to the so called *reciprocal lattice* (see Chapter 3).

2.5 Deviations from the infinite perfect crystal model

Real crystals are nor perfect nor infinite. The ability of dealing with the deviations from the model of the infinite perfect crystals is important for a number of reasons. For example: *a)* the finite size of crystals can affect the shape of diffraction patterns; *b)* surface properties are different from bulk properties; *c)* systems of very small size exhibit structural and functional properties different from the bulk properties; *d)* defects of crystals, whose presence is practically unavoidable, play a fundamental role in explaining some physical properties, such as heat conduction or plastic deformation.
We consider below separately:

- the treatment of finite-size crystals (§ 2.5.1)
- the effect of atomic vibrations (§ 2.5.2)
- the nano-structures (§ 2.5.3)
- the defects of crystals (§ 2.5.4)

2.5.1 Finite-size effects in crystals

Periodic boundary conditions

A real finite crystal lacks the full translational symmetry. When the crystal size is much larger than the nearest-neighbours interatomic distances, so that the number of surface atoms is much lower than the number of bulk atoms, surface effects can be neglected.

In that case, to exploit the properties connected to the full translational invariance of an infinite crystal, it is convenient to introduce suitable boundary conditions at the surfaces of the real finite crystals. In general one resorts to the *Born-von Karman periodic boundary conditions*.
Let's consider a crystal of parallelepiped shape, whose edges are $N_1\vec{a}_1, N_2\vec{a}_2, N_3\vec{a}_3$ (where N_1, N_2, N_3 are very large numbers). This "fundamental" crystal is considered as a part of an infinite series of identical crystals extending in all directions with the same orientation.
Once a point is chosen in the fundamental crystal, any translation

$$\vec{T}_{N_1\vec{a}_1}\,, \quad \vec{T}_{N_2\vec{a}_2}\,, \quad \vec{T}_{N_3\vec{a}_3}\,,$$

say any translation of length equal to one of the edges of the fundamental crystal, leads to a point in one of the replied crystals where the physical properties are supposed to be exactly the same as in the original point of the fundamental crystal.
Otherwise stated, according to the periodic boundary conditions,

$$\vec{T}_{N_1\vec{a}_1} \;=\; \vec{T}_0\,, \quad \vec{T}_{N_2\vec{a}_2} \;=\; \vec{T}_0\,, \quad \vec{T}_{N_3\vec{a}_3} = \;\vec{T}_0\,, \tag{2.14}$$

where T_0 is the null translation.
In the following chapters, a number of applications of the periodic boundary conditions (PBC) can be found.

Shape function

In some applications, typically diffraction studies, it is convenient to account for the finite size of a real crystal by a *shape function* $M(\vec{r})$.
Example 1: For a spherical crystal of radius r_0,

$$M(\vec{r}) \;=\; \begin{cases} 1 & \text{if } |\vec{r}| \le r_0 \\ 0 & \text{if } |\vec{r}| > r_0 \end{cases} \tag{2.15}$$

Example 2: For a crystal with parallelepiped shape,

$$M(\vec{r}) \;=\; \begin{cases} 1 & \text{if } \begin{cases} x_1 \le x < x_2 \\ y_1 \le y < y_2 \\ z_1 \le z < z_2 \end{cases} \\ 0 & \text{elsewhere} \end{cases} \tag{2.16}$$

To take into account the finite size of a crystal, the lattice function $L(\vec{r})$ is multiplied by the shape function and the expression (2.12) of the structure function is modified into

$$\boxed{S(\vec{r}) \;=\; [L(\vec{r})\, M(\vec{r})] \,*\, f(\vec{r}) \;=\; \sum\nolimits_n f(\vec{r} - \vec{T}_n)} \tag{2.17}$$

where now the sum is over a finite number of displacement vectors $\vec{T}_n$.

Surfaces

In some cases, we can be interested in studying the surface properties of a crystal. The atomic structure at the surface of a crystal is generally different from the the structure in the bulk; correspondingly, different are the vibrational and electronic properties.
Atoms at or near the surface only experience inter-atomic forces with one component directed towards the interior of the crystal. As a result of this anisotropy, atoms near the surface assume relative positions with different spacing and/or symmetry with respect to the bulk atoms, creating a different surface structure. This change in equilibrium positions near the surface can be of two different types.

1. *Relaxation* refers to a change in the position of entire layers of atoms relative to the bulk positions. A typical effect is the contraction of the interatomic distances normal to the surface, leading to a smaller inter-planar spacing with respect to the bulk (Fig. 2.13, left). Some surfaces also experience relaxations in the lateral direction: the upper layers become shifted with respect to the bulk layers in order to minimize the positional energy (Fig. 2.13, centre).
2. *Reconstruction* refers to a change in the two-dimensional structure of the surface layers, in addition to relaxation effects (Fig. 2.13, right). The general symmetry of a layer may also change.

Figure 2.13: Schematic examples of relaxation and reconstruction of a crystal surface.

Heterostructures, interfaces

Alternate layers of two or more different substances (typically semiconductors) can be grown coherently on a substrate. Heterostructures offer flexible possibilities in the design and performance of semiconductor devices.

Example: Ge and GaAs have the same structure (diamond-zincblende) and very similar lattice constants a=5.646 Å and 5.653 Å, respectively. The two structures fit coherently to form parallel layers. The general interest of such kind of heterostructures stems from the different electronic properties (for example, the electronic band gaps are 0.76 eV and 1.43 eV for Ge and GaAs, respectively) and the possibility of engineering the electronic energy bands in many solid state device applications.

The structural and physical properties of crystals at interfaces are different from the properties within the bulk.

2.5.2 Atomic vibrations

In a real crystal, atoms are not frozen at the crystallographic positions; they are instead affected by vibrations around the crystallographic positions, which should then be considered as equilibrium positions. Only equilibrium positions form a periodic lattice.
The amplitude of atomic vibrations is finite even at zero temperature (zero point energy) and increases when temperature increases. Atomic vibrations are accounted for by the probability density $w(\vec{r})$ of finding the atomic nucleus at the position $\vec{r}$ with respect to the equilibrium position. The contribution of each atom to the basis function has to be convoluted with the probability density $w(\vec{r})$. This problem is considered in more detail in Chapter 11, dedicated to structural probes.

2.5.3 Nano-particles

In the last years, crystals whose size is of the order of a few nanometers, typically between 1 and 100 nm, have aroused much interest for their peculiar properties, which can vary dramatically with the particle size and are intermediate between the properties of molecules and of bulk matter. Nanoparticles properties are exploited for applications in many scientific and technological fields, such as medicine, optics, electronics.

An important parameter is the ratio between the number of surface atoms and the number of bulk atoms; for sufficiently small particles, the number of surface atoms can be of the same order of the number of bulk atoms (see § 2.7.4 for a calculation).

Problem: Verify that for a cubic crystal of copper of 1 μm size the ratio between the numbers of surface and bulk atoms is about 1/8000. Show that the ratio scales as $1/d$, where d is the linear size. (See § 2.7.4 for some hints).

The regular crystalline structure (when present) can be highly perturbed; the lattice parameters and the inter-atomic distances are generally different in the bulk and at the surface, so that they can be defined only in terms of distributions of values for a given nanoparticle.
One should add that nanoparticles are generally produced with a distribution of sizes and shapes, which introduces a further contribution to the width of the distributions of interatomic distances. When the size distribution of nanoparticles is narrow, one speaks of nanoclusters.

2.5.4 Defects in crystals

Defects of various types are always present in real crystals, and can be responsible for important physical properties.
Crystal defects can be classified, according to their dimensionality, as:

- Zero-dimensional defects (point defects) (Fig. 2.14): vacancies and interstitials; substitutional or interstitial doping; isotopic defects.
- One-dimensional defects, such as edge dislocations, which consist in the termination of a plane of atoms in the middle of a crystal; dislocations account for the plasticity and hardness properties of metals.
- Two-dimensional defects, such as grain boundaries, interfaces, stacking faults; also the surface of a finite crystal can be considered a two-dimensional defect.
- Three-dimensional defects, such as micro-cavities and inclusions of heterogeneous phases.

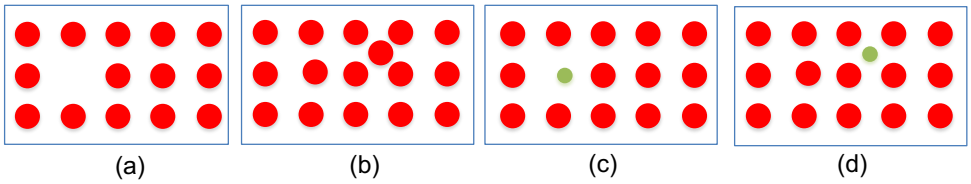

Figure 2.14: Zero-dimensional defects: (a) vacancy, (b) interstitial, (c) substitutional doping, (d) interstitial doping.

A different classification can be based on the origin of defects:

- defects due to the mechanism of crystal growth
- thermally activated defects
- defects created by impact of swift ions
- defects due to the stoichiometry, say to the deviation from the nominal composition

Thermally activated point defects

Let us focus the attention on two very common types of thermally activated point defects (Fig. 2.15):

a) Frenkel pair defects.
One atom (or ion) thermally excited is displaced from its regular crystal site to an interstitial site, leaving a hole at the crystal site.
Frenkel defects are more easily found in relatively open structures (diamond, zincblende) than in close-packed structures.

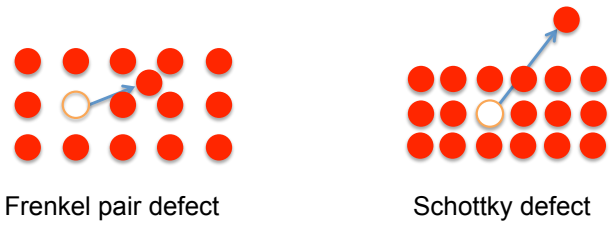

Figure 2.15: A Frenkel pair defect (left) and a Schottky defect (right).

b) Schottky defects.
One atom (or ion) thermally excited near the crystal surface exits from the surface (sublimation) leaving a hole at it previous the crystal site; the hole can migrate in the interior of the crystal.
Schottky defects are more easily found in close-packed structures.

Statistics of thermally-activated point defects

Let us consider a crystal in thermodynamic equilibrium at constant temperature and pressure. The concentration of thermally activated defects can be determined by minimising the Gibbs free energy

$$G = U + pV - TS \tag{2.18}$$

with respect to the number n of defects.
It is convenient to decompose the entropy S as the sum of two terms, a thermal one and a configurational one:

$$S = S_{\text{th}} + S_{\text{conf}} = k_B \ln \Omega_{\text{th}} + k_B \ln \Omega_{\text{conf}} , \tag{2.19}$$

where the configurational entropy S_{conf} is determined by the number Ω_{conf} of possible different modes by which n defects can be created.
Let us assume the following approximations:

- the volume V doesn't depend on the number n of defects,
- the thermal entropy S_{th} doesn't depend on the number n of defects.

The minimum of the Gibbs function corresponds thus to the condition

$$\frac{\partial G}{\partial n} = \frac{\partial U}{\partial n} - k_B T \frac{\partial \ln \Omega_{\text{conf}}}{\partial n} = 0 . \tag{2.20}$$

The first term in (2.20)

$$\frac{\partial U}{\partial n} = E_f \tag{2.21}$$

corresponds to the energy of formation of a single defect (typically of the order of the eV). For a low defects concentration, we can assume that E_f is independent of n.

For Schottky defects, the number of possible configurations is

$$\Omega_{\text{conf}}^{(S)} = \frac{N!}{(N-n)!\, n!} , \tag{2.22}$$

where N is the number of crystals sites where the defect can be created.

For Frenkel pair defects, the number of possible configurations is

$$\Omega_{\text{conf}}^{(F)} = \frac{N!}{(N-n)!\, n!} \frac{N'!}{(N'-n)!\, n!} , \tag{2.23}$$

where again N is the number of crystals sites where the defect can be created and N' is the number of available interstices.

Making use of the Stirling formula ($\ln x! = x \ln x - x$ for large x) and approximating $n \ll N, N'$, is it easy to verify that eq. (2.20) leads to

$$n^{(S)} = N \exp(-E_f/k_B T) \tag{2.24}$$

for Schottky defects and to

$$n^{(F)} = \sqrt{NN'} \exp(-E_f/2k_B T) \tag{2.25}$$

for Frenkel defects.

Problem: Assuming $N = N'$ for simplicity and $E_f = 3\,\text{eV}$, verify that the concentration n/N of Schottky and Frenkel defects is, respectively, 10^{-50} and 10^{-25} at 300 K and 10^{-15} and $10^{-7.5}$ at 1000 K. The concentration of Frenkel pair defects is larger than the concentration of Schottky defects because Ω_{conf} is larger for Frenkel defects than for Schottky defects.

2.6 Radial distributions

A different approach for describing the structure of atomic aggregates is based on the concept of radial distribution.
This approach, which for crystals is by far less powerful than the approach based on lattice periodicity, is the most viable for non-crystalline systems, such as liquids and glasses, which lack the periodic structure of crystals. For these systems, the positions of the atoms cannot be characterised in terms of Bravais lattices and unit cells, and the structural description is made in terms of one-dimensional statistical distributions.

2.6.1 Correlation functions

It is convenient to begin with a short introduction to correlation functions.

Let us consider a density function $\rho(\vec{r})$, such as

- the distributions of atomic positions in a condensed-matter system (possibly affected by vibrational disorder)
- the distribution of electronic charge density (measured for example by X-ray diffraction experiments).

The density-density equal-time correlation function is defined as

$$C(\vec{r}) = \langle \rho(\vec{r}\,') \, \rho(\vec{r}\,' + \vec{r}) \rangle_{\vec{r}\,'} \,, \tag{2.26}$$

where the average $\langle ... \rangle$ is performed over all possible values of $\vec{r}\,'$. The function $C(\vec{r})$ takes on large values when the argument $\vec{r}$ corresponds to a vector joining regions of large density ρ.

Note: The expression "equal time" means that we are considering the system at a fixed time, thus neglecting its possible time evolution. More refined correlation functions can evaluate the correlation between the density functions at different positions and at different times.

If we consider a system of N particles, whose positions $\vec{r}_n$ (where $1 \le n \le N$) are well defined, the density function becomes

$$\rho(\vec{r}) = \sum_{n=1}^{N} \delta(\vec{r} - \vec{r}_n) \tag{2.27}$$

and is different from zero only when $\vec{r}$ corresponds to one of the positions $\vec{r}_n$.
The density-density correlation function becomes

$$C(\vec{r}) = \langle \rho(\vec{r}_\kappa) \, \rho(\vec{r}_\kappa + \vec{r}) \rangle_\kappa \,, \tag{2.28}$$

where now the average $\langle \ldots \rangle_\kappa$ is performed over the discrete κ sites where the density is different from zero.
Substituting now (2.27) with $\vec{r} = \vec{r}_\kappa$ in (2.28) one obtains

$$C(\vec{r}) = \left\langle \sum_{m=1}^{N} \delta(\vec{r}_\kappa - \vec{r}_m) \sum_{n=1}^{N} \delta(\vec{r}_\kappa - \vec{r}_n + \vec{r}) \right\rangle_\kappa . \tag{2.29}$$

Defining $\vec{r}_{mn} = \vec{r}_n - \vec{r}_m$, the correlation function (2.29) becomes

$$C(\vec{r}) = \left\langle \sum_{n} \delta(\vec{r} - \vec{r}_{mn}) \right\rangle_m = \frac{1}{N} \sum_{m=1}^{N} \sum_{n=1}^{N} \delta(\vec{r} - \vec{r}_{mn}) . \tag{2.30}$$

The correlation function $C(\vec{r})$ is zero when $\vec{r} \neq \vec{r}_{mn}$. Otherwise stated, $C(\vec{r}) \neq 0$ when $\vec{r}$ corresponds to a vector distance between any two pair of particles.

Note: The sums in (2.30) include the term $m = n$, which corresponds to $r_{mn} = 0$, say to a singularity at the origin.

Normalised correlation function

Let $\rho_0 = N/V$ be the number density, say the number of particles per unit volume.
The normalised correlation function is defined as

$$g(\vec{r}) = \frac{C(\vec{r})}{\rho_0} = \frac{V}{N^2} \sum_{m} \sum_{n} \delta(\vec{r} - \vec{r}_{mn}) . \tag{2.31}$$

Correlation function for a crystal

For a perfect infinite crystal, the sums in (2.30) have no upper limit. If the crystal has only one atom per primitive cell, all atomic positions are equivalent, and (2.30) becomes

$$C(\vec{r}) = \sum_{n=1}^{\infty} \delta(\vec{r} - \vec{r}_n) . \tag{2.32}$$

The function $C(\vec{r})$ is different from zero only when $\vec{r}$ corresponds to a vectorial distance between any two atoms of the crystal.

(?) How should (2.32) be modified for a crystal with $s \neq 1$ atoms per primitive cell?
(?) Consider a two-dimensional crystal (e.g. a graphene sheet) and draw the corresponding correlation function for a finite number of low $\vec{r}$ values.

2.6.2 Radial distribution function

Let us now focus the attention on isotropic systems, such as liquids and non-crystalline solids.

In the absence of long-range translational symmetry, the directions of the interatomic distances $\vec{r}_{mn}$ are randomly distributed. The correlation function (2.30) is transformed into the radial distribution function (RDF) $R(r)$:

$$R(r) = \frac{1}{N} \sum_{m} \sum_{n \neq m} \delta(r - r_{mn}) , \tag{2.33}$$

where r_{mn} is the scalar distance between atom m and atom n and where the terms $m = n$ are eliminated from the sums.

The Radial Distribution Function $R(r)$ can ideally be built as follows (Fig. 2.16 left):

1. choose an atom m and calculate how many atoms are included in a spherical shell of radius r and thickness dr centred on m;

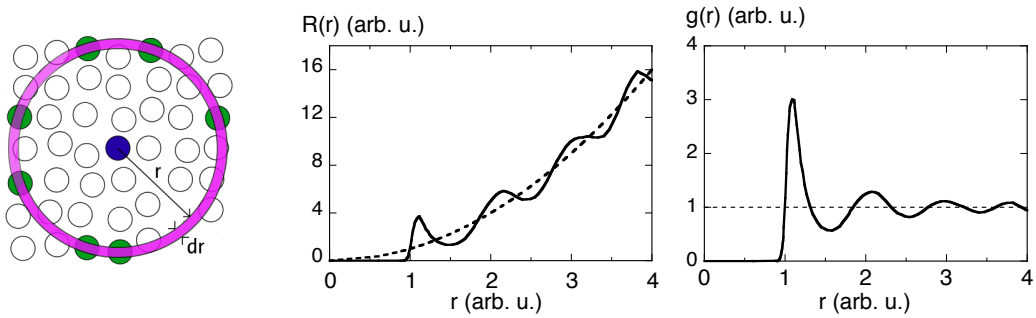

Figure 2.16: Left: how to build up the radial distribution function. Centre: example of radial distribution function $R(r)$ (continuous line), the dashed line is $4\pi\rho_0 r^2$. Right: corresponding pair distribution function $g(r)$.

2. repeat now the operation for any other atom of the system and calculate the average.

In view of structural disorder, one expects a large number of very closely spaced distances r_{mn}. The further effect of thermal disorder is to transform the RDF into a continuous distribution (Fig. 2.16 centre).
Alternatively, the RDF is defined by the following statement: the average number of atoms inside a spherical shell centred on a randomly chosen atom and included between the radii r_1 and r_2 is given by the integral of the RDF

$$n(r_1, r_2) \;=\; \int_{r_1}^{r_2} R(r)\, dr\,. \tag{2.34}$$

From (2.34) one can see that the RDF $R(r)$ has the dimension of an inverse length.

The RDF is zero for short distances (Fig. 2.16 centre), because there is a distance of minimum approach between nearest-neighbour atoms. The RDF increases dramatically, forming a peak in correspondence of the nearest-neighbour distance, which is generally quite well defined even in non-crystalline systems. When the distance r increases, the RDF becomes less and less structured; its behaviour becomes a progressively damped oscillation around the average value $4\pi r^2\rho_0$, where $\rho_0 = N/V$ is the average number density, say the number of atoms per unit volume.
Groups of atoms whose distances from the central atom form well distinguishable peaks in the RDF are grouped into coordination shells. The first coordination shell is generally well defined for all system; outer coordination shells are less and less evident.

For non monatomic systems, one can define different partial RDFs, corresponding to the different pairs of atomic species. The total RDF is the sum of the partial RDFs (Fig. 2.17).

In conclusion, the RDF only contains one-dimensional information, and in non-crystalline systems one can find well defined correlations, corresponding to peaks or bumps of the RDF, only for short-range and sometimes medium-range distances.

2.6.3 Pair distribution function

An alternative expression of the radial structure is obtained by considering the fluctuation $g(r)$ with respect to the smooth behaviour $4\pi r^2\rho_0$ for a homogenous system (ρ_0 is the average number density). The radial distribution function $R(r)$ can be expressed as

$$R(r) \;=\; 4\pi r^2\rho_0\, g(r)\,. \tag{2.35}$$

The function $g(r)$ is called pair distribution function (PDF). It is the quantity most frequently used when the structure of non-crystalline systems is measured by diffraction experiments.
The properties of the PDF are strictly connected to the properties of the RDF (Fig. 2.16, right):

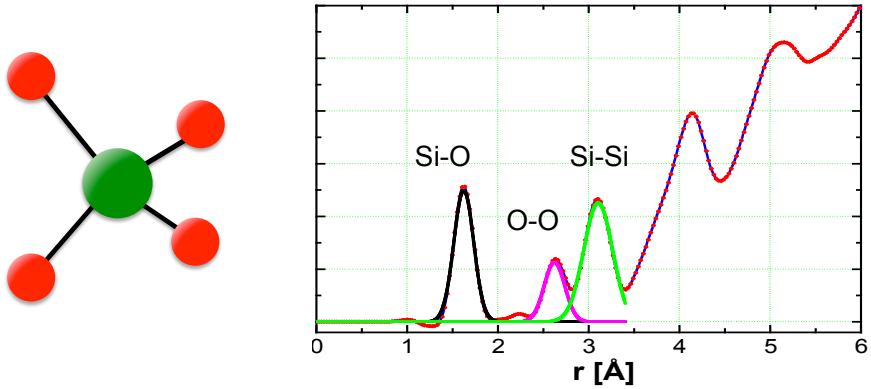

Figure 2.17: Structure of non-crystalline SiO_2. Left: each Si atoms is coordinated to four O atoms. Right: RDF determined from X-ray diffraction experiments: the partial RDFs due to the Si–O, O–O and Si–Si nearest neighbours are distinguished.

- for $r \to 0$, below the distance of closest approach, $g(r) = 0$
- at and above the distance of closest approach $g(r)$ oscillates around the value 1
- the oscillations damp out when r increases, and $g(r) \to 1$ for $r \to \infty$

The PDF $g(r)$ is an adimensional quantity.
From (2.33) and (2.35), one can express the pair distribution function $g(r)$ for discrete distributions as

$$g(r) \;=\; \frac{1}{4\pi r^2}\,\frac{V}{N^2}\sum_m \sum_{n\neq m} \delta(r - r_{mn})\,. \tag{2.36}$$

Note: Be careful, the acronym PDF has no unique meaning; elsewhere it can mean "probability distribution function".

2.6.4 Pair density function

The oscillations of the pair distribution function $g(r)$ can be expressed in terms of the number density oscillations by means of the pair density function $\rho(r)$

$$\rho(r) \;=\; \rho_0\, g(r)\,, \quad \text{so that} \quad R(r) \;=\; 4\pi r^2\, \rho(r)\,. \tag{2.37}$$

From (2.37) one can see that the pair density function $\rho(r)$ has the dimension of an inverse volume.
From (2.36), one can find

$$\rho(r) \;=\; \frac{1}{4\pi r^2}\,\frac{1}{N}\sum_{mn} \delta(r - r_{mn})\,. \tag{2.38}$$

The properties of the pair density function $\rho(r)$ are strictly connected to the properties of $g(r)$:

- for $r \to 0$, $\rho(r) = 0$,
- at and above the distance of closest approach $\rho(r)$ oscillates around the value ρ_0
- $\rho(r) \to \rho_0$ for $r \to \infty$

2.6.5 Radial functions and crystalline solids

The RDF $R(r)$ can be evaluated for crystalline solids too.
For an ideal crystal with atoms fixed at their equilibrium positions, the RDF would be made by delta-like contributions due to the different coordination shells. Actually, a real crystal is always affected by vibrational disorder, and the instantaneous atomic positions are spread around the

equilibrium positions. The RDF is thus characterised by peaks whose finite width increases with temperature (Fig. 2.18).
In general, the RDF is considered of limited interest for crystals, since it only gives one-dimensional information on the distribution of distances, and misses a great deal of the richness of information embedded in the crystallographic formalism.

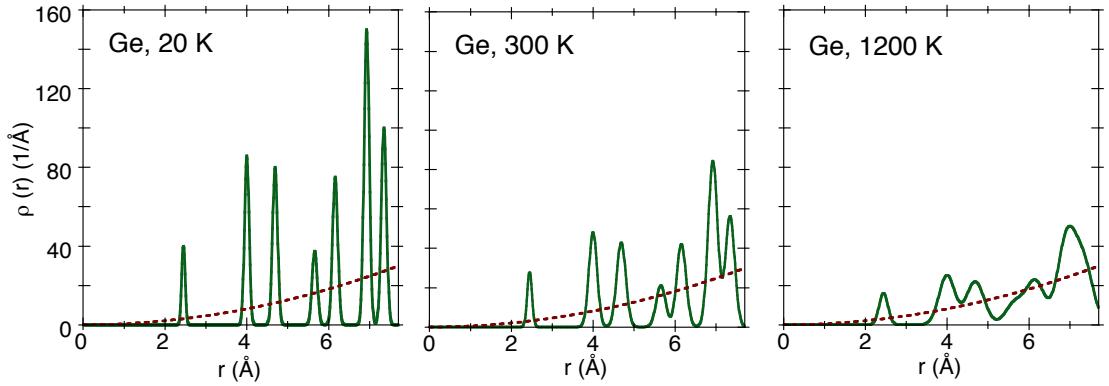

Figure 2.18: Radial Distributiuon Function of crystalline Germanium at different temperatures (20, 300 and 1200 K). The dashed lines is $4\pi r^2\rho_0$, where $\rho_0 = 4.42$ atoms/Å^3 is the number density.

However, in some cases the RDF contains original information, complementary to the crystallographic information. The basic difference between the RDF and the crystallographic picture resides in the fact that crystallography gives information on average atomic positions, while the RDF gives information on the relative inter-atomic distances. Due to vibrational disorder, the two types of information, average and local, are different and complementary. The difference between average and local properties can be enhanced by other sources of structural disorder.
Crystallographic information in crystals is generally obtained by analysing the Bragg peaks in diffraction experiments of X-rays, neutrons, electrons (see § 11.6). Local information on the RDF can be gained from EXAFS spectroscopy (see § 11.8) or from the quite difficult analysis of the thermal diffuse scattering in diffraction patterns.

Problem: Determine the coordination numbers and the interatomic distances for the first six coordination shells of copper. The structure is fcc, the lattice parameter of the cubic conventional cell is $a = 3.61491$ Å at 25°C [value from Acta Crystallographica A25, 676 (1969)].

2.7 Complements and problems

2.7.1 Primitive vectors for fcc and bcc Bravais lattices

We show here some relatively simple choices of primitive vectors for the fcc and bcc Bravais lattices (see Fig. 2.8).

fcc lattices

Let's take a corner of the conventional cubic cell as origin of the coordinate system. The simplest choice are primitive vectors whose tips are at the centre of the three cube faces that meet at the origin corner:

$$\vec{a}_1 = (a/2)\,(\hat{x} + \hat{y}) \tag{2.39}$$
$$\vec{a}_2 = (a/2)\,(\hat{y} + \hat{z}) \tag{2.40}$$
$$\vec{a}_3 = (a/2)\,(\hat{z} + \hat{x}) \tag{2.41}$$

The three primitive vectors share the same length $|\vec{a}_i| = a/\sqrt{2}$. The distance between the tips of any pair of primitive vectors is again $a/\sqrt{2}$. The angle between primitive vectors is 60°.
This primitive cell is rhombohedral.
The volume of the conventional unit cell is $V_c = a^3$.
The volume of the primitive cell is

$$V_p = (\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3 = a^3/4 = V_c/4\,. \tag{2.42}$$

Actually, the conventional cell contains 4 Bravais lattice points.

bcc lattices

Let's take a corner of the conventional cubic cell as origin of the coordinate system.

1.
We can choose two primitive vectors corresponding to two edges of the cubic conventional cell and the third vector connecting a corner to the centre of the cube:

$$\begin{array}{ll} \vec{a}_1 = a\,(1\hat{x} + 0\hat{y} + 0\hat{z}) & |\vec{a}_1| = a \\ \vec{a}_2 = a\,(0\hat{x} + 1\hat{y} + 0\hat{z}) & |\vec{a}_2| = a \\ \vec{a}_3 = (a/2)\,(1\hat{x} + 1\hat{y} + 1\hat{z}) & |\vec{a}_3| = \sqrt{3}a/2 \end{array} \tag{2.43}$$

The three primitive vectors have different lengths.

2.
If one prefers that the three primitive vectors share the same length, one can connect the origin with three of the eight nearest-neighbour points (centres of the eight cubes that meet at the origin):

$$\begin{array}{ll} \vec{a}_1 = (a/2)\,(-\hat{x} + \hat{y} + \hat{z}) & |\vec{a}_1| = \sqrt{3}a/2 \\ \vec{a}_2 = (a/2)\,(\hat{x} - \hat{y} + \hat{z}) & |\vec{a}_2| = \sqrt{3}a/2 \\ \vec{a}_3 = (a/2)\,(\hat{x} + \hat{y} + -\hat{z}) & |\vec{a}_3| = \sqrt{3}a/2 \end{array} \tag{2.44}$$

The corresponding primitive cell is rombohedral; the angle between the primitive vectors is 109.47° (as in tetrahedra).
The situation is better visualised if the origin is taken at the centre of the cubic conventional cell and the primitive vectors connect the origin to three of the eight corners.
The volume of the primitive cell is (for whichever choice of primitive vectors)

$$V_p = (\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3 = a^3/2 = V_c/2\,. \tag{2.45}$$

Actually, the conventional cell contains 2 Bravais lattice points.

2.7.2 Convolution and delta functions

The convolution of any two one-dimensional functions $f(x)$ and $g(x)$, symbolically $f(x) * g(x)$, is a third function $h(x)$ defined as

$$h(x) \;=\; f(\xi) * g(\xi) \;=\; \int_{-\infty}^{+\infty} f(\xi)\, g(x-\xi)\, d\xi \;=\; \int_{-\infty}^{+\infty} f(x-\xi)\, g(\xi)\, d\xi\,. \tag{2.46}$$

The convolution $h(x) = f * g$ is the integral of the product of the two functions $f(\xi)$ and $g(\xi)$ after one is reversed and shifted by x.

Problem: Calculate the convolution $h(x)$ of the two square functions

$$f(\xi) = \begin{cases} 0 & \text{if } \xi \le -2 \\ 1 & \text{if } -2 < \xi \le +2 \\ 0 & \text{if } \xi > +2 \end{cases} \qquad g(\xi) = \begin{cases} 0 & \text{if } \xi \le 9 \\ 1 & \text{if } 9 < \xi \le 11 \\ 0 & \text{if } \xi > 11 \end{cases} \tag{2.47}$$

Investigate the effect on the convolution $h(x)$ of: *a)* the progressive narrowing of the function $g(\xi)$, *b)* a shift of the centre of the function $f(\xi)$.

In one dimension, the delta function is implicitly defined by

$$\int_{-\infty}^{+\infty} dx\, f(x)\, \delta(x - x_0) \;=\; f(x_0)\,, \tag{2.48}$$

where $f(x)$ is any function. In three dimensions

$$\int_{-\infty}^{+\infty} dr^3\, f(\vec{r})\, \delta(\vec{r} - \vec{r}_0) \;=\; f(\vec{r}_0)\,. \tag{2.49}$$

Let us now identify the function $g(\xi)$ of (2.46) with the delta function, $g(\xi) = \delta(\xi - x_0)$, and calculate the convolution of the delta function with another function $f(\xi)$:

$$h(x) \;=\; f * \delta \;=\; \int_{-\infty}^{+\infty} f(x - \xi)\, \delta(\xi - x_0)\, d\xi \;=\; f(x - x_0)\,. \tag{2.50}$$

The effect of the convolution with the delta function is to shift the function $f(\xi)$ along the x axis. Suppose that the function $f(\xi)$ attains its maximum value in correspondence of the coordinate $\xi = x_1$. The shifted function h exhibits the maximum when $x - x_0 = x_1$, say for $x = x_1 + x_0$.

2.7.3 Interplanar distance in orthogonal lattices

Let us consider three orthogonal axes (x, y, z) whose directions correspond to the directions of the edges (a, b, c) of the unit cell.
A plane of Miller indices (h, k, ℓ) intersects the three edges at $(a/h, b/k, c/\ell)$, respectively.
The components of the unit vector normal to the plane are the directional cosines $(\cos\alpha, \cos\beta, \cos\gamma)$, which obey the relation

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1\,. \tag{2.51}$$

The distance of the plane (h, k, ℓ) from the origin, which corresponds to the interplanar distance, is connected to the directional cosines by

$$d_{hk\ell} = \frac{a}{h}\,\cos\alpha = \frac{b}{k}\,\cos\beta = \frac{c}{\ell}\,\cos\gamma\,. \tag{2.52}$$

Eq. (2.51) can be re-written as

$$\left(\frac{h d_{h\,k\ell}}{a}\right)^2 + \left(\frac{k\, d_{hk\ell}}{b}\right)^2 + \left(\frac{\ell\, d_{hk\ell}}{c}\right)^2 = 1\,, \tag{2.53}$$

so that

$$d^2_{hk\ell} = \frac{1}{(h/a)^2 + (k/b)^2 + (\ell/c)^2} \tag{2.54}$$

and finally

$$d_{hk\ell} = \frac{1}{\sqrt{(h/a)^2 + (k/b)^2 + (\ell/c)^2}}\,. \tag{2.55}$$

For cubic crystals, where $a = b = c$,

$$d_{hk\ell} = \frac{a}{\sqrt{h^2 + k^2 + \ell^2}}\,. \tag{2.56}$$

2.7.4 Ratio between surface and bulk atoms

To roughly evaluate the ratio between the number of surface and bulk atoms in a small crystal, let us take a model system, say a cubic crystal of linear size d, so that the volume is $V = d^3$, and let us consider atoms as small cubes of linear size ℓ and volume ℓ^3.
The total number of atoms in the cubic crystal is

$$N_{\text{tot}} = \frac{V}{\ell^3} = \frac{d^3}{\ell^3} \,. \tag{2.57}$$

To evaluate the number of surface atoms, let us consider a monolayer of thickness ℓ, corresponding to a volume $d^2\ell$ for each one of the cube faces. Taking into account that there are 6 faces and that the atoms along the edges shouldn't be counted twice, one finds

$$N_{\text{surf}} = 6\,\frac{d^2\ell}{\ell^3} - 12\frac{d\ell^2}{\ell^3} + 8\frac{\ell^3}{\ell^3} \,. \tag{2.58}$$

The number of bulk atoms is

$$N_{\text{bulk}} = N_{\text{tot}} - N_{\text{surf}} = \frac{d^3}{\ell^3} - 6\,\frac{d^2\ell}{\ell^3} + 12\frac{d\ell^2}{\ell^3} - 8\frac{\ell^3}{\ell^3} \,. \tag{2.59}$$

By introducing the ratio $x = \ell/d$, the ratio between the number of surface and bulk atoms can be expressed as

$$\frac{N_{\text{surf}}}{N_{\text{bulk}}} \;=\; \frac{6x - 12x^2 + 8x^3}{1 - 6x + 12x^2 - 8x^3} \tag{2.60}$$

(?) Plot the above function for realistic values of x.

One can easily verify that for $x \ll 1$, say for $d \gg \ell$,

$$\frac{N_{\text{surf}}}{N_{\text{bulk}}} \;\propto \frac{1}{d} \,. \tag{2.61}$$

2.8 Bibliography of Chapter 2

On crystalline structures

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions). Chapter 4.
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapter 1.

On the radial distribution functions and non-crystalline structures

- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapter 19.
- Y. Waseda: *The structure of non-crystalline materials: liquids and amorphous solids*, Mcgraw-Hill 1980. ISBN-13: 978-0070684263.
- T. Egami and S. J. L. Billinge: *Underneath the Bragg peaks*, Pergamon 2003 (introduction to radial distributions and to diffuse scattering).

Chapter 3

Reciprocal space

The possibility of describing a given phenomenon in two equivalent conjugate spaces connected by the Fourier transform is a powerful tool in many branches of Physics as well as for many technological applications. In Solid State Physics and in Crystallography, the two spaces are called direct (or real) space and reciprocal (or inverse or wavevector) space.
In these lectures, we are interested in the applications of the reciprocal space to diffraction experiments (Chapter 11), atomic vibrations in crystals (Chapter 7), motion of electrons in a periodic crystal potential (Chapter 9).

The conjugate of the Bravais lattice in the direct space is the reciprocal lattice in the reciprocal space. In this Chapter, $\vec{a}_1, \vec{a}_2, \vec{a}_3$ are the primitive vectors of the direct lattice and $\vec{b}_1, \vec{b}_2, \vec{b}_3$ are the primitive vectors of the reciprocal lattice. Crystallographers prefer instead the notation $\vec{a}, \vec{b}, \vec{c}$ for the primitive vectors of the direct lattice and $\vec{a}^*, \vec{b}^*, \vec{c}^*$ for the primitive vectors of the reciprocal lattice.

To introduce the basic concepts in the simplest form, we start in § 3.1 from the description of the reciprocal space in one dimension; the extension to the three-dimensional case is made in § 3.2. The fundamental relationships connecting the three-dimensional direct and reciprocal lattices are presented in § 3.3. The physical meaning of the reciprocal space is discussed in § 3.4, while § 3.5 contains mathematical complements.

3.1 Reciprocal space in one dimension

Let us first consider the connection between direct and reciprocal space in one dimension. (A tutorial introduction to some basic concepts of the Fourier Transform can be found in § 3.5).

3.1.1 Periodic functions. Fourier series. Reciprocal lattice.

A periodic function in the direct space $f(x) = f(x + na)$, where a is the period and n an integer number, can be expanded in a Fourier series as

$$f(x) = \sum_{m=-\infty}^{+\infty} A_m \, e^{imbx} = \sum_G A_G \, e^{iGx} \tag{3.1}$$

where $b = 2\pi/a$, the integer index m can be positive, zero or negative, and the quantities $G = mb$ form a lattice of equally spaced points in the one-dimensional reciprocal space, say a *reciprocal lattice*.
The coefficients $A_m = A_G$ are generally complex quantities, defined by:

$$A_m = \frac{1}{a} \int_{-a/2}^{+a/2} dx \, f(x) \, e^{-imbx} \, . \tag{3.2}$$

The coefficient A_0 is the average value of the function $f(x)$. The coefficients A_m with low $|m|$ values give the broader features of $f(x)$. The larger is the number of A_m coefficients considered (the higher is $|m|$) the better is the description of the fine details of the function $f(x)$.

Example: The periodic electrostatic potential created by a linear array of positive ions can be expanded in Fourier series. The main effects of the potential on a negative charge moving along the linear array are given by the low-order terms of the expansion.

Let us now consider some properties of the Fourier coefficients.

a) If the function $f(x)$ *is real* (e.g. an electron density), one can easily verify from (3.2) that $A_{-m} = A_m^*$, so that $|A_{-m}|^2 = |A_m|^2$. In diffraction experiments, one measures the squared moduli of the Fourier coefficients $|A_m|^2$, so that diffraction patterns are centro-symmetric and no information on the phase of the coefficients A_m can be directly obtained.

b) If the real function $f(x)$ *is symmetrical*, $f(-x) = f(x)$, one can easily verify from (3.2) that the Fourier coefficients A_m are real, and (3.1) becomes $f(x) = \sum_m A_m \cos(mbx)$. Diffraction experiments measure $|A_m|^2$, and the sign of A_m cannot be determined.

3.1.2 Non-periodic functions. Fourier integral.

For non-periodic functions, the Fourier series (3.1) transforms into the Fourier integral:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{f}(k)\, e^{ikx}\, dk\,, \tag{3.3}$$

where $\tilde{f}(k)$ is a function of a continuous variable k in the reciprocal space,

$$\tilde{f}(k) = \int_{-\infty}^{+\infty} f(x)\, e^{-ikx}\, dx\,. \tag{3.4}$$

As for the Fourier series, also for the Fourier integral the following properties hold:

a) If the function $f(x)$ is real, one can easily verify that $\tilde{f}(-k) = \tilde{f}(k)^*$.

b) If the real function $f(x)$ is symmetrical, $f(-x) = f(x)$, one can easily verify that the function $\tilde{f}(k)$ is real.

Note on the 2π factor

The two expressions (3.3) and (3.4) of the Fourier transform and back transform are unsymmetrical, the factor $1/\sqrt{2\pi}$ appearing only in the first one.
In Crystallography one generally prefers to separate the factor 2π in the expression of the wavevector $k = 2\pi/\lambda$ and consider the reciprocal space variable $x^* = 1/\lambda = k/2\pi$. By this choice, the Fourier integrals become symmetrical, since dk of (3.3) becomes $2\pi\, dx^*$:

$$\tilde{f}(x^*) = \int_{-\infty}^{+\infty} f(x)\, e^{i2\pi x x^*}\, dx \quad , \quad f(x) = \int_{-\infty}^{+\infty} \tilde{f}(x^*)\, e^{-i2\pi x x^*}\, dx^*\,.$$

In Physics, it is customary to use the wavevector $k = 2\pi/\lambda$. This choice introduces the factor 2π, which is sometimes symmetrically attributed to direct and inverse transforms as $\sqrt{2\pi}$.
The non-symmetric choice of (3.3) and (3.4) gives to the $x \to k$ transform (3.4) the same expression as the $x \to x^*$ transform of crystallographers, which is useful for the interpretation of diffraction patterns.

Example: PDF in non-crystalline materials

A one-dimensional non-periodic function is the pair distribution function $g(r)$ introduced in § 2.6 to describe the structure of non-crystalline materials. The PDF $g(r)$ typically oscillates around the value 1, the amplitude of oscillations being damped when r increases.
Due to the isotropy of non-crystalline materials, one can consider $g(r)$ as a symmetric function, $g(r) = g(-r)$. The Fourier transform of the PDF $g(r)$ is an oscillating real function in the reciprocal space, which can be sampled by X-ray or neutron scattering experiments (see § 11.6).

3.1.3 Convolution theorem

The convolution theorem states that:

> The Fourier transform of the convolution of two functions is the product of the Fourier transforms of the two functions.

Let $\mathcal{F}$ represent the Fourier transform operator. The convolution theorem can be analytically stated as

$$\mathcal{F}[f(x) * g(x)] \;=\; \mathcal{F}[f(x)] \cdot \mathcal{F}[g(x)]\,. \tag{3.5}$$

where the symbol $\star$ represents the convolution.

3.1.4 Application: Periodic crystal functions

According to equation 2.12 of § 2.4, the periodic function $S(x)$ which describes a one-dimensional crystalline structure can be expressed as the convolution of a periodic lattice function $L(x) = \sum_T \delta(x - T)$ and a non-periodic basis function $f(x)$:

$$S(x) \;=\; L(x) \,*\, f(x)\,. \tag{3.6}$$

According to (3.5), the Fourier transform of the periodic function $S(x)$ is the product of the separate Fourier transforms of the periodic lattice function and of the non periodic basis function:

$$\mathcal{F}[L(x) * f(x)] \;=\; \mathcal{F}[L(x)] \cdot \mathcal{F}[f(x)] \tag{3.7}$$

Fourier transform of the lattice function

Let us first consider the Fourier transform of a single delta function centred on $x = a$, which is

$$\mathcal{F}[\delta(x-a)] = \int_{-\infty}^{+\infty} \delta(x-a)\, e^{-ikx}\, dx = e^{-ika}\,, \tag{3.8}$$

say a complex function of the variable k, with modulus $|e^{-ika}| = 1$. When $a = 0$, the Fourier transform collapses to a constant real value, $\mathcal{F}[\delta(x-a)] = 1$.

The lattice function of a finite lattice of $N = 2p+1$ points (a "Dirac comb") is the sum of N delta functions

$$L(x) = \sum_{n=-p}^{p} \delta(x - na)\,, \tag{3.9}$$

which has been here chosen to be symmetric, $L(x) = L(-x)$.
The Fourier transform of the lattice function $L(x)$ is

$$\tilde{L}(k) = \mathcal{F}\left[\sum_{n=-p}^{p} \delta(x-na)\right] = \int_{-\infty}^{+\infty} \left[\sum_{n=-p}^{p} \delta(x-na)\right] e^{-ikx}\, dx = \sum_{n=-p}^{p} e^{-ikna}\,. \tag{3.10}$$

Since the lattice function $L(x)$ of equation (3.9) is symmetric, one expects the Fourier transform be real. Actually, one can show that

$$\tilde{L}(k) = \sum_{n=-p}^{p} e^{-ikna} = \frac{\sin(Nka/2)}{\sin(ka/2)} \tag{3.11}$$

which is a function with principal maxima at $k = 2m\pi/a$ separated by ripples of decreasing intensity as N increases. The demonstration of (3.11) is given in § 3.5.4, page 51. A practical example is given in Fig. 3.1 (top).

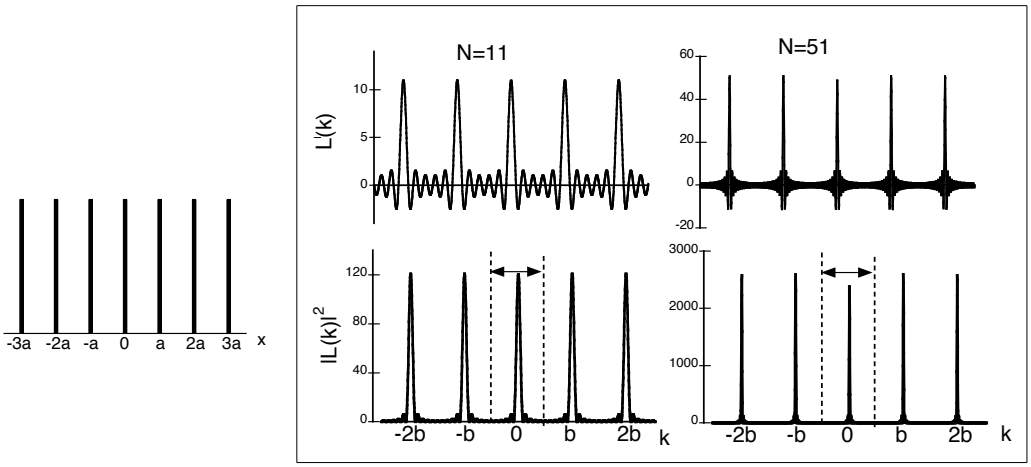

Figure 3.1: Left: a Dirac comb $L(x)$ with $N = 7$. The right panel shows the Fourier transforms $\tilde{L}(k)$ for $N = 11$ and $N = 51$ (top figures); the bottom figures show the squared values $|\tilde{L}(k)|^2$. The period of the reciprocal lattice is $b = 2\pi/a$; the vertical dashed lines limit the first Brillouin Zone.

In diffraction experiments one measures the intensity of the scattered radiation in correspondence of the reciprocal lattice points; the intensity is proportional to the squared modulus of $\tilde{L}(k)$; the function

$$|\tilde{L}(k)|^2 = \frac{\sin^2(Nka/2)}{\sin^2(ka/2)} \tag{3.12}$$

is often called Laue function (bottom of Fig. 3.1).

For the limiting case of an infinite crystal, $N \to \infty$, the Fourier transform of the lattice function is

$$\tilde{L}(k) = \lim_{N\to\infty} \frac{\sin(Nka/2)}{\sin(ka/2)} = \frac{1}{a} \sum \delta(k - 2\pi m/a) \tag{3.13}$$

The Fourier transform of an infinite Bravais lattice in the real space is an infinite Bravais lattice in the reciprocal space, of period $b = 2\pi/a$. The points of the reciprocal lattice are $G = mb$. The spacing b between the points of the reciprocal lattice is inversely proportional to the spacing a between the points of the direct space.

Fourier transform of the structure function

Let $\mathcal{F}[f(x)] = \tilde{f}(k)$ be the Fourier transform of the (non periodic) basis function $f(x)$. In general, $\tilde{f}(k)$ is a continuous function, significantly different from zero in a large interval of the variable k.

According to the convolution expression (3.7), the Fourier transform of the periodic function $S(x) = L(x) \star f(x)$ for an infinite lattice is different from zero only for the k values corresponding to points of the reciprocal lattice. For these points, $\mathcal{F}[S(x)] = \mathcal{F}[f(x)]$.
Otherwise stated, the function $\tilde{f}(k)$ can be sampled (e.g. in diffraction experiments) only in correspondence of the points of the reciprocal lattice.

3.1.5 Example: lattice plus basis

Let us consider a periodic function $S(x) = S(x + na)$ in the real space, defined as:

$$S(x) = \sum_n \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x - na)^2}{2\sigma^2}\right] . \tag{3.14}$$

The function $S(x)$ is the convolution of a lattice function $L(x)$ of period a and a gaussian function $f(x)$ of average position zero and standard deviation σ. Let us assume that $\sigma < a/2$ (Fig. 3.2, left). The function $S(x)$ could describe the spread of nuclear positions in a linear array of atoms in thermal motion (in harmonic approximation, the distribution function is gaussian). The function $S(x)$ could also approximate the distribution of electronic charge around the nuclei (which is in principle non-gaussian).

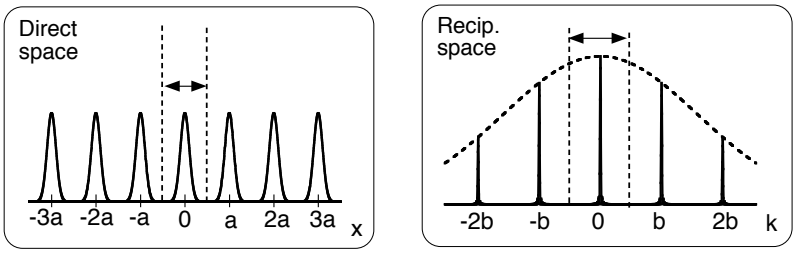

Figure 3.2: Left: unidimensional crystal with lattice parameter a and gaussian basis function. Right: squared modulus of its Fourier transform. The vertical lines define one Wigner-Seitz cell in the direct space (left) and the first Brillouin cell in the reciprocal space (right).

The Fourier transform $\tilde{L}(k)$ of the lattice function $L(x)$ is again a lattice function; the period $b = 2\pi/a$ of the reciprocal lattice is inversely proportional to the period a of the direct lattice.
The Fourier transform of the gaussian basis function $f(x)$ is a gaussian function $\tilde{f}(k) = \exp(-k^2\sigma^2/2)$, whose standard deviation is $\propto 1/\sigma^2$; the more peaked is the basis function $f(x)$, the larger is the Fourier transform $\tilde{f}(k)$.
The Fourier transform of the entire function $S(x)$, calculated making use of the convolution theorem, is the product $\tilde{S}(k) = \tilde{L}(k)\ \tilde{f}(k)$. The squared modulus of $\tilde{f}(k)$ is shown as a dashed line in Fig 3.2, right; the squared modulus of $\tilde{S}(k)$ is given by the values of the squared modulus of $\tilde{f}(k)$ in correspondence of the peaks of $\tilde{L}(k)$. Otherwise stated, the amplitude of the $\tilde{L}(k)$ function is modulated by the gaussian function $\tilde{f}(k)$.

(?) Determine the Wigner-Seitz cells in direct space (one is limited by the vertical lines in Fig. 3.2, left). Compare the basis function $f(x)$ within the Wigner-Seitz cells of different lattice points of direct space.

(?) How is the reciprocal lattice modified if the period a of the direct lattice is increased?

(?) How are the Fourier coefficients modified if the width σ of the gaussian function $f(x)$ is reduced? What happens to the reciprocal lattice if $\sigma \to 0$, say if the periodic function $S(x)$ becomes a pure lattice function? and if $\sigma \to \infty$?

Note: While the basis function $f(x)$ is the same for the Wigner-Seitz cell of every lattice point of the direct space, the Fourier-transformed function $\tilde{f}(k)$ is different for different Wigner-Seitz cells of the reciprocal lattice (Fig. 3.2). The Wigner Seitz cell centred on the origin of the reciprocal lattice is called "first Brillouin Zone". For a further discussion, see Section 3.4.

3.1.6 Effects of the finite crystal size

The finite size of the crystal has been accounted for by a suitable shape function in § 2.5. Adapting (2.17) to the one-dimensional case, the crystal structure becomes

$$S(x) \ = \ [L(x)\, M(x)] \ * \ f(x)\,. \tag{3.15}$$

According to the convolution theorem, the product of the lattice and shape functions is Fourier transformed into their convolution:

$$\mathcal{F}[L(x)M(x)] \ = \ \mathcal{F}[L(x)] * \mathcal{F}[M(x)]\,. \tag{3.16}$$

The effect is an enlargement of the peaks of the lattice function in reciprocal space.

3.2 Reciprocal space in three dimensions

The one-dimensional theory is easily generalised to the three-dimensional case.

Periodic functions, Fourier series

Let us consider a three-dimensional periodic function in the direct space (such as an electronic density or a potential energy)

$$f(\vec{r}+\vec{T}) = f(\vec{r})\,, \tag{3.17}$$

where

$$\vec{T} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \tag{3.18}$$

is the generic translation vector.

The three-dimensional periodic function can be expanded in a Fourier series, by analogy with the one-dimensional case. The one-dimensional quantities $G = mb$ become three-dimensional vectors

$$\vec{G} = m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3\,, \tag{3.19}$$

where $\vec{b}_1, \vec{b}_2, \vec{b}_3$ are primitive vectors of the reciprocal space and m_i are integer numbers.
The Fourier expansion is a sum over all the reciprocal lattice vectors $\vec{G}$

$$f(\vec{r}) = \sum_{\vec{G}} B_{\vec{G}}\, e^{i\vec{G}\cdot\vec{r}}\,. \tag{3.20}$$

The quantities $B_{\vec{G}}$ are the (generally complex) Fourier coefficients of the three-dimensional periodic function, which share some properties with the one-dimensional Fourier coefficients A_G of (3.2):

a) If $f(\vec{r})$ *is real*, then $B_{-\vec{G}} = B^*_{\vec{G}}$, so that $|B_{-\vec{G}}|^2 = |B_{\vec{G}}|^2$. In diffraction experiments, one measures the squared moduli of the Fourier coefficients $|B_{\vec{G}}|^2$: diffraction patterns are centrosymmetric and no information on the phase of the coefficients $B_{\vec{G}}$ can be directly obtained.
b) If the real function $f(\vec{r})$ *is symmetrical*, $f(-\vec{r}) = f(\vec{r})$, the Fourier coefficients $B_{\vec{G}}$ are real. Again, diffraction experiments measure $|B_{\vec{G}}|^2$, and the sign of $B_{\vec{G}}$ cannot be determined.

Non-periodic functions, Fourier integrals

For non-periodic functions, the Fourier integrals are defined in the three-dimensional space. To the function $f(\vec{r})$ of the continuous variable $\vec{r}$ in the direct space it corresponds a function $\tilde{f}(\vec{k})$ of the continuous variable $\vec{k}$ in the reciprocal space. The two functions are connected by the Fourier integrals:

$$\tilde{f}(\vec{k}) \;=\; \int f(\vec{r})\, e^{-i\vec{k}\cdot\vec{r}}\, d^3r \tag{3.21}$$

and

$$f(\vec{r}) \;=\; \frac{1}{(2\pi)^3}\int \tilde{f}(\vec{k})\, e^{i\vec{k}\cdot\vec{r}}\, d^3k\,. \tag{3.22}$$

Convolution theorem and crystal structure

Also in three dimensions, the convolution theorem holds.
Let us consider the periodic crystal structure $S(\vec{r})$ as the convolution of the periodic lattice function $L(\vec{r})$ and a non-periodic basis function $f(\vec{r})$:

$$S(\vec{r}) \;=\; L(\vec{r}) * f(\vec{r}) \tag{3.23}$$

The Fourier transform of the periodic crystal structure is

$$\tilde{S}(\vec{k}) \;=\; \tilde{L}(\vec{k}) \cdot \tilde{f}(\vec{k})\,. \tag{3.24}$$

For an infinite crystals, the Fourier transform $\tilde{L}(\vec{k})$ of the periodic lattice function $L(\vec{r})$ is different from zero only in correspondence of the tips of the reciprocal lattice vectors $\vec{G}$:

$$\tilde{L}(\vec{k}) = \frac{1}{V} \sum\nolimits_{\vec{G}} \delta(\vec{k} - \vec{G}) \tag{3.25}$$

As for the one-dimensional case, the Fourier transform of the periodic function $S(\vec{r})$ is different from zero only for the $\vec{k}$ values corresponding to points $\vec{G}$ of the reciprocal lattice.

Effects of the finite crystal size

By introducing the shape function $M(\vec{r})$, the crystal structure has been expressed in (2.17) as

$$S(\vec{r}) \;=\; [L(\vec{r})\, M(\vec{r})] \,*\, f(\vec{r})\,. \tag{3.26}$$

According to the convolution theorem, the product of the lattice and shape functions is Fourier transformed into their convolution:

$$\mathcal{F}[L(\vec{r})M(\vec{r})] \;=\; \mathcal{F}[L(\vec{r})] * \mathcal{F}[M(\vec{r})]\,. \tag{3.27}$$

The effect is an enlargement of the peaks of the lattice function.

3.3 Relation between direct and reciprocal lattices

Our problem is now to understand how the primitive vectors of the reciprocal lattice are defined and connected to the primitive vectors of the direct for three-dimensional structures.

Primitive vectors: orthogonal lattices

If the primitive vectors of the direct space are mutually orthogonal, the extension of the one-dimensional case to the three-dimensional case is trivial.
For each direction,

$$\vec{b}_i || \vec{a}_i\,, \qquad \vec{b}_i \perp \vec{a}_j \;\; (j \neq i)\,, \qquad b_i = \frac{2\pi}{a_i}\,, \tag{3.28}$$

so that

$$\vec{a}_i \cdot \vec{b}_j = 2\pi\delta_{ij}\,. \tag{3.29}$$

Fig. 3.3 shows a two-dimensional case. The primitive vectors of the reciprocal space $\vec{b}_1$ and $\vec{b}_2$ are orthogonal to the planes (10) and (01), respectively, and their lengths are inversely proportional to the corresponding inter-planar distances, $2\pi/a_1$ and $2\pi/a_2$, respectively.

One can easily verify that the condition of periodicity in the real space $f(\vec{r}+\vec{T}) = f(\vec{r})$ is fulfilled by the Fourier expansion

$$f(\vec{r}+\vec{T}) = \sum\nolimits_{\vec{G}} B_{\vec{G}}\, e^{i\vec{G}\cdot(\vec{r}+\vec{T})} = \sum\nolimits_{\vec{G}} B_{\vec{G}}\, e^{i\vec{G}\cdot\vec{r}}\, e^{i\vec{G}\cdot\vec{T}} = f(\vec{r})\,, \tag{3.30}$$

because the dot product of any pair of direct and reciprocal lattice vectors

$$\vec{T}\cdot\vec{G} = (n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3)\cdot(m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3) \tag{3.31}$$

is an integer multiple of 2π, so that

$$e^{i\vec{G}\cdot\vec{T}} = 1\,. \tag{3.32}$$

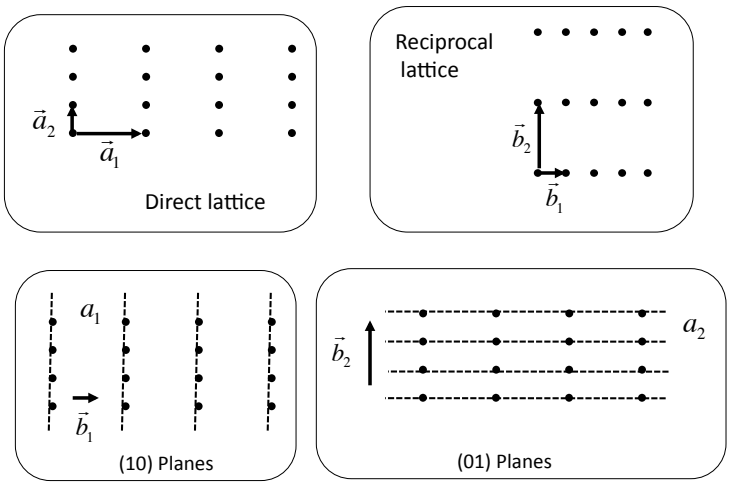

Figure 3.3: Top panels: an orthogonal direct lattice (left) and the corresponding reciprocal lattice. Bottom panels: relation between primitive vectors of reciprocal lattice and direct lattice planes.

Primitive vectors: two-dimensional oblique lattices

To better grasp the relation between direct and reciprocal non-orthogonal lattices, it is convenient to examine first the two dimensional case (Fig 3.4).
As for the orthogonal case, the reciprocal lattice primitive vectors $\vec{b}_1$ and $\vec{b}_2$ are perpendicular to the (10) and (01) planes, respectively, say they are oriented along the [10] and [01] directions, respectively. Their directions are thus different from the directions of the primitive vectors $\vec{a}_1$ and $\vec{a}_2$ of the direct lattice.
The moduli of the reciprocal primitive vectors $\vec{b}_1$ and $\vec{b}_2$ are again inversely proportional to the inter-planar spacing, but now the inter-planar spacing doesn't coincide with the length of the direct primitive vectors, as for the orthogonal lattices.
If θ is the angle between $\vec{a}_1$ and $\vec{a}_2$, the relations between the moduli of primitive vectors are

$$b_1 = 2\pi \frac{1}{a_1 \sin\theta}\,, \qquad b_2 = 2\pi \frac{1}{a_2 \sin\theta}\,. \tag{3.33}$$

When $\theta = \pi/2$, (3.33) reduces to (3.28).
Introducing the area of the primitive cell $a_1 a_2 \sin\theta$ at the denominator, we obtain an alternative expression which facilitates the transition from the two-dimensional to the three-dimensional case:

$$b_1 = 2\pi \frac{a_2}{a_1\, a_2 \sin\theta}\,, \qquad b_2 = 2\pi \frac{a_1}{a_1\, a_2 \sin\theta}\,. \tag{3.34}$$

Taking into account that the angle between $\vec{a}_i$ and $\vec{b}_i$ is $\pi/2 - \theta$, it is easy to see that the primitive vectors obey the orthogonality relation (3.29) found for orthogonal lattices

$$\vec{a}_i \cdot \vec{b}_j = 2\pi\delta_{ij}\,. \tag{3.35}$$

Primitive vectors: general rule for three-dimensional lattices

For the three-dimensional case, the general relation between direct and reciprocal primitive vectors, valid for both orthogonal and oblique lattices, is:

$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}\,; \quad \vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}\,; \quad \vec{b}_3 = 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}\,. \tag{3.36}$$

From (3.36) one can see that:

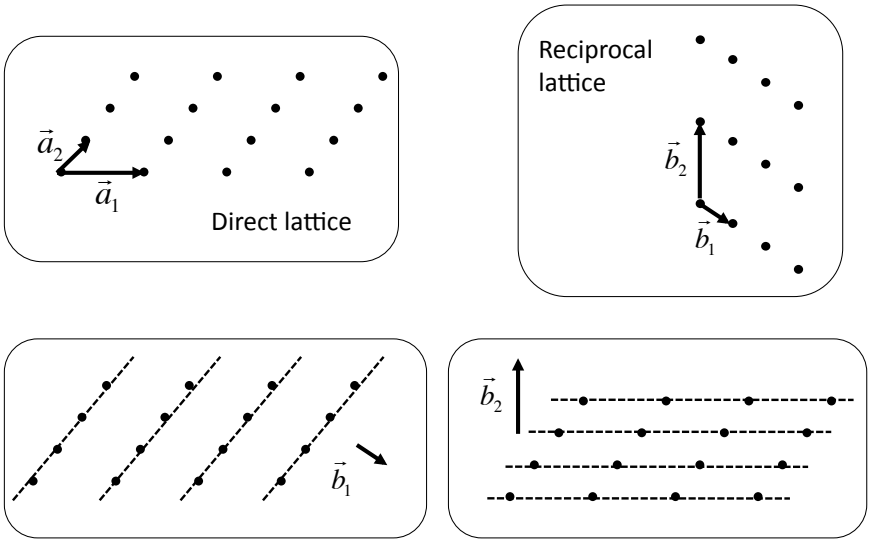

Figure 3.4: Top panels: an oblique direct lattice (left) and the corresponding reciprocal lattice. Bottom panels: the primitive vectors $\vec{b}_1$ and $\vec{b}_2$ are orthogonal to the planes (10) and (01), respectively, and their lengths are inversely proportional to the corresponding inter-planar distances.

- The denominator in each equation is the volume of the direct space cell. The larger the unit cell of the direct space, the smaller is the unit cell of the reciprocal space.
- The primitive vectors obey the orthogonality relation (3.29) found for orthogonal lattices

$$\vec{a}_i \cdot \vec{b}_j = 2\pi\delta_{ij} \,. \tag{3.37}$$

- The primitive vector $\vec{b}_1$ is perpendicular to the (100) planes, and its modulus is proportional to the inverse of the ratio between the volume of the direct primitive cell and the surface of the face defined by $\vec{a}_2$ and $\vec{a}_3$, say is proportional to the interplanar spacing (more exactly, 2π divided by the inter-planar spacing). Similar considerations hold for $\vec{b}_2$ and $\vec{b}_3$.
- The general relations (3.36) collapse to (3.28) for orthogonal lattices.

One can easily verify that the condition of periodicity in real (direct) space $f(\vec{r}+\vec{T}) = f(\vec{r})$ is fulfilled by the Fourier expansion, since the general definition again leads to

$$e^{i\vec{G}\cdot\vec{T}} = 1 \tag{3.38}$$

also for oblique lattices.
The connection between primitive vectors $\vec{a}_i$ and $\vec{b}_i$ ensures that the two lattices are rigidly connected: a rotation of the direct lattice corresponds to a similar rotation of the reciprocal lattice.

Problem: Verify that the reciprocal lattice of a fcc lattice is bcc (see § 3.5.5). Verify that the reciprocal lattice of a bcc lattice is fcc.

3.4 Physical interpretation of the reciprocal space

In the previous sections we have studied the properties of the reciprocal space in connection with the structure of crystalline matter. To that purpose, we have focussed our attention on the reciprocal lattice, say on a discrete grid of points of the reciprocal space.
Actually, the reciprocal space consists in a three-dimensional continuous distribution of points, corresponding to vectors $\vec{k}$. Every wavevector $\vec{k}$ can be connected to a plane wave in real space,

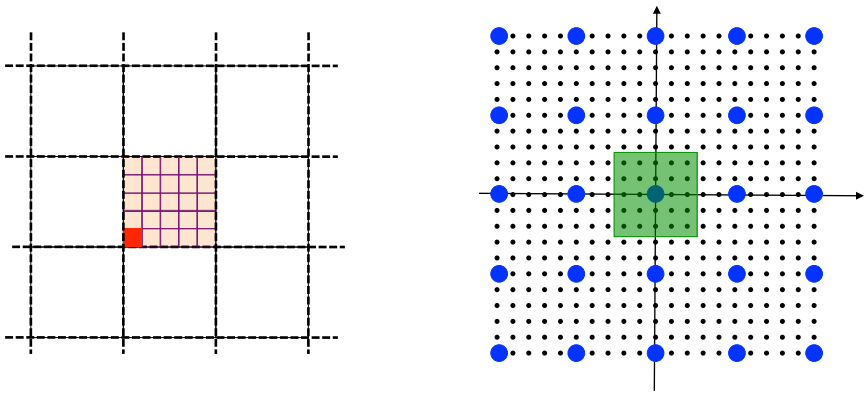

Figure 3.5: Schematic two-dimensional representation of direct and reciprocal spaces for a square lattice. *Left:* Direct space: the primitive cell (red) is the building block of the real finite crystal (central pink square) whose repetition according to the periodic boundary conditions simulates the infinite crystal. *Right:* Reciprocal space: grid of reciprocal space points (small black circles), reciprocal lattice points (large blue circles) and first Brillouin Zone (green square).

whose wave-fronts are normal to $\vec{k}$ and whose wavelength is $\lambda = 2\pi/k$. The reciprocal lattice vectors $\vec{G}$ represents a discrete subset of the reciprocal vectors $\vec{k}$.
The entire reciprocal space, not limited to the reciprocal lattice, is a powerful tool for studying important phenomena, such as the atomic vibrational modes in crystals (see Chapter 7) and the electron wavefunctions in crystals (see Chapter 9).
In this section we introduce some basics concepts which are extensively used in those later chapters.

The discussion is illustrated by the two-dimensional schemes of Fig. 3.5. In the left part of the figure, representing the direct space of a two-dimensional crystal, one can distinguish three levels of scale: the primitive cell (the little red square), the entire real crystal (the central pink square), also referred to as "fundamental" crystal, and the ideal infinite crystal made up by the repetition of the fundamental crystal according to the rules of the periodic boundary conditions introduced in § 2.5.1.

Reciprocal lattice and lattice planes

Let us first recall some basic concepts concerning the structure of crystals, as depicted in previous sections.
According to (3.20), the Fourier expansion of the periodic function $f(\vec{r})$ only involves the points of the reciprocal lattice (large blue circles in Fig. 3.5, right).
The reciprocal lattice vectors $\vec{G}$ are linear combinations of the primitive vectors $\vec{b}_1, \vec{b}_2, \vec{b}_3$ with integer coefficients, according to (3.19):

$$\vec{G} = m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3 \,.$$

A reciprocal lattice point $\vec{G}$ can be connected to a plane wave whose periodicity is related to the periodicity of the direct lattice in the real space.

a) For every family of lattice planes in the direct space, separated by a distance d, there are reciprocal lattice vectors perpendicular to the planes, the shortest of which have a length $2\pi/d$.
b) For every reciprocal lattice vector $\vec{G}$, there is a family of lattice planes in the direct space, normal to $\vec{G}$ and separated by a distance d, where $2\pi/d$ is the length of the shortest reciprocal lattice vector parallel to $\vec{G}$.

It is important to stress here that the wavevectors of the reciprocal lattice correspond to wavelengths of the same order or smaller than the size of the primitive cell.

Note: In Chapter 1 we have introduced two alternative notations for direct-space translation vectors:

$$\vec{T} \;=\; n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \;=\; u\vec{a} + v\vec{b} + w\vec{c}. \tag{3.39}$$

The corresponding alternative notations for the reciprocal-space translation vectors are:

$$\vec{G} \;=\; m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3 \;=\; h\vec{a}^* + k\vec{b}^* + \ell\vec{c}^* \,. \tag{3.40}$$

The second expression evidences the fact that the integer coordinates of the reciprocal lattice points correspond to the Miller indices $(hk\ell)$ of the direct lattice planes.

Periodic boundary conditions and reciprocal space

When treating the normal modes of atomic vibrations, we are dealing with wavelengths larger than the primitive cell size. When treating the electron wavefunctions in crystals, we are dealing with wavefunctions which can be larger or smaller than the primitive cell size. In both cases, one has to consider the entire reciprocal space, or at least a portion of it, not limited to the discrete lattice points.

As it will be seen later on, the treatment of atomic vibrations and electron dynamics is greatly simplified if the translational symmetry of the crystal is considered. However, the full translational symmetry is only possible for infinite crystals. Real crystals are finite, and thus lack translational symmetry.
However, if the real crystal is sufficiently large, so that surface effects can be neglected, the properties connected to the full translational symmetry can be recovered by using the periodic boundary conditions, introduced in § 2.5.1. We want to investigate here the effects of periodic boundary conditions on the reciprocal space.

As in § 2.5.1, let's consider a crystal of parallelepiped shape, whose edges are $N_1\vec{a}_1, N_2\vec{a}_2, N_3\vec{a}_3$, so that it contains $N = N_1N_2N_3$ primitive cells (pink square in Fig. 3.5, left). This "fundamental" crystal is considered as a part of an infinite series of identical crystals extending in all directions with the same orientation.
The periodic boundary conditions (2.14)

$$\vec{T}_{N_1\vec{a}_1} = \vec{T}_{N_2\vec{a}_2} = \vec{T}_{N_3\vec{a}_3} = \vec{T}_0 \,,$$

where T_0 is the null translation, represent a constraint on the wavevectors k, so that:

$$e^{i\vec{k}\cdot\vec{r}} \;=\; e^{i[\vec{k}\cdot(\vec{r}\,+\,\vec{a}_i N_i)]} \,, \quad (i = 1,2,3) \tag{3.41}$$

which amounts to say that

$$e^{i[\vec{k}\,\cdot\,\vec{a}_i N_i]} \;=\; 1\,. \quad (i = 1,2,3) \tag{3.42}$$

Let us now express a given wavevector $\vec{k}$ of the reciprocal space in terms of the primitive vectors $\vec{b}_i$ of the reciprocal lattice and take into account the orthogonality relation (3.28) between primitive vectors of the direct and reciprocal lattice

$$\vec{k} = \sum_{i=1}^{3} x_i\vec{b}_i \,, \qquad \vec{a}_i \cdot \vec{b}_j = 2\pi\,\delta_{ij} \qquad (i = 1,2,3) \tag{3.43}$$

where x_i are in principle real coefficients.
The periodic boundary conditions (3.42) can be rewritten as

$$e^{i\,2\pi\,x_i N_i} \;=\; 1 \;=\; e^{i\,2\pi\,n_i} \qquad (i = 1,2,3) \tag{3.44}$$

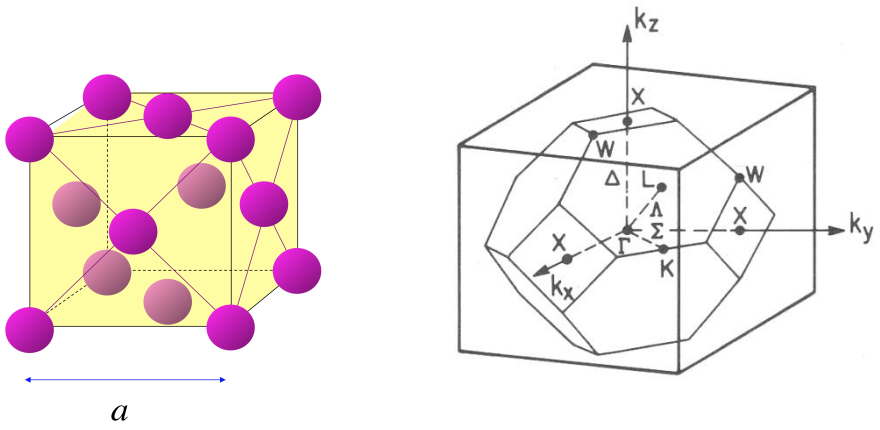

Figure 3.6: Left: fcc Bravais lattice in real space. Right: its First Brillouin Zone (in reciprocal space); the highest symmetry directions are conventionally labelled by specific letters.

where n_i are integer numbers, so that

$$\vec{k} = \sum_i x_i \vec{b}_i = \sum_i \frac{n_i}{N_i} \vec{b}_i \,. \tag{3.45}$$

The possible values of $\vec{k}$ represent a grid of points extended over all the reciprocal space (small black circles in Fig. 3.5, right). The density of the points of the grid increases if the size of the real ("fundamental") crystal increases.

First Brillouin zone

In principle, the choice of a primitive cell is arbitrary in the reciprocal space as well as in the direct space.
Particularly important is the Wigner-Seitz primitive cell centred on the origin of the reciprocal lattice, which is called First Brillouin Zone (green square in Fig. 3.5, right).

Example: The fcc Bravais lattice supports a number of important structures (fcc, diamond, zincblende, NaCl). One can easily see that the reciprocal lattice is bcc (§ 3.5.5). The First Brillouin Zone of the fcc Bravais lattice, say the Wigner-Seitz cell centred on the origin of the bcc lattice, is shown in Fig. 3.6.

The First Brillouin Zone contains only one reciprocal lattice point, $\vec{G} = 0$, which is conventionally called the Γ (gamma) point. The Γ point corresponds to a wave of infinite wavelength and thus has generally no physical interest.

The other points $\vec{k}$ within the first BZ in reciprocal space correspond to plane waves whose wavelength $\lambda = 2\pi/k$ is larger than the periods of the direct-space lattice. These long-wavelength waves can describe the normal modes of vibrations of atoms or the dynamical properties of electrons moving in the periodic crystal potential.

It will be shown in forthcoming chapters that, as a consequence of translational symmetry, a number of physical properties of crystals can be exhaustively described even if attention is limited to the $\vec{k}$ wavevectors of the First Brillouin Zone.

Constructing Brillouin zones

Let's consider in some detail the procedure for constructing the first Brillouin Zone, with reference to Fig. 3.7.
Draw lines connecting the origin of the reciprocal space (Γ point) to all the other points of the reciprocal lattice and bisect each line with a plane (these planes are named "Bragg planes").

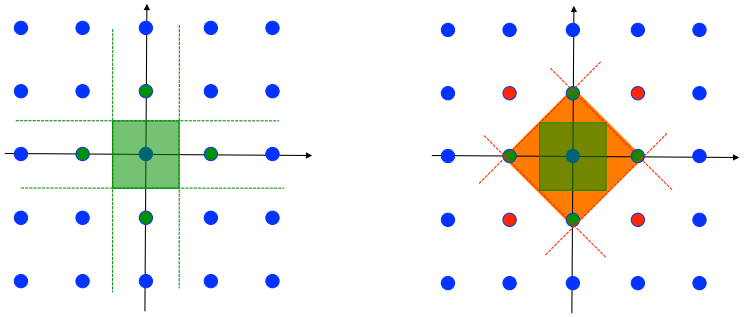

Figure 3.7: Left: First Brillouin Zone for a square lattice. Right: First and second Brillouin Zones for the same square lattice.

The first Brillouin Zone is the smallest polyhedron containing the origin of the reciprocal space and bounded by Bragg planes (green square in Fig. 3.7, left).
Other Brillouin Zones can be singled out by considering farther reciprocal lattice points and farther Bragg planes. For example, in Fig. 3.7 (right) the four triangular orange regions represent the 2nd Brillouin Zone for the considered square lattice.
The outer Brillouin Zones are the sum different disconnected regions. The total volume of every outer Brillouin zone is always equal to the volume of the first Brillouin Zone.

3.5 Complements and problems

3.5.1 Time periodicity and space periodicity

Time periodicity

A periodic function of time is characterized, in the direct space t, by the analytic expression $f(t) = f(t + \mathcal{T})$, where $\mathcal{T}$ is the period.
The conjugate variable of the time t is the frequency ν, measured in hertz (inverse seconds). In Physics, it is customary to consider the angular frequency $\omega = 2\pi\nu$ as conjugate variable.
To a given angular frequency ω it corresponds, in the direct space, a sinusoidal function, whose period is $\mathcal{T} = 2\pi/\omega$.

$$f(t) = C\,\cos(\omega t) + S\,\sin(\omega t) = A\,\sin(\omega t + \phi)\,. \tag{3.46}$$

Spatial periodicity

Let us now consider a periodic function of the linear spatial coordinate x, expressed analytically by $f(x) = f(x + a)$, where a is the period.
The conjugate variable of the position x (direct space) is the inverse length. In Physics, it is customary to consider the inverse length multiplied by 2π as conjugate variable ("spatial frequency").
To a given "spatial frequency" b it corresponds, in the direct space, a sinusoidal function, whose period is $a = 2\pi/b$:

$$f(x) = C\,\cos(bx) + S\,\sin(bx) = A\,\sin(bx + \phi)\,. \tag{3.47}$$

This description is quite familiar when dealing with sinusoidal plane waves, where a can be identified with the wavelength λ and b with the wavenumber $k = 2\pi/\lambda$:

$$f(x) = A\,\sin(kx + \phi) = A\,\sin(2\pi x/\lambda + \phi)\,. \tag{3.48}$$

3.5.2 Spatial periodic functions

Let us consider a periodic function of the spatial position x

$$f(x+T) = f(x+na) = f(x)\,, \tag{3.49}$$

where a is the period, n is an integer number and $T = na$ is a translation quantity.
The quantities $T = na$ form a lattice of points on the real one-dimensional axis x.
(Don't confuse the translation T with the time period $\mathcal{T}$.)

Cosine and sine Fourier series

The periodic function $f(x)$ can be expanded into a Fourier series

$$f(x) = f_0 + \sum_{m>0} [C_m \cos(mbx) + S_m \sin(mbx)] \tag{3.50}$$

$$= f_0 + \sum_{G} [C_G \cos(Gx) + S_G \sin(Gx)] \tag{3.51}$$

where $b = 2\pi/a$ and m is an integer mumber. In (3.51), $G = mb$.
The average value of the function (3.49) is

$$f_0 = \frac{1}{a}\int_0^a f(x)\,dx\,. \tag{3.52}$$

The Fourier coefficients C_m and S_m of (3.50) are real numbers defined as:

$$C_m = \frac{2}{a}\int_0^a f(x)\,\cos(mbx)\,dx \tag{3.53}$$

$$S_m = \frac{2}{a}\int_0^a f(x)\,\sin(mbx)\,dx \tag{3.54}$$

The actual shape of the periodic function $f(x)$ depends on the values of the Fourier coefficients C_m and S_m.

Reciprocal lattice

The quantities $G = mb$ of (3.51) form a lattice of points in the reciprocal one-dimensional space, which is called *reciprocal lattice*; the quantity $b = 2\pi/a$ is the period of the reciprocal lattice.
The period a of the function $f(x)$ determines the reciprocal lattice. The larger is a, the smaller is b, and viceversa (whence the name "reciprocal space").
Any periodic function (3.49) in real space can be reproduced with the desired approximation by summing up of a sufficiently large number of Fourier terms with appropriate coefficients.

Problem: Consider the behaviour of a sum of cosine terms with all coefficients $C_m = 1$

$$f(x) = \sum_{m>0} \cos(mbx) = \sum_{m>0} \cos\left(m\,\frac{2\pi}{a}\,x\right) \tag{3.55}$$

as a function of the number of terms. For one term (m=1), $f(x)$ is a cosine function. When more and more terms are added, corresponding to cosine functions of higher and higher frequencies, the function $f(x)$ progressively increases at the values $x = a, 2a, 3a, \ldots$, where all cosine functions are in phase. In the intermediate points the amplitude of the function $f(x)$ becomes progressively negligible with respect to the values at $x = a, 2a, 3a, \ldots$. Asymptotically, for $m \to \infty$, the function $f(x)$ tends to a lattice function $f(x) = \sum_n \delta(x - na)$.
To a Bravais lattice in the reciprocal space it corresponds a Bravais lattice in the real space, and viceversa.

Note: The Fourier expansion of a periodic spatial function (3.49) is formally equal to the more familiar case of the time and frequency domains: a periodic function $f(t)$ of period $\mathcal{T}$ is Fourier expanded in terms of a fundamental frequency $\omega_0 = 2\pi\nu = 2\pi/\mathcal{T}$ and its harmonics $m\omega_0$.

Exponential Fourier series

The Fourier expansion (3.50) is frequently expressed in more compact complex exponential form,

$$f(x) = \sum_{m=-\infty}^{+\infty} A_m \, e^{imbx} = \sum_G A_G \, e^{iGx} \tag{3.56}$$

where now the index m can be positive, zero or negative, and the coefficients A_m are complex quantities, given by:

$$A_m = \frac{1}{a} \int_0^a dx \, f(x) \, e^{-imbx} \, . \tag{3.57}$$

The condition that $f(x)$ is real imposes $A_{-m} = A_m^*$.
The connection between the A_m coefficients of (3.56) and the C_m and S_m coefficients of (3.50) is given by

$$A_m = \left[\frac{C_m}{2} + \frac{S_m}{2i} \right] , \quad A_{-m} = \left[\frac{C_m}{2} - \frac{S_m}{2i} \right] , \quad (m > 0) \, . \tag{3.58}$$

3.5.3 Non-periodic functions, Fourier integral

For non-periodic functions, the Fourier series (3.56) transforms into the Fourier integral. To the function $f(x)$ of the continuous variable x in the direct space it corresponds a function $\tilde{f}(k)$ of the continuous variable k in the reciprocal space. The two function are connected by the Fourier integrals:

$$\tilde{f}(k) \; = \; \int_{-\infty}^{+\infty} f(x) \, e^{-ikx} \, dx \, , \qquad f(x) \; = \; \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{f}(k) \, e^{ikx} \, dk \, . \tag{3.59}$$

Demonstration

Let us rewrite (3.56) and (3.57) for a periodic function of period a in the real space as

$$f_a(x) = \sum_{m=-\infty}^{+\infty} A_m \, e^{ik_m x} \, , \quad A_m = \frac{1}{a} \int_{-a/2}^{a/2} f(x) \, e^{-ik_m x} \, dx \, , \tag{3.60}$$

where now $mb = k_m$. For each value of m one has $a = 2\pi/(k_{m+1} - k_m)$. It is convenient to rewrite (3.60) as

$$f_a(x) = \sum_{m=-\infty}^{+\infty} \left[\frac{k_{m+1} - k_m}{2\pi} \int_{-a/2}^{a/2} f(\xi) \, e^{-ik_m \xi} \, d\xi \right] e^{ik_m x} \, , \tag{3.61}$$

For a non-periodic function, let $a \to \infty$, $k_{m+1} - k_m \to dk$ and substitute the sum in (3.61) by an integral

$$\begin{aligned} f(x) \; &= \; \int_{-\infty}^{+\infty} \frac{dk}{2\pi} \left[\int_{-\infty}^{+\infty} f(\xi) \, e^{-ik\xi} \, d\xi \right] e^{ikx} \\ &= \; \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \left[\tilde{f}(k) \right] e^{ikx} \end{aligned} \tag{3.62}$$

3.5.4 Demonstration of equation (3.11)

Let first re-adjust the sum

$$\sum_{n=-p}^{p} e^{-ikna} \; = \; e^{+ikpa} \sum_{n=-p}^{p} e^{-ik(p+n)a} \; = \; e^{+ikpa} \sum_{s=0}^{2p} \left(e^{-ika} \right)^s \, , \tag{3.63}$$

where the variable $s = p + n$ has been introduced for convenience.

Now, since the number of lattice points is $N = 2p+1$, so that $2p = N-1$ and $p = N/2-1/2$, and remembering that

$$\sum_{s=0}^{N-1} a^s = \frac{a^N - 1}{a-1}\,, \tag{3.64}$$

one finds

$$\sum_{n=-p}^{p} e^{-ikna} = e^{+ik(N-1)a/2}\,\frac{e^{-ikNa}-1}{e^{-ika}-1} = \frac{e^{-ikNa/2}-e^{+ikNa/2}}{e^{-ika/2}-e^{+ika/2}}\,, \tag{3.65}$$

whence finally

$$\sum_{n=-p}^{p} e^{-ikna} = \frac{\sin(Nka/2)}{\sin(ka/2)} \tag{3.66}$$

3.5.5 Reciprocal lattice of an fcc Bravais lattice

We have seen in § 2.7 that the simplest choice of primitive vectors for an fcc Bravais lattice is

$$\vec{a}_1 = (a/2)\,(\hat{x}+\hat{y}) \tag{3.67}$$
$$\vec{a}_2 = (a/2)\,(\hat{y}+\hat{z}) \tag{3.68}$$
$$\vec{a}_3 = (a/2)\,(\hat{z}+\hat{x}) \tag{3.69}$$

and that the volume of the primitive cell of the direct lattice is

$$V_p = (\vec{a}_1 \times \vec{a}_2)\cdot\vec{a}_3 = a^3/4 = V_c/4\,. \tag{3.70}$$

Making use of (3.36) one can easily find the primitive vectors of the reciprocal space

$$\vec{b}_1 = 2\pi\,\frac{\vec{a}_2\times\vec{a}_3}{\vec{a}_1\cdot(\vec{a}_2\times\vec{a}_3)} = \frac{2\pi}{a}\,(\hat{x}+\hat{y}-\hat{z}) \tag{3.71}$$
$$\vec{b}_2 = 2\pi\,\frac{\vec{a}_3\times\vec{a}_1}{\vec{a}_1\cdot(\vec{a}_2\times\vec{a}_3)} = \frac{2\pi}{a}\,(-\hat{x}+\hat{y}+\hat{z}) \tag{3.72}$$
$$\vec{b}_3 = 2\pi\,\frac{\vec{a}_1\times\vec{a}_2}{\vec{a}_1\cdot(\vec{a}_2\times\vec{a}_3)} = \frac{2\pi}{a}\,(\hat{x}-\hat{y}+\hat{z})\,. \tag{3.73}$$

Looking again at § 2.7, we can check that the $\vec{b}_i$ are primitive vectors of a bcc lattice.
The volume of the primitive cell of the reciprocal lattice is

$$V^* = (\vec{b}_1\times\vec{b}_2)\cdot\vec{b}_3 = \frac{32\pi^3}{a^3}a^3/4 = V_c/4\,. \tag{3.74}$$

The volume of the bcc conventional cell of the reciprocal lattice is

$$2V^* = \frac{64\pi^3}{a^3} \tag{3.75}$$

and the lattice parameter of the bcc conventional cell is

$$b = (2V^*)^{1/3} = \frac{4\pi}{a}\,. \tag{3.76}$$

3.6 Bibliography of Chapter 3

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions). Chapter 5 (The reciprocal lattice).
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapter 2 (Wave diffraction and the reciprocal lattice).

Chapter 4

Symmetry in molecules and crystals

The term symmetry originates from ancient Greek ($\sigma\upsilon\nu$ $\mu\epsilon\tau\rho\iota\alpha$, with measure). The original meaning was mainly aesthetic: a symmetric object is characterised by right proportions. Typical examples are the five platonic solids, appreciated for their regular shape (Fig. 4.1).
Only in the last two centuries has the idea of symmetry gained a clear mathematical meaning, connected to the concepts of transformation and of the invariance of the objects (or of their properties) under symmetry transformations (Section 4.1). The possibility of describing the symmetry properties within the framework of the group theory has dramatically enhanced the scientific applications of symmetry (Section 4.2).
Invariance properties under symmetry transformation represent a powerful tool in many fields of Physics. In this Chapter the attention is mainly focused on the symmetry properties of molecules (Sections 4.3 and 4.5 and crystals (Section 4.6).

Figure 4.1: The five regular (platonic) solids.

4.1 Introduction

Many natural objects exhibit some forms of approximate symmetry: flowers are often invariant with respect to some rotation angles or with respect to one reflection plane. The external appearance of animals is generally symmetrical with respect to a reflection plane; a starfish exhibits invariance with respect to a rotation of $2\pi/5$.
Symmetry has been often considered a requisite for aesthetic quality of human buildings, from Egyptian pyramids to gothic churches to stadiums. Some degree of symmetry characterises many man-made objects, such as cars, planes, ships, bottles.
In the following, we consider the symmetry properties of ideally perfect objects, such as geometric solids, molecules, crystals.

4.1.1 Phenomenological definition

A simple geometrical definition of symmetry is: a symmetry transformation (for short "a symmetry") of an object is a transformation that brings the object into a form indistinguishable from the

original one.

Example: The rotation of a cube by an angle of $2\pi/3$ around one of its main diagonals leaves it unchanged; it is a symmetry transformation (Fig. 4.2, left).

Alternatively, one can think of objects as composed of equal parts, which can be converted one into the other by symmetry transformations.

Example: In the ammonia molecule NH_3, the N atom is at the vertex of a triangular pyramid, and the three equal H atoms are at the corners of the base (Fig. 4.2, right). The rotation of the molecule by an angle of $2\pi/3$ around the axis perpendicular to the base and including the N atom converts each H atom into another H atom.

This definition is particularly appropriate when one is interested in the invariance of the shape and atomic structure of molecules and crystals.

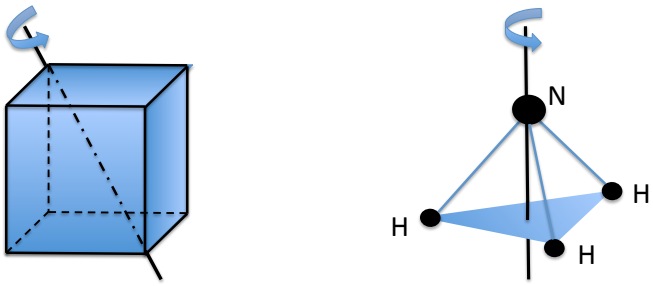

Figure 4.2: Left: rotation axis corresponding to a main diagonal of a cube. Right: the ammonia molecule NH_3.

On more general grounds, one is interested in the invariance, with respect to symmetry transformation, of some physical properties of systems, such as elastic, electric or magnetic properties. In Quantum Mechanics, one is interested in the invariance of the Hamiltonian operator and of the stationary Schrödinger equation.

The possible symmetry transformation of Physics are not limited to geometrical transformations. One can be interested in symmetry with respect to inversion of spatial or time coordinates, with respect to permutation of the elements of a set, with respect to the inversion of electric charge or with respect to more sophisticated transformations (e.g. the gauge invariance connected to the choice of the scalar and vector potentials of the electromagnetic fields).

4.1.2 Formal definition of symmetry

Let $\{\vec{x}\} = \{x_1, x_2, \dots x_m\}$ be a set of variables describing the properties of an object F. The variables can be cartesian coordinates or other types of variables; the property can be any physical property, including the aspect of the object.
Consider a transformation

$$g[\{x_1, x_2, \dots x_m\}] \;=\; \{x'_1, x'_2, \dots x'_m\} \;=\; \{\vec{x}\,'\}\,. \tag{4.1}$$

The object F is symmetric with respect to the transformation g if

$$F(\{x_1, x_2, \dots x_m\}) \;=\; F(g[\{x_1, x_2, \dots x_m\}]) \;=\; F(\{x'_1, x'_2, \dots x'_m\})\,, \tag{4.2}$$

or

$$F(\{\vec{x}\}) \;=\; F(\{\vec{x}\,'\})\,, \tag{4.3}$$

say if one cannot distinguish whether the object has undergone or not the transformation.
For each symmetry transformation $g[\{\vec{x}\}] = \{\vec{x}'\}$ there is an inverse transformation $g^{-1}[\{\vec{x}\,'\}] = \{\vec{x}\}$.

The identity transformation $g = e$ (or $g = E$), which leaves all the variables unaltered, is a symmetry operation too.

It is interesting to note that symmetry transformations have two equivalent interpretations: *a)* a change of coordinates with the object remaining unmoved, *b)* a change of the points of the object within an unmodified coordinate system.

4.1.3 Useful classifications

When dealing with molecules and crystals, it is convenient to distinguish two main types of symmetries:

A. Point symmetries, say operations that leave at least one point of the object unmoved.
We can further distinguish:

   a1. Discrete symmetry transformations
   e.g. rotations and reflections of a cube, rotations and reflections of an ammonia molecule
   a2. Continuous symmetry transformations
   e.g. rotations of a sphere, rotations of a spherically symmetric atom, rotations of a linear molecule around the bond axis

B. Translational symmetries, say operations that move all points by the same amount along the same direction.
Translational symmetry is in principle possible only for infinite objects. For finite crystals, the properties connected to full translational symmetry can be recovered by imposing the periodic boundary conditions (Section 2.5).
We can further distinguish:

   b1. Discrete symmetry transformations
   e.g. translation of an ideal (infinite) crystal lattice by a lattice vector $\vec{T}$
   b2. Continuous symmetry transformations
   e.g. continuous translations in time or in space

In the following, we focus our attention on two types of systems subject to discrete transformations:

1. molecules (point discrete symmetries), Section 4.3
2. crystals (point + translational discrete symmetries), Section 4.6

4.1.4 Composition of symmetry transformations and groups

Let us suppose that all the possible symmetry transformations $A, B, C, \ldots$ of a given system have been found; they form the set of transformations of the system.

Any two symmetry transformations can be composed, say performed one after the other. One finds that the result of the composition corresponds to one of the transformations of the set. Otherwise stated, the set of symmetry transformation is closed with respect to the composition of transformations.

Moreover, one can verify that the composition is always associative, $A(BC) = (AB)C$, and that for each transformation A there is an inverse transformation A^{-1} such that $AA^{-1} = E$, where E is the null transformation. For each system, the set of symmetry transformations has the properties of an algebraic group.

Notice, however, that in general the composition (which is often called multiplication) is not commutative.

In the next Section we well shortly review some of the basic properties of groups.

4.2 Group theory, basic elements

4.2.1 Group axioms

A group $\mathcal{G}$ is a set of elements $\{g_1, g_2, g_3, \ldots\}$ (or $\{A, B, C, \ldots\}$), finite or infinite, characterised by the following axioms:

1. A composition law (for short "multiplication") is defined, such that the product of any two elements of the set is again an element of the set (*closure property*): $AB = C$.
2. The multiplication is *associative*: $(AB)C = A(BC)$.
3. There exists an *identity element* E such that $EA = AE = A$ for every element A.
4. There is an *inverse element* A^{-1} for every element A of the group, such that $A^{-1}A = AA^{-1} = E$. One can easily show that $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.

The number of elements is the *order* of the group. The order can be finite or infinite.

Example 1: The only group of order one is the group containing only the identity element E.

Example 2: There is only one possibility of a group of order two: $\mathcal{G} = \{A, E\}$, where $AA = E$, so that $A = A^{-1}$.

Example 3: The only group of order 3 is $\mathcal{G} = \{A, B, E\}$, where $AA = B$ and $A^{-1} = B$.

Example 4: The set of natural numbers is not a group with respect to addition, since there is no identity element and there are no inverse elements.

Note: In the applications of group theory to crystallographic symmetry, don't confuse the symmetry elements (axes, planes, points) with the group elements, which actually correspond to symmetry transformations (rotations, reflections, inversion).

Commutative groups

If, for any pair A, B of group elements, the equality $AB = BA$ holds, the group is said to be *commutative* or *abelian*.

Example 1: An example of commutative group of infinite order is the set of all positive and negative integers, including zero, with respect to the ordinary addition operation. The zero is the identity element.

Example 2: An example of a non-commutative group of infinite order is the set of all $m \times m$ matrices with non-zero determinant. The composition law is the ordinary matrix multiplication. The identity element is the unit matrix $D_{ij} = \delta_{ij}$.

Multiplication tables

A finite group can be represented by a multiplication table, such as Table. 4.1. In the multiplication tables shown here, by convention the first column to the left of the table contains the first transformation and the first row on top of the table contains the second transformation. (In the product AB, B is the first transformation.)
In every multiplication table, each column and each row contains each element of the group once and only once (rearrangement theorem).
The group whose multiplication table is represented in Table. 4.1 is non-commutative; one can see that, for example, $CA = D \neq AC = F$. One can also see that $A^{-1} = B$, and so on.

4.2.2 Order of an element. Cyclic groups

Let us consider an element g of a finite group $\mathcal{G}$. The sequence

$$g,\ g^2,\ g^3,\ \ldots,\ g^p = E \tag{4.4}$$

Table 4.1: Multiplication table of a possible non-commutative group of order 6; as we will see later on, this corresponds to the multiplication table of the point symmetry group C_{3v}. From the Table we can see for example that $DB = C$, the first element B of the product being in the leftmost column, the second element D being in the upmost row.

	E	A	B	C	**D**	F
E	E	A	B	C	D	F
A	A	B	E	D	F	C
B	B	E	A	F	**C**	D
C	C	F	D	E	B	A
D	D	C	F	A	E	B
F	F	D	C	B	A	E

is called the period of the element g, and the integer p is the order of g. Otherwise stated, the order of an element is the minimum integer value p for which $g^p = E$.
If there is an element g whose period (4.4) exhausts all the elements of the group, then the group $\mathcal{G}$ is called cyclic,

$$\mathcal{G} \;=\; \{g, g^2, g^3, \ldots, g^p = E\}\,, \tag{4.5}$$

and the element g is the "generator" of the cyclic group.
If the period (4.4) is only a subset of the elements of the group $\mathcal{G}$, it is evident that it has the properties of a group; it a subgroup of $\mathcal{G}$ (see below).

All cyclic groups are commutative. Let g be the generator and $A = g^p$ and $B = g^q$ two elements; then $AB = g^{p+q} = g^{q+p} = BA$.

Example 1: The group of order 3, say $\mathcal{G} = \{A, B, E\}$, is cyclic, the generators being both A and B.

Example 2: The complex roots of order N of the number one, $1^{1/N}$, form a cyclic group of order N with respect to ordinary multiplication, $\mathcal{G} = \{e^{i2\pi/N}, \ldots, e^{i2\pi k/N}, \ldots, e^{i2\pi}\}$, where $e^{i2\pi/N}$ is the generator element and $e^{i2\pi} = 1$ is the identity element.

Problem 1: The 4-th order roots of the number one form the cyclic group $\mathcal{G} = \{i, -1, -i, 1\}$, where $i = \sqrt{-1}$ is the generator and 1 is the identity element. Write the multiplication table.

Problem 2: Verify that the group whose multiplication table is shown in Table 4.1 is not cyclic.

4.2.3 Subgroups - Cosets

A subset $\mathcal{S} = \{s_1 = e, s_2, \ldots s_k\}$ of a group $\mathcal{G} = \{g_1 = e, g_2, \ldots g_n\}$ $(k < n)$ is called a subgroup of $\mathcal{G}$ if its elements satisfy the group axioms. In particular, every subgroup $\mathcal{S}$ of $\mathcal{G}$ contains the identity element E (or e).

Let g_i be an element of the group $\mathcal{G}$ not belonging to the subgroup $\mathcal{S}$ $(g_i \in \mathcal{G}, g_i \notin \mathcal{S})$. The sets

$$g_i\,\mathcal{S} \;=\; \{g_i s_1, g_i s_2, \ldots, g_i s_k\}\,, \qquad \mathcal{S}\,g_i \;=\; \{s_1 g_i, s_2 g_i, \ldots, s_k g_i\}$$

are called a left coset and a right coset, respectively, of the subgroup $\mathcal{S}$ (in italian, "laterale sinistro" and "laterale destro").

One can verify the following properties:

a) A coset shares no elements with the subgroup.
Dem: Let be $s_j g_i = s_m \in \mathcal{S}$ for some g_i; then $g_i = s_j^{-1} s_m \in \mathcal{S}$, contrary to the assumption $g_i \notin \mathcal{S}$.
b) The cosets are no subgroups (they don't contain the identity element E).
c) A coset of a subgroup contains the same number k of elements as the subgroup.
Dem.: The number of elements of a coset cannot obviously be larger than k. It cannot be smaller too: actually, let $g_i s_j = g_i s_m$ for some j, m, then $s_j = g_i^{-1} g_i s_m = s_m$.

d) Two (right or left) cosets of a subgroup $\mathcal{S}$ of the group $\mathcal{G}$ either are identical or have no elements in common.
Dem.: Let us consider two cosets $g_1\mathcal{S}$ and $g_2\mathcal{S}$, and assume that they have a single element in common, $g_1 s_j = g_2 s_m$; then $g_2^{-1} g_1 = s_m s_j^{-1} \in \mathcal{S}$, so that $g_2^{-1} g_1 \mathcal{S} = \mathcal{S}$ and finally $g_1\mathcal{S} = g_2\mathcal{S}$.

e) Each element of the group $\mathcal{G}$ belongs to the subgroup $\mathcal{S}$ or to one of its cosets.
The order k of a subgroup $\mathcal{S}$ is always an integer divisor of the order n of the group G: $k = n/\ell$, where the integer ℓ is the index of the subgroup $\mathcal{S}$ in $\mathcal{G}$ (Lagrange theorem).

Example 1: In the group of Table 4.1 the subgroups are $\{E\}$, $\{E, A, B\}$, $\{E, C\}$, $\{E, D\}$, $\{E, F\}$.

Example 2: In the group of Table 4.1, the set $\{C, D, F\}$ is a coset of the subgroup $\{E, A, B\}$. Find possible cosets of the other subgroups.

Problem: Verify that any group of prime order must be cyclic (otherwise the period of some element would appear as a subgroup whose order is a divisor of a prime number).

4.2.4 Conjugate elements - Classes

Two elements A and B of a group $\mathcal{G}$ are said to be conjugate if there exists an element C of the group such that $A = C^{-1}BC$.
One can show that the two following properties hold for the conjugation relation:

- *Symmetric:* if B is conjugate of A, then A is conjugate of B.
  Dem.: If $B = C^{-1}AC$, then, by multiplying by C on the left and by C^{-1} on the right one gets $CBC^{-1} = A$.
- *Transitive:* if C is conjugate of B and B is conjugate of A, then C is conjugate of A.
  Dem.: Let be $B = D^{-1}AD$ and $C = F^{-1}BF$. By substituting B from the first into the second equation, $C = F^{-1}D^{-1}ADF$. But $F^{-1}D^{-1} = (DF)^{-1}$.

As a consequence of these properties, a group can be partitioned into *classes*, say into subsets without intersections and whose union is the entire group. All the elements within a class are conjugate, elements of two different classes are not conjugate.
One can verify that:

a) The identity E (e) is a class by itself.
Dem.: For any C, $C^{-1}EC = C^{-1}C = E$.

b) For commutative groups, any element is a class by itself.
Dem.: For any C and any A, $C^{-1}AC = AC^{-1}C = A$.

c) All elements of a class share the same order p.
Dem.: Let A and B belong to the same class, so that $B = C^{-1}AC$. Since $B^p = (C^{-1}AC)^p = (C^{-1}AC)\,(C^{-1}AC)\cdots = C^{-1}A^pC$, if $A^p \equiv E$, then $B^p = E$. Notice that E doesn't belong to the class of A and B.

Don't confuse classes with subgroups. A class, with the exception of the identity E, is not a sub-group.

The formal criteria for finding the classes of a group can be found in Landau-Lifshitz, *Quantum mechanics*, Chapter XII, or in more specialised textbooks.

Example: In the group of Table 4.1 there are three classes: $\{E\}$, $\{A, B\}$ and $\{C, D, F\}$.

The partition of a group into classes is important for the discussion of irreducible group representations (Chapter 5).

4.2.5 Isomorphism and homomorphism

Two groups $\mathcal{G}$ and $\mathcal{H}$ are said to be isomorphic (or isomorphous) if one can establish a one-to-one correspondence between their elements such that to the product of any two elements of $\mathcal{G}$ it corresponds the product of the two corresponding elements of $\mathcal{H}$:

$$\mathcal{G} \leftrightarrow \mathcal{H} \quad \text{if} \quad g_i \leftrightarrow h_i, \quad g_j \leftrightarrow h_j, \quad g_i g_j \leftrightarrow h_i h_j \tag{4.6}$$

Different isomorphic groups are different realisations of the same abstract group.

Example: The group C_N of rotations is isomorphic to the group of N complex numbers $\exp(2\pi ik/N$, where $0 \leq k \leq n$.

Two groups $\mathcal{G}$ and $\mathcal{H}$ may be in a unidirectional correspondence (homomorphism), when $\mathcal{G}$ is of higher order than $\mathcal{H}$, several elements of $\mathcal{G}$ are mapped into one element of $\mathcal{H}$ and the group operation is preserved.

Isomorphism and homomorphism are at the base of the theory of group representation (§ 5.2).

4.2.6 Invariant subgroups - Factor group

In the previous subsections we have introduced two kinds of possible partitions of a group $\mathcal{G}$: *a)* the partitions in terms of semigroups and cosets and *b)* the (unique) partition in term of classes.

A subgroup $\mathcal{S}$ consisting of entire classes is called an *invariant subgroup* (or normal subgroup).
It means that for any $s_i \in \mathcal{S}$ and for any $g_j \in \mathcal{G}$, one has $g_j^{-1} s_i g_j \in \mathcal{S}$. More synthetically, $g^{-1}\mathcal{S}g = \mathcal{S}$ for every g.
As a first consequence, $\mathcal{S}g = g\mathcal{S}$ for every g. In particular, right and left cosets of an invariant subgroup are equal (for the case of cosets, remember that $g \notin \mathcal{S}$).

Example: In the group of Table 4.1 the subgroup $\{E, A, B\}$ is invariant, since it is formed by the two classes $\{E\}$ and $\{A, B\}$. The subgroups $\{E, C\}$, $\{E, D\}$ and $\{E, F\}$ are not invariant.

It is convenient now to introduce the concept of *complex* $\mathcal{C}$ for any collection (subset) of group elements. Subgroups and cosets are examples of complexes.

The complexes represented by an invariant subgroup $\mathcal{S}$ and its cosets $g_i\mathcal{S} = \mathcal{S}g_i$ can be considered as the elements of a group of higher level, called the *factor group* $\mathcal{F}$

$$\mathcal{F} \;=\; \{\mathcal{S}, \mathcal{S}g_1, \mathcal{S}g_2, \ldots\} \;=\; \{\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2, \ldots\} \tag{4.7}$$

The order of the factor group is $\ell = n/k$, where n and k are the orders of the original group $\mathcal{G}$ and of the subgroup $\mathcal{S}$, respectively.
The invariant subgroup $\mathcal{S}$ is also called the *normal divisor* of the factor group $\mathcal{F}$.

One can verify that the factor group (4.7) has the properties of a group.
Each element of the factor group $\mathcal{F}$ is a complex $\mathcal{C}_i = g_i\mathcal{S} = \mathcal{S}g_i$; in particular, $\mathcal{C}_0 = e\mathcal{S} = \mathcal{S}$.
The composition of any two elements $\mathcal{C}_i$ of the factor group $\mathcal{F}$ is based on the invariance with respect to conjugation of the normal divisor subgroup $\mathcal{S} = \mathcal{C}_0$, which corresponds to the equality of left and right cosets:

$$\mathcal{C}_i\,\mathcal{C}_j \;=\; (g_i\mathcal{S})\,(g_j\mathcal{S}) \;=\; (g_i\mathcal{S}\mathcal{S}g_j) \;=\; g_i\mathcal{S}g_j \;=\; g_ig_j\mathcal{S}\,. \tag{4.8}$$

One can easily check that the identity element of the factor group $\mathcal{F}$ is the normal divisor $\mathcal{S}$ and that there exists an inverse for each element.

4.2.7 Direct product of two groups

Let us consider two groups:

$$\{\mathcal{H}\} = \{E_{\mathcal{H}} = h_1, h_2, h_3, \ldots, h_s\} \qquad \{\mathcal{K}\} = \{E_{\mathcal{K}} = k_1, k_2, k_3, \ldots, k_t\}$$

where the composition laws $h_i \circ h_v$ and $k_j \star k_w$ are defined. In general, the elements of the two groups can be of different nature and the composition laws can be different.
The direct product $\mathcal{G} = \mathcal{H} \times \mathcal{K}$ is defined as follows:

1. The underlying set is the cartesian product $\mathcal{H} \times \mathcal{K}$, whose $s \times t$ elements are the *ordered pairs* (h_i, k_j)

$$\mathcal{G} = \mathcal{H} \times \mathcal{K} = \{(h_1, k_1), (h_1, k_2), \ldots (h_i, k_1), (h_i, k_2), \ldots, (h_s, k_t)\} \tag{4.9}$$

2. The composition law in the direct product $\mathcal{G}$ is defined as

$$(h_i, k_j) \cdot (h_v, k_w) = (h_i \circ h_v, k_j \star k_w) \tag{4.10}$$

so that $E = (h_1, k_1) = (E_\mathcal{H}, E_\mathcal{K})$ is the identity element of $\mathcal{G}$ and $(h_i, k_j)^{-1} = (h_i^{-1}, k_j^{-1})$.

Example: Let $\mathcal{R}$ be the group of real numbers under addition. The direct product $\mathcal{R} \times \mathcal{R}$ is the group of all two-component vectors (x, y) under the operation of vector addition.

Let us now consider the two subsets $\mathcal{H}' = \{(h_1, E_\mathcal{K}) \dots (h_s, E_\mathcal{K})\}$ and $\mathcal{K}' = \{(E_\mathcal{H}, k_1) \dots (E_\mathcal{H}, k_t)\}$ of the direct product $\mathcal{G}$. The two subsets are isomorphic to $\mathcal{H}$ and to $\mathcal{K}$, respectively, and are thus proper subgroups of $\mathcal{G}$. (Sometimes, $\mathcal{H}'$ and $\mathcal{K}'$ are simply identified with $\mathcal{H}$ and $\mathcal{K}$, respectively).
One can easily verify the following properties of the two subgroups $\mathcal{H}'$ and $\mathcal{K}'$:

a) only the identity $(E_\mathcal{H}, E_\mathcal{K})$ is shared by $\mathcal{H}'$ and $\mathcal{K}'$
b) all elements of $\mathcal{H}'$ commute with all elements of $\mathcal{K}'$

As a consequence, both $\mathcal{H}'$ and $\mathcal{K}'$ are composed of entire classes, and are thus both invariant subgroups of the newly formed group $\mathcal{G}$.

Direct products of symmetry groups

Let us now apply the properties of the direct product to the specific case of symmetry groups.
If $\mathcal{H}$ and $\mathcal{K}$ are symmetry groups, their elements have the same nature (symmetry transformations) and the composition laws are the same (composition of symmetry transformations, in the following represented as a product $\cdot$). The identity element is the same for both groups, $E_\mathcal{H} = E_\mathcal{K} = E$.
It is customary to consider a more restrictive definition of the internal product with respect to the general case (4.9), say

$$\mathcal{G} = \mathcal{H} \otimes \mathcal{K} = \{(h_1 \cdot k_1), (h_1 \cdot k_2), \dots (h_i \cdot k_1), (h_i \cdot k_2), \dots, (h_s \cdot k_t)\} \tag{4.11}$$

where each ordered pair of (4.9) is substituted the product of its two transformations.
The composition law (4.10) is substituted, in the newly defined direct product $\mathcal{G}$, by

$$(h_i \cdot k_j) \otimes (h_v \cdot k_w) = h_i \cdot h_v \cdot k_j \cdot k_w \,, \tag{4.12}$$

say by the product of four symmetry transformations.
The above properties (a) and (b) of the subsets $\mathcal{H}'$ and $\mathcal{K}'$ are still valid and both initial groups $\mathcal{H}$ and $\mathcal{K}$ are again invariant subgroups of the direct product $\mathcal{G}$.
A simple example of direct product is the C_{2v} group (see §). Other more important examples of direct product for point symmetry groups are given in § 4.5.
For some applications (e.g. when crystal space groups are formed from point and translational groups, see Section 4.6), a *semi-direct product* is defined by relaxing the property (b), so that only one of the two subgroups is invariant. The resulting product still consists of ordered pairs $(h_i \cdot k_j)$, but with a multiplication rule slightly more complicated than (4.12).

4.3 Discrete point symmetries of molecules

Let us consider the isometric transformations of non-linear molecules and of geometrical bodies where at least one point remains fixed. Isometric means that the mutual distances between any two points of the objects remain unchanged. If at least one point remains fixed, the transformation is said to be orthogonal.
The orthogonal transformations can be analytically expressed as

$$\vec{r}\,' \;=\; \mathbf{R}\,\vec{r}, \tag{4.13}$$

where $\vec{r}$ and $\vec{r}'$ are the position vectors of any point before and after the transformation and $\mathbf{R}$ is a 3×3 orthogonal matrix.
In this Section, we consider the transformations (4.13) for which the object maintains unchanged its properties, $F(\vec{r}') = F(\vec{r})$, say the *point symmetry* transformations (or simply point symmetries). Furthermore, we focus our attention mainly on *discrete* symmetry transformations.

Five types of point symmetry transformations are distinguished, which are here labelled by their Schoenflies symbols. They can be grouped into two categories:

a) transformations of the 1st kind, or proper rotations, which lead to congruent equalities of the objects;
b) transformations of the 2nd kind, or improper rotations, which lead to mirror equalities of the objects.

It is also important to distinguish between

a) symmetry elements (planes, axes, points)
b) symmetry transformations related to a given symmetry element

In this Section, we consider the symmetry operations in a descriptive mode, without reference to the analytic form (4.13). In Chapter 5, we consider the mathematical representation of symmetry operations by means of matrices; eq. (4.13) exemplifies a representation in the basis of cartesian coordinates.

4.3.1 Identity E (or e)

Identity (symbol E or e) does actually correspond to no transformation at all. Considering identity is however necessary for two reasons:

1. there are some bodies, and in particular some molecules, without any symmetry element, e.g. the molecule C H Cl Br F;
2. identity is formally necessary to the group theory.

Identity can be considered as a transformation of the 1st kind (proper rotation).

4.3.2 Rotations C_N^k

By the symbol C_N one labels an N-fold rotation axis (symmetry element). The symmetry transformations supported by a C_N axis are the $N-1$ rotations by an angle $2\pi/N$ and its multiples $2\pi k/N$, where $k = 1, \ldots, N-1$. For $k = N$ the rotation reduces to the identity E.
Rotations are transformations of the 1st kind (proper rotations).

Example 1: The water molecule H_2O has one C_2 rotation axis (symmetry element) that supports one symmetry operation, the rotation by π. The rotation by 2π coincides with the identity E.

Example 2: The ammonia molecule NH_3 (triangular pyramid) has one C_3 rotation axis (symmetry element, which supports two symmetry operations, the rotation C_3^1 by $2\pi/3$ and the rotation C_3^2 by $4\pi/3$. Alternatively, the two symmetry operations can be described as the clockwise and anti-clockwise rotations by $2\pi/3$, C_3^- and C_3^+, respectively.

Example 3: A cube has 13 symmetry rotation axes, all of them passing through the cube center:

- three C_4 axes (perpendicular to the faces), each one supporting three rotations: C_4^1, $C_4^2 = C_2$ and C_4^3.
- four C_3 axes (through the main diagonals), each one supporting two rotations, C_3^1 and C_3^2.
- six C_2 axes (perpendicular to the edges), each one supporting a C_2 rotation.

The axes of highest symmetry, say the axes supporting the largest number of rotations, are the C_4 ones.

Example 4: The benzene molecule C_6H_6 has seven rotation axes:

- 1 C_6 axis (perpendicular to the plane of the molecule), which supports five rotations: $C_6^1, C_6^2 = C_3^1, C_6^3 = C_2, C_6^4 = C_3^2, C_6^5$.
- 3 C_2' axes (in the plane, through the C atoms)
- 3 C_2'' axes (in the plane, bisecting the C–C bonds).

Example 5: A sphere has an infinite number of C_∞ rotation axes, each one of them supporting an infinite number of possible rotations. The sphere is an example of continuous symmetry transformations.

4.3.3 Reflection σ

By the symbol σ one labels a mirror plane (symmetry element) as well as the transformation consisting in the reflection through the plane.
Reflections are transformations of the 2nd kind (improper rotations).
Three different types of mirror planes are distinguished:

1. σ_v: the mirror plane contains a rotation axis of maximum rotational symmetry
2. σ_h: the mirror plane is perpendicular to a rotation axis of maximum rotational symmetry
3. σ_d: the vertical mirror plane bisects the angle between two mutually perpendicular horizontal C_2 axes

Example 1: The water molecule H_2O has two orthogonal mirror planes (symmetry elements) containing the C_2 axis: we label them σ_v (corresponding to the plane of the molecule) and σ_v' (perpendicular to the plane of the molecule).

Example 2: The ammonia molecule NH_3 (triangular pyramid) has three σ_v planes through the C_3 axis, the angle between the planes being $2\pi/3$.

Example 3: A cube has three mirror planes parallel to the faces (three σ_v or σ_h) and six mirror planes through the face diagonals (six σ_v or σ_h).

Example 4: The benzene molecule C_6H_6 has one σ_h plane (the plane of the molecule), three σ_v planes (containing the C_3' axes) and three σ_d planes (containing the C_3'' axes).

Example 5: A sphere has an infinite number of mirror planes.

4.3.4 Inversion I

By the symbol I one labels a geometrical point, the centre of inversion (symmetry element), as well as the symmetry transformation consisting in the inversion through the point.
Inversion is a transformation of the 2nd kind (improper rotation).

Example 1: Objects that posses an inversion centre:
cube, sphere, octahedron, the molecule of benzene C_6H_6

Example 2: Objects that *do not* posses an inversion centre:
tetrahedron, the molecules of water H_2O, ammonia NH_3, methane CH_4

4.3.5 Roto-reflections S_N

By the symbol S_N one labels the roto-reflection (or improper rotation), say the transformation composed of

a rotation C_N + a reflection σ_h

The rotoreflection is a transformation supported by two symmetry elements: a mirror plane and an axis. It is a transformation of the 2nd kind.

Example 1: The methane molecule CH_4 (that has the geometrical form of a tetrahedron) is invariant with respect to six possible S_4 transformations. To check, inscribe the tetrahedron inside a

cube; the axes to be considered are the three C_4 axes of the cube, and the rotations are C_4^1 and C_4^3.

Example 2: A cube is invariant with respect to six S_4 transformations (from rotations C_4^1 and C_4^3) and eight S_6 transformations (supported by the four C_3 axes).

Example 3: The benzene molecule C_6H_6 is invariant with respect to two S_3 transformations (from rotations C_6^2 and C_6^4 and two S_6 transformations (from rotations C_6^1 and C_6^5).

Note: Not for all values of N a roto-reflection represents a new symmetry transformation. For example, it is easy to verify that $S_1 \equiv \sigma_h$ and $S_2 \equiv I$. The last relation can be checked for the to last examples of the cube and of the benzene molecule.

4.4 Point symmetry groups, examples

As we have seen in the previous Section 4.3, a given system can possess a number of symmetry elements that give rise to a number of different symmetry transformations.

Problem 1: Consider all the symmetry elements of the planar hexagonal molecule of benzene, C_6H_6 and the corresponding symmetry operations.

Problem 2: Consider all the symmetry elements of a cube, an octahedron and a tetrahedron and the corresponding symmetry operations.

The composition (multiplication) of any two symmetry transformations of a given system gives rise again to a symmetry transformation.
By convention, the product AB means that the transformation is obtained by first performing the B transformation and then A transformation.

One can easily verify that the symmetry transformations and their compositions have the *basic properties of the algebraic groups*, which have been summarised in Section 4.2.
Different systems are characterised by different sets of symmetry transformations, which correspond to different symmetry groups. For orthogonal transformations, one speaks of *point-symmetry groups*.

The different point-symmetry groups are usually described by means of the Schönflies notation. A general introduction to the classification of point groups and to the Schönflies notation is given in the next Section 4.5.
In this Section 4.3 we present some tutorial examples of simple algebraic point groups and of the isomorphic symmetry groups.

4.4.1 Example: the C_3 group

The C_3 group contains only three symmetry transformations, identity and two rotations by 120 and 240 degrees around a given axis: $C_3 = \{e, C_3^1, C_3^2\}$. It is thus a group of order $n = 3$.
A simple system whose symmetry properties are represented by the C_3 group is the molecule of boric acid $B(OH)_3$ (Fig. 4.3, left).

Figure 4.3: The molecules of boric acid (left) and of water (right).

The multiplication table of the C_3 group is given in Table 4.2. The C_3 group is cyclic, the generator element is C_3^1. Being cyclic, it is commutative and each element forms a class by itself.

Table 4.2: Multiplication table of the C_3 group.

C_3	e	C_3^1	C_3^2
e	e	C_3^1	C_3^2
C_3^1	C_3^1	C_3^2	e
C_3^2	C_3^2	e	C_3^1

4.4.2 Example: groups of order $n = 4$

There are two algebraic groups of order $n = 4$, the cyclic group and the non-cyclic "Klein four-group".

1) The cyclic group of order 4 is isomorphic to the C_4 point-symmetry group (Schönflies notation), whose only symmetry element is a C_4 rotation axis. The multiplication table is shown in Table 4.3, left. Being cyclic, the C_4 group is commutative.

2) The Klein four-group K_4 is characterised by the multiplication table shown in Table 4.3, centre. Each group element coincides with its inverse, $A^{-1} = A$.
It is easy to see that the Klein four-group K_4 is the *direct product* of two copies of the (cyclic) group of order two, $K_4 = \{E, a\} \times \{E, b\}$: the elements of the direct product are the ordered pairs $(E \cdot E), (a \cdot E), (E \cdot b), (a \cdot b)$, which correspond to E, A, B, C, respectively.
Every non-cyclic group of order 4 is isomorphic to the Klein four-group. There are three point symmetry groups isomorphic to the Klein four-group:
   a) the D_2 group, whose only symmetry elements are three perpendicular C_2 axes;
   b) the C_{2h} group, whose only symmetry elements are a vertical C_2 axis and a horizontal mirror plane;
   c) the C_{2v} group, whose symmetry elements are a C_2 axis and two vertical mirror planes.

The C_{2v} group - the water molecule

A simple system whose symmetry properties are represented by the C_{2v} group is the water molecule H_2O (Fig. 4.3, right). The molecule has three symmetry elements (one rotation axis and two mirror planes) and four symmetry transformations: $E, C_2, \sigma_v, \sigma_v'$. The multiplication table is given in Table 4.3, right; the isomorphism with the Klein group is evident.

Table 4.3: Multiplication tables of the groups of order $n = 4$. Left: $n = 4$ cyclic group, represented by the C_4 symmetry group. Centre and right: $n = 4$ non-cyclic "Klein four-group" K_4 and the isomorphic C_{2v} symmetry group.

C_4	E	C_4^1	C_4^2	C_4^3
E	E	C_4^1	C_4^2	C_4^3
C_4^1	C_4^1	C_4^2	C_4^3	E
C_4^2	C_4^2	C_4^3	E	C_4^1
C_4^3	C_4^3	E	C_4^1	C_4^2

K_4	E	A	B	C
E	E	A	B	C
A	A	E	C	B
B	B	C	E	A
C	C	B	A	E

C_{2v}	E	σ_v	σ_v'	C_2
E	E	σ_v	σ_v'	C_2
σ_v	σ_v	E	C_2	σ_v'
σ_v'	σ_v'	C_2	E	σ_v
C_2	C_2	σ_v'	σ_v	E

One can easily verify the following properties of the C_{2v} group, which is isomorphic to the Klein four-group:

- The C_{2v} group is *non cyclic*.
- All diagonal elements of the table correspond to the identity E: each transformation corresponds to its own inverse: $(C_2)^{-1} = C_2$, etc.
- The multiplication table is *symmetric*: the C_{2v} group is commutative.
- Besides the trivial subgroup $\{E\}$, one can distinguish three cyclic subgroups: $\{E, C_2\}$, $\{E, \sigma_v\}$, $\{E, \sigma'_v\}$.
- There are *four classes*: $\{E\}$, $\{C_2\}$, $\{\sigma_v\}$, $\{\sigma'_v\}$
- *All subgroups are invariant*, since all of them are formed by entire classes.
  For example, the subgroup $\mathcal{S} = \{E, C_2\} = \mathcal{C}_0$ is formed by the two classes $\{E\}$ and $\{C_2\}$.
  The set $\{\sigma, \sigma'\} = \mathcal{C}_1$ is the coset of the invariant subgroup $\mathcal{S}$.
  The factor group is $\mathcal{F} = \{\mathcal{C}_0, \mathcal{C}_1\}$. By means of (4.8), one can verify the group properties of $\mathcal{F}$.
- The C_{2v} group is the *direct product* of the two cyclic groups of order two, $C_{2v} = \{E, \sigma\} \times \{E, \sigma'\}$: the elements of the direct product are the pairs $(E \cdot E = E), (\sigma \cdot E = \sigma), (E \cdot \sigma' = \sigma'), (\sigma \cdot \sigma' = C_2)$, which correspond to E, A, B, C, respectively.

4.4.3 Example: the C_{3v} group (the ammonia molecule)

Let us now consider point groups of order $n = 6$.
An interesting group of order 6 is the C_{3v} group, whose properties are here studied by considering a real system, the ammonia molecule NH_3. The ammonia molecule has 6 symmetry transformations: E, $C_3^1 = C_3^+$ (anticlockwise rotation), $C_3^2 = C_3^-$ (clockwise rotation), $\sigma'_v, \sigma''_v, \sigma'''_v$, which constitute the 6 elements of the C_{3v} group. The multiplication table of the C_{3v} is shown in Table 4.4, left.

One can easily verify the following properties of the C_{3v} group:

- The C_{3v} is formed by three 1st-kind transformations (proper rotations) and three 2nd -kind transformations (mirror reflections).
- The C_{3v} group is not cyclic; one can verify that
  - the order of E is $n = 1$, since $E^1 = E$;
  - the order of σ_v is $n = 2$, since $(\sigma_v)^2 = E$;
  - the order of $C_3^{\pm}$ is $n = 3$, since $(C_3^{\pm})^3 = E$.
- The group C_{3v} is not commutative: for example, $\sigma'_v \sigma''_v \neq \sigma''_v \sigma'_v$.
- Each reflection corresponds to its own inverse. For rotations: $(C_3^+)^{-1} = C_3^-$ and $(C_3^-)^{-1} = C_3^+$, say $(C_3^1)^{-1} = C_3^2$ and $(C_3^2)^{-1} = C_3^1$.

Table 4.4: Multiplication table of the C_{3v} group (left) and of the C_6 group (right).

C_{3v}	E	C_3^1	C_3^2	σ'_v	σ''_v	σ'''_v
E	E	C_3^1	C_3^2	σ'_v	σ''_v	σ'''_v
C_3^1	C_3^1	C_3^2	E	σ''_v	σ'''_v	σ'_v
C_3^2	C_3^2	E	C_3^1	σ'''_v	σ'_v	σ''_v
σ'_v	σ'_v	σ'''_v	σ''_v	E	C_3^2	C_3^1
σ''_v	σ''_v	σ'_v	σ'''_v	C_3^1	E	C_3^2
σ'''_v	σ'''_v	σ''_v	σ'_v	C_3^2	C_3^1	E

C_6	E	C_6^1	C_6^2	C_6^3	C_6^4	C_6^5
E	E	C_6^1	C_6^2	C_6^3	C_6^4	C_6^5
C_6^1	C_6^1	C_6^2	C_6^3	C_6^4	C_6^5	E
C_6^2	C_6^2	C_6^3	C_6^4	C_6^5	E	C_6^1
C_6^3	C_6^3	C_6^4	C_6^5	E	C_6^1	C_6^2
C_6^4	C_6^4	C_6^5	E	C_6^1	C_6^2	C_6^3
C_6^5	C_6^5	E	C_6^1	C_6^2	C_6^3	C_6^4

Subgroups of C_{3v}

One can easily verify that the three transformations $\{E, C_3^1, C_3^2\}$ form a cyclic (and hence commutative) subgroup of the group C_{3v}. It corresponds to the group of rotations without mirror reflections C_3.
Four other subgroups are present in C_{3v}: $\{E\}, \{E, \sigma_v'\}, \{E, \sigma_v''\}, \{E, \sigma_v'''\}$.

Classes of C_{3v}

One can easily verify that C_3^1 and C_3^2 are conjugate elements, since

$$\sigma_v' \, C_3^1 \, \sigma_v' = C_3^2 \, . \tag{4.14}$$

Similarly, one can verify that $\sigma_v', \sigma_v'', \sigma_v'''$ are conjugate elements:

$$C_3^1 \, \sigma_v' \, C_3^2 = \sigma_v''' \, , \qquad C_3^1 \, \sigma_v'' \, C_3^2 = \sigma_v' \, , \qquad C_3^1 \, \sigma_v''' \, C_3^2 = \sigma_v'' \, . \tag{4.15}$$

The C_{3v} group contains three classes: $\{E\}, \{{C_3}^1, {C_3}^2\}, \{\sigma_v', \sigma_v'', \sigma_v'''\}$. Only the class $\{E\}$ is a subgroup. All elements within a given class have the same period.

The subgroup $\mathcal{S} = \{E, C_3^1, C_3^2\} = \mathcal{C}_0$ is invariant, since it is formed by the two classes $\{E\}$ and $\{C_3^1, C_3^2\}$. The subset $\{\sigma, \sigma', \sigma''\} = \mathcal{C}_1$ is the coset of the invariant subgroup $\mathcal{S}$.

The subgroups $\{E, \sigma\}$ are not invariant, since they are not formed by entire classes.
The factor group is $\mathcal{F} = \{\mathcal{C}_0, \mathcal{C}_1\}$. By means of (4.8), one can verify the group properties of $\mathcal{F}$.

Problem: Find the classes of the rotation group C_3. Remember that C_3 is cyclic and thus commutative. Discuss the difference with respect to the classes of C_{3v}.

The D_3 group, isomorphic to C_{3v}

A triangle has four symmetry elements: a threefold axis perpendicular to the triangle and three two-fold axes in the plane of the triangle. The corresponding symmetry operations form the 6-ordered D_3 symmetry group (characterised by 3 axes perpendicular to the main axis): $D_3 = \{E, C_3^1, C_3^2, C_2', C_2'', C_2'''\}$.
One can easily calculate the multiplication table of the D_3 group and verify that the D_3 group is isomorphic to the C_{3v} group.
As already observed for the case of the $n = 4$ non-cyclic group, symmetry groups composed by different symmetry operations can be isomorphic; they don't only have multiplication tables with the same structure but, as we will see in Chapter 5, have the same character table of irreducible representations.

4.4.4 Example: the point group C_6

Another interesting group of order $n = 6$ is the group of rotations C_6, whose multiplication table is shown in Table 4.4, right. In the group C_6, the only symmetry transformations are proper rotations around a C_6 axis.

One can easily verify the following properties of the C_6 group and compare them with the corresponding properties of the C_{3v} group:

- The C_6 is formed by six 1st kind-transformations (proper rotations).
- The C_6 group is cyclic; the generator element is C_6^1 (rotation by 60°).
  Being cyclic, the group C_6 is commutative. As a consequence, each element represents a class by itself; there are thus six one-element classes.
- The inverse of each transformation is give by $(C_6^k)^{-1} = C_6^{6-k}$.

The cyclic group C_6 includes two cyclic subgroups: $\{E, C_6^2, C_6^4\}$, isomorphic to the group C_3, and $\{E, C_6^3\}$, isomorphic to the group C_2.

Since each element represents a class by itself, both subgroups of C_6 are composed of entire classes and are thus invariant.

Problem: Find the cosets of the two subgroups of C_6.

4.5 Point symmetry groups, classification

In Section 4.3 all the possible discrete symmetry operations of molecules have been listed and a little number of concrete examples have been studied.
We want now to give a systematic account of all the possible point symmetry groups and of their classification.

4.5.1 The $O(3)$ and $SO(3)$ groups

In the three-dimensional Euclidean space, all the point symmetry groups can be considered as subgroups of the $O(3)$ group:

> $O(3)$ is the *orthogonal group*, say the group of all distance-preserving transformations (isometries) that leave the origin fixed; otherwise stated, $O(3)$ is the group of all the symmetries of a sphere. The elements of $O(3)$ can be represented by orthogonal matrices.

A number of possible point symmetry groups can be considered as subgroups of the smaller group $SO(3)$:

> $SO(3)$ is the *special orthogonal group*, subgroup of $O(3)$ containing only proper rotations. The elements of $SO(3)$ can be represented by orthogonal matrices with determinant $+1$.

Note: An orthogonal matrix $\mathbf{R}$ is a real square matrix with the property that its transpose is equal to its inverse, $\mathbf{R}^T = \mathbf{R}^{-1}$, say $\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. The determinant of an orthogonal matrix is either 1 or -1. Orthogonal matrices preserve the dot product of vectors, so that isometric transformations (4.13) are described by orthogonal matrices.

All point groups, subgroups of $O(3)$ and, in some cases, even of $SO(3)$, can be grouped into:

a) Seven *simple* discrete point groups, characterised by the possible presence of only one axis of N-fold symmetry, with $N > 2$.
b) Seven *high symmetry* discrete point groups, characterised by the presence of more than one axis of order $N > 2$.
c) *Continuous* groups.

4.5.2 The seven families of simple discrete point groups

In these groups, only one axis C_N with $N > 2$ can be present. Some groups have additional axes of 2-fold symmetry (dihedral axes) perpendicular to the main C_N axis.

1. Groups C_N. The only symmetry element is an axis of N-fold symmetry. These groups are cyclic Abelian groups.
2. The C_{Nv} groups contain a C_N axis and N σ_v reflection planes.
3. The C_{Nh} groups contain a C_N axis and a σ_h plane normal to it.
4. The groups S_N contain an N-fold roto-reflection axis.
   If N is odd, these groups are equivalent to C_{Nh}.
   The group S_2 contains only two elements, the identity E and the roto-reflection S_2, that corresponds to the inversion I. This group is often called the *inversion group*, and symbolised by $C_i = \{E, I\}$.

5. The groups D_N have N two-fold axes perpendicular to the main C_N axis.
6. The groups D_{Nd} contain the elements of D_N together with dihedral reflection planes σ_d bisecting the angles between the two-fold axes perpendicular to C_N.
7. The groups D_{Nh} contain the elements of D_N plus the horizontal reflection plane σ_h.

Some of the point groups listed above and containing improper rotations can be obtained from simpler groups containing only proper rotations and the inversion group $C_i = \{E, I\}$ by means of the *direct product* defined in eqs. (4.11) and (4.12). For example:

$$\begin{array}{ll} C_{2h} = C_2 \otimes C_i & D_{2h} = D_2 \otimes C_i \\ S_6 = C_3 \otimes C_i & D_{3d} = D_3 \otimes C_i \end{array} \tag{4.16}$$

In crystals (see Section 4.6), the presence of translational symmetry imposes some limitations to the possible pint groups: only the values $N = 1, 2, 3, 4, 6$ fare allowed.

Example: the benzene molecule

The benzene molecule belongs to the D_{6h} group. The symmetry elements are:

a) Seven rotation axes: 1 C_6, 3 C_2' and 3 C_2''
b) One inversion center I (to be considered in the three-dimensional space)
c) Seven mirror planes: 1 σ_h, 3 σ_v and 3 σ_d

The 24 symmetry transformations (group elements) supported by the 15 symmetry elements can be grouped in two sets.

1. A set of 12 *proper rotations*, forming the group D_6:
   1a. The identity E
   1b. Five rotations around the C_6 axis: two C_6, two C_3, one C_2
   1c. One rotation for each of the three C_2' axes
   1d. One rotation for each of the three C_2'' axes
2. A set of 12 additional *improper rotations*, which can be obtained by multiplying each of the proper rotations of the group D_6 by the inversion I:
   2a. $EI = I$
   2b. $C_2 I = \sigma_h$
   $2\ C_6 I = 2\ S_3$
   $2\ C_3 I = 2\ S_6$
   2c. $3\ C_2' I = 3\ \sigma_d$
   2d. $3\ C_2'' I = 3\ \sigma_v$

The group D_{6h} can thus be obtained as the direct product of the proper rotations group D_6 and the inversion group $C_i = \{E, I\}$, say $D_{6h} = D_6 \times C_i$.

4.5.3 The seven high-symmetry discrete groups

These groups are said to be of very high symmetry, since they have more than one rotation axis of order greater than 2.

1. Group T, of order 12. It is the group of the proper rotations of the tetrahedron.
   The 12 transformation are proper rotations, say: identity E, three rotations around the three C_2 axes and eight rotations around the four C_3 axes.
2. Group T_d, of order 24. It is the group of all the symmetry transformations of the tetrahedron.
   The 24 transformations of T_d are: the 12 proper rotations + 6 mirror reflections + 6 S_4 roto-reflections. An example of T_d group is the methane molecule CH_4.

3. Group T_h, of order 24. It is obtained by the direct product of the group T and the inversion group $C_i = \{E, I\}$, say $T_h = T \otimes C_i$. Note that a regular tetrahedron cannot belong to the T_h group, since it has no inversion symmetry.
4. Group O, of order 24. It is the group of the proper rotations of the cube (or equivalently of the octahedron). See below the complete list.
5. Group O_h, of order 48. It is the group of all symmetry operations of the cube (or of the octahedron). It can be obtained as a direct product $O_h = O \otimes C_i$, since the cube has inversion symmetry. See below the complete list.
6. Group Y (or i), of order 60. It is the group of the proper rotations of the icosahedron, the regular solid with 20 triangular faces, 30 edges and 12 vertices (see Fig. 4.1).
   The 60 transformations are: identity E, 20 rotations around the ten C_3 axes normal to the faces, 15 rotations around the 15 C_2 axes normal to the edges and 24 rotations around the six C_5 axes.
7. Group Y_h (or I_h), of order 120. It is the group of all symmetry operations of the icosahedron. It can be obtained as a direct product $Y_h = Y \otimes C_i$, since the icosahedron has inversion symmetry.

Only the tetrahedral and octahedral groups T, T_d, T_h, O, O_h are compatible with the translational symmetry of crystals.
The interest in icosahedral symmetry has recently arisen as a consequence of the experimental observation of atomic icosahedral packing in some liquids.

The symmetry of the cube

The cube and the octahedron belong to the O_h group, with 48 elements (symmetry transformations).
The symmetry elements are:

a) Rotation axes: three C_4 (normal to the faces), four C_3 (main diagonals) and six C_2
b) Inversion center I
c) Mirror planes: three σ_h (parallel to the faces) and six σ_v (through face diagonals)

The symmetry operations can be grouped in two sets.

1. A set of 24 proper rotations, forming the group O (5 classes):
   1a. The identity E
   1b. Nine rotations around the three C_4 axes: three C_4^1, three C_4^2 and three $C_4^2 = C_2$
   1c. Six rotations around the six C_2 axes
   1d. Eight rotations around the four C_3 axes: four C_3 and four C_3^2
2. A set of 24 improper rotations, which can be obtained by multiplying each of the proper rotations of the group O by the inversion I:
   - I
   - 3 σ_h
   - six σ_v (through face diagonals)
   - eight S_6 around C_3 axes (four S_6 and S_6^5)
   - six S_4 around C_4 axes ($C_4\sigma_h$ and $C_4^2\sigma_h$)

The group O_h is thus the direct product of the proper rotations group O and the inversion group $C_i = \{E, I\}$, say $O_h = O \otimes C_i$.
One can verify that the group O_h has 10 classes.

4.5.4 The continuous groups

Let us conclude this Section 4.5 dedicated to point symmetry with the possible continuous groups.

1. The group $C_{\infty v}$ is the group of a linear molecule which has only rotational symmetry around the axis containing the atoms.
2. The group $D_{\infty h}$ is the group of homonuclear linear molecules or of symmetric linear molecules, such as CO_2. In addition to the symmetry axis containing the atoms, there are infinite rotational symmetry axes and a mirror plane passing through the centre and perpendicular to the axis containing the atoms.
3. The group $O(3)$ is the group of all symmetries of a sphere, e.g. of a spherically symmetric atom.

4.6 Symmetry in crystals

In crystals, one has to consider two types of symmetry transformations;

1. *Point symmetry transformations*, say transformations where at least one point remains fixed. The types of symmetry transformations are the same as for molecules (§ 4.3), with some restrictions imposed by the long range translational order of crystals: only the values $N = 1, 2, 3, 4, 6$ are allowed for C_N axes.
2. *Translation symmetry transformations.*
   For finite crystals, the properties connected to full translational symmetry can be recovered by imposing the periodic boundary conditions (Section 2.5).

The possible symmetry transformations of crystals include not only pure rotations (proper and improper) or pure translations but also combinations of translations and rotations.
A symmetry transformation of a crystal can be analytically expressed, in terms of the coordinates of a vector $\vec{r}$, as

$$\vec{r}\,' \;=\; \mathbf{R}\,\vec{r} \;+\; \vec{T}\,, \tag{4.17}$$

where $\mathbf{R}$ is an orthogonal matrix representing a proper or improper rotation as in (4.13) and $\vec{T} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$ is a lattice translation vector.
In a more compact form, (4.17) is often expressed as

$$\vec{r}\,' \;=\; \{\mathbf{R}|\vec{T}\}\,\vec{r}\,. \tag{4.18}$$

All the possible transformations (4.18) for a given system form a group. One can easily verify, by successive applications of (4.17), the multiplication rule

$$\{\mathbf{R}'|\vec{T}'\}\;\{\mathbf{R}|\vec{T}\} \;=\; \{\mathbf{R}'\mathbf{R}|\mathbf{R}\,\vec{T} + \vec{T}'\} \tag{4.19}$$

and the inversion

$$\{\mathbf{R}|\vec{T}\}^{-1} \;=\; \{\mathbf{R}^{-1}|-\mathbf{R}^{-1}\,\vec{T}\}\,. \tag{4.20}$$

If $\vec{T} = 0$ (no translations) we recover the subgroup of proper and improper rotations considered in previous Section 4.3.

If $\mathbf{R} = E$ (no proper or improper rotations) we obtain the infinite subgroup of translations $\vec{T}$, where $E = 0$ and $\vec{T}^{-1} = -\vec{T}$.
The subgroup of translations is commutative; each element is a class by itself. As a consequence, the subgroup of translations is made by entire classes, and is thus an invariant subgroup of the full symmetry group.
The composition law (4.19) is more complicated than the composition law (4.12) of a group obtained by a simple direct product.

For classification purposes, it is convenient to consider first only the symmetries of the Bravais lattices, then the symmetries of the complete crystals structures (Bravais lattice + basis).

4.6.1 Symmetries of the Bravais lattices

The Bravais lattices can be classified *a)* considering only the point transformations or *b)* considering also translations in addition to point transformations.

a) Point symmetry of the lattice: 7 crystal systems

There is only a finite number of different possible point groups for three-dimensional Bravais lattices. Actually, not all the symmetry transformations that are possible for molecules are possible for crystal lattices too, because the repetition of the unit cells has to fill up the space. The possible point symmetry elements are:

$$E,\ I,\ \sigma,\ C_1,\ C_2,\ C_3,\ C_4,\ C_6,\ S_4$$

The rotations axes C_N with $N = 5$ and $N > 6$ are incompatible with the crystal structure.

Only seven different types of polyhedra can fill-up the three-dimensional space, corresponding to the *seven crystal systems*

cubic, hexagonal, trigonal, tetragonal, orthorhombic, monoclinic, triclinic

The 7 crystal systems are listed in Table 4.5. The cells are chosen so as to exhibit the highest symmetry; they are conventional cells: only in some cases they correspond to primitive cells and define a Bravais lattice.

Note: The hexagonal and trigonal systems are strictly connected and are often grouped as a unique hexagonal crystal family

Table 4.5: Point symmetry: the 7 crystal systems and the repartition of 32 point groups of crystal symmetry. See also Fig. 4.4 for the cell shapes.

7 Crystal systems	Cell parameters	Point symmetry requirements	Point groups
Cubic	$a = b = c$; $\alpha = \beta = \gamma = 90°$	4 C_3 axes	5
Hexagonal	$a = b \neq c$; $\alpha = \beta = 90°$, $\gamma = 120°$	1 C_6 axis	7
Trigonal	$a = b = c$; $\alpha = \beta = \gamma \neq 90°$	1 C_3 axis	5
Tetragonal	$a = b \neq c$; $\alpha = \beta = \gamma = 90°$	1 C_4 axis	7
Orthorhombic	$a \neq b \neq c$; $\alpha = \beta = \gamma = 90°$	3 C_2 axes or 1 C_2 axis and 2 σ planes	3
Monoclinic	$a \neq b \neq c$; $\alpha = \beta = 90°$, $\gamma \neq 120°$	1 C_2 axis or 1 σ plane	3
Triclinic	$a \neq b \neq c$; $\alpha \neq \beta \neq \gamma \neq 90°$	none	2

b) Space symmetry: 14 Bravais lattices

Let us now consider, in addition to the point group symmetry, the translational symmetries of an infinite lattice. The infinite lattice is invariant with respect to translations $\vec{T} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$. As a consequence, to the seven point groups (crystals systems) there correspond *14 Bravais lattices*, say 14 different groups (point transformations + translation). The 14 Bravais lattices are listed in Table 4.6.

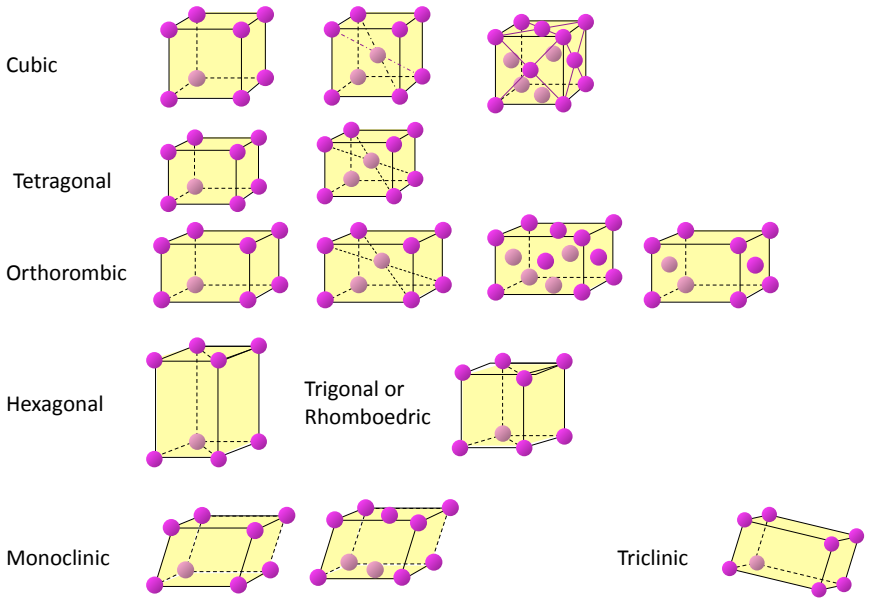

Figure 4.4: The 14 Bravais lattices.

The conventional unit cells correspond to the unit cells of the seven crystal systems with the possible addition of lattice points at the centre of the cells or at the centre of all or some of their faces.

Note: Some crystals of the trigonal crystal system exhibit hexagonal Bravais lattice symmetry. In an attempt to clarify the situation, one distinguishes crystal systems and lattice systems. Crystal systems and lattice systems coincide for all systems except the hexagonal and trigonal ones. The Bravais lattices of the trigonal crystal system can be hexagonal or orthorhombic (the orthorhombic cell has the same symmetry as trigonal cell).

Example 1: Let us consider a 2-dimensional case. A rectangular lattice and a centred-rectangular lattice share the same point-symmetry properties (check!), and belong to the same system. The translational symmetry is however different.

Example 2: The cubic system gives rise to three Bravais lattices, which share the cubic conventional cell but differ for the number of lattice points per cell and the primitive translation vectors (Fig. 4.4). The following table lists the symbols and the primitive vectors of the three cubic lattices.

	Simple cubic sc (P)	Body centred cubic bcc (I)	Face centred cubic fcc (F)
$\vec{a}_1 =$	$a\hat{x}$	$(a/2)(-\hat{x}+\hat{y}+\hat{z})$	$(a/2)(\hat{x}+\hat{y})$
$\vec{a}_2 =$	$a\hat{y}$	$(a/2)(+\hat{x}-\hat{y}+\hat{z})$	$(a/2)(\hat{y}+\hat{z})$
$\vec{a}_3 =$	$a\hat{z}$	$(a/2)(+\hat{x}+\hat{y}-\hat{z})$	$(a/2)(\hat{x}+\hat{z})$

4.6.2 Symmetries of Crystal structures

A crystal structure is obtained by adding a basis to each point of a Bravais lattice. Only a finite number of point groups and space groups are possible for crystal structures.

Let us again first consider *a)* only the points symmetries and then *b)* the full space symmetries (including translations).

Table 4.6: The 7 crystal systems, the 7 lattice systems and the 14 Bravais lattices and the repartition of 230 space groups of crystal symmetry. The ravais lattices are shown in Fig. 4.4.

7 Crystal systems	7 Lattice systems	Space groups	14 Bravais lattices
Cubic	Cubic	36	Simple cubic Cubic body centred Cubic face centred
Hexagonal	Hexagonal	27	
Trigonal	Hexagonal	18	Simple
	Orthorhombic	7	Simple
Tetragonal	Tetragonal	68	Simple Body centred
Orthorhombic	Orthorhombic	59	Simple Body centred Face centred Base centred
Monoclinic	Monoclinic	13	Simple Face centred
Triclinic	Triclinic	2	Simple

a) Point group transformations in crystal structures

One can show that all the possible crystals structures can be classified in *32 point symmetry groups* (sometimes called 32 crystallographic classes; don't confuse with the classes of group theory). Each one of the 32 point symmetry groups is characterised by a symbol. Two different notations are currently in use:

1. The Schönflies symbols, which have been introduced in § 4.3 and are used throughout these lecture notes, are particularly intuitive for the classification of point groups. They are mainly used in the Physics and Chemistry of molecules. They are less useful when considering also translational symmetry of crystals, since it is not possible to construct Schönflies symbols for space groups in a logical way.
2. The International (or Hermann-Mauguin) symbols are preferentially used by crystallographers, since they can account in a logical way not only for point symmetries but even for space symmetries. Their interpretation can however be less intuitive than the interpretation of the Schönflies symbols.
   Basically, the International symbols distinguish rotations n (where n=1, 2, 3, 4, 6), reflections m, inversion $\overline{1}$ and roto-inversions $\overline{n}$. Roto-inversion, the combination of a rotation and inversion, is not present in Schönflies simbols; by converse, roto-reflection is not present in International symbols. A more complete account is given in Section 4.8.

The number of point symmetry groups corresponding to each crystal system is given in Table 4.5.

Of the 32 point symmetry groups:

- 27 are low symmetry groups: only one C_N axis is present with $N > 2$.
- 5 are high symmetry groups compatible with the translational symmetry: O_h, O, T_d, T, T_h.

b) 230 space groups of crystal structures

When all possible symmetry transformations are considered for crystals structures, say when translations are added to point-symmetry transformations, *230 space groups* are identified. The number of space groups corresponding to each lattice system is given in Table 4.6.
Two different types of space groups are distinguished:

1. 73 space groups, called *symmorphic*, include only translations equal to Bravais lattice vectors. As already observed, the composition law (4.19) is more complicated than the composition law (4.12) of a group obtained by a simple direct product. A symmorphic space group can be obtained as semi-direct product of the point group and of the translation group. The point group is a subgroup of the symmorphic space group. Since translations represent an invariant subgroup, the point group is isomorphic to the factor group of which the translation subgroup is the divisor.

2. 157 space groups, called *non symmorphic*, include also two additional types of transformations:
   a) rotations C_n + translations a/n around and along *screw axes*
   b) reflections σ + translations a/n through and along *glide planes*

   These two types of transformations cannot be obtained as the product of a transformation of the point group and a transformation of the translation subgroup. The space group is not the direct product of the point group and the translation group.

Example: Let us consider a simple one-dimensional example of a non-symmorphic group (Fig. 4.5). Let $\mathcal{R} = \{E, \sigma\}$ be a point group containing only the identity E and the mirror symmetry σ and let $\mathcal{T} = \{0, a, 2a, 3a, \ldots\}$ be the translation group of period a (Fig. 4.5, top).
The direct product $\mathcal{D} = \mathcal{R} \times \mathcal{T}$ (Fig. 4.5, middle) has translational period a.
In (Fig. 4.5, bottom), a symmetry transformation is represented where the translation is $a/2$, not belonging to the translation subgroup $\mathcal{T}$. The symmetry transformation doesn't belong to the direct product $\mathcal{D}$.

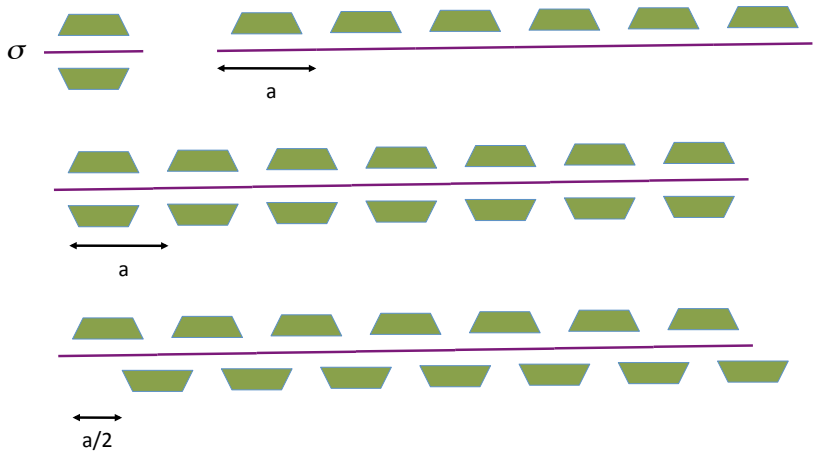

Figure 4.5: Top: mirror symmetry (left) and translation symmetry (right). Middle: product group. Bottom: subgroup of glide plane symmetry.

4.7 Symmetry breaking

The previous classifications of point and translation groups allows one to evaluate whether a system is more or less symmetric than another system.
The higher is the number of symmetry elements and the higher the index N of the symmetry axes C_N, the higher is the symmetry of an object or of a system.
Many phenomena, spontaneous or induced, are characterised by a modification of the symmetry properties. When the symmetry is reduced, one speaks of "symmetry breaking".

Example 1: When a stone is thrown into the water of a pond, a train of circular waves starts propagating from the impact point. The perfect translational symmetry of the water surface

is transformed into a lower circular symmetry. The initial symmetry has been broken by the impact of the stone.

Example 2: Compare the symmetry elements of the cubic and tetragonal cells (fig. 4.4).
The cubic cell has three equal edges a. The tetragonal cell is obtained by expanding or compressing one of the edges of the cube, so that there are two edges a and one edge b. The symmetry elements of the two cells are listed below, where the corresponding elements are in the same column:

cubic	$3C_4$	$4C_3$	$6C_2$	$3\sigma_h$	$6\sigma_v$	I
tetragonal	$1C_4, 2C_2$	–	$6C_2$	$3\sigma_h$	$2\sigma_v$	I

Some symmetry elements of the cube are lost in going from the cubic to the tetragonal cell. The cubic cell has a higher symmetry than the tetragonal cell.
The transition from the cubic lattice to the tetragonal lattice corresponds to a symmetry breaking event

Induced and spontaneous symmetry breaking

Symmetry breaking can be induced or spontaneous. For example, with reference to the cubic to tetragonal transition:

a) if the transition cubic → tetragonal is induced by an anisotropic stress, e.g a variation of pressure applied in a given direction which gives rise to a modification of the cube edges along that direction, one speaks of an *induced* symmetry breaking;
b) if the transition cubic → tetragonal is induced by an isotropic stress, e.g. by a variation of the external idrostatic pressure, or by a variation of temperature, one speaks of a *spontaneous* symmetry breaking (see for example Fig. 4.6).

Example 1: A phase transition liquid → crystalline solid induced by a reduction of the temperature is characterised by a lowering of symmetry, from the perfect isotropy of the liquid state to the crystalline arrangement of atoms. It corresponds to a spontaneous symmetry breaking.

Example 2: A drop of milk vertically falls on a bowl full of milk. If the experiment is carefully performed, one can observe that the splash has a crown-like shape, with exactly 24 equally spaces spikes. The initial C_∞ circular symmetry of the falling drop is modified into a lower C_{24} symmetry. It's again a spontaneous symmetry breaking.

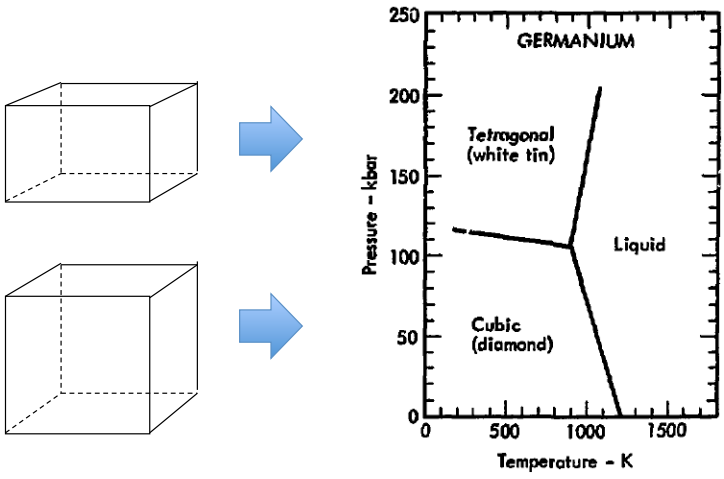

Figure 4.6: Schematic $p-T$ phase diagram of Germanium. At about 120 kbar, the cubic structure transforms to the lower symmetry tetragonal structure. Note the anomalous (water-like) behaviour of the liquid solid equilibrium line at low pressures.

Solid-solid phase transitions

The phase diagram of simple substances is generally characterised by the presence of various solid phases of different crystalline structure. In general, lowering the temperature or increasing the pressure leads to a lowering of the crystal symmetry (symmetry breaking). The example of Germanium is shown in Fig. 4.6: at room temperature, below about 120 kbar the symmetry is cubic, above 120 kbar the symmetry is tetragonal.

The structural symmetry of crystals is connected to the symmetry properties of atomic vibrations and electronic states, and with the onset of phenomena such as ferro-magnetism or ferro-electricity. For example, Ge at low pressure has the open cubic structure of diamond and is a semiconductor, while in the more densely packed orthorhombic phase it exhibits metallic behaviour.

The availability of high-pressure devices, such a the diamond anvil cell, the refinement of characterisation probes, such as X-ray diffraction with Synchrotron Radiation and the development of theoretical tools, such as the density functional theory (DFT) are allowing a deeper insight on high-pressure crystalline phases and on their physical properties. Pressures as high as 100 GPa = 1Mbar can be obtined in the laboratory, simulating the high pressures naturally present inside planets and stars.

The behaviours that are emerging from experiments and theoretical investigations performed on different systems are much more complex than the one schematically depicted in Fig. 4.6 for Ge (dating back to 1975).

Curie Principle and symmetry breaking

In 1894, the French physicist Pierre Curie proposed a principle, whose simplest formulation is "the symmetries of the causes are to be found in the effects".

The Curie principle seems to be contradicted by the spontaneous symmetry breaking phenomena. Actually, this is not the case, if the Curie principle is properly interpreted, as a few examples can explain.

Example 1: Let's consider the liquid → crystal transition induced by a reduction of temperature. The isotropic symmetry is broken in going from the liquid to a single mono-crystal. However, the orientation of a single crystal is a-priori unpredictable; the original symmetry of the liquid state is preserved in the isotropic possible orientation of single crystals.

Example 2: Let's consider the cubic → tetragonal transition. It is unpredictable in which one of the three spatial direction will the cube be deformed; the original cubic symmetry is preserved in the equivalence of the three possible direction of deformation of the cubic cell to originate the tetragonal cell.

4.8 Complements on group symbols

4.8.1 Theorems on symmetry elements

Symmetry elements (axes, mirror planes, inversion centre) are not completely independent. The presence of some symmetry elements implies the presence of other elements. The relations between symmetry elements are given by the following theorems (whose demonstration can be found in specialised textbooks).

The knowledge of the relations among symmetry elements are fundamental to interpret the synthetic symbols of both the Schönflies and the International (Hermann-Mauguin) systems.

Theorems on point symmetry elements

Theorem 1: If a 2-fold symmetry axis is perpendicular to an n-fold symmetry axis, there would be on the whole n 2-fold axes perpendicular to the n-fold axis.

Theorem 2: The line of intersection of two symmetry planes is always a symmetry axis with an angle of rotation twice as large as the angle between the planes.

Theorem 3: The intersection point of an even-fold symmetry axis with a symmetry plane perpendicular to it is the centre of symmetry.

Theorem 4: If a symmetry plane passes through an n-fold symmetry axis, there will be n such planes.

Theorem 5: The resultant of two intersecting symmetry axes is a third axis passing through the point of their intersection.

Theorem 6: If symmetry planes pass through an even-fold inversion axis, 2-fold symmetry axes will be located between them.

Theorems on space symmetry

Theorem 7: A consecutive reflection in two parallel symmetry planes is equivalent to a translation by the parameter $T = 2d$, where d is the inter-planar distance.

Theorem 8: A symmetry plane and a translation $\vec{T}$ perpendicular to it generated new ("inserted") symmetry planes parallel to the generating one and removed from it by a distance $d = T/2$.

Theorem 9: A symmetry plane and a translation $\vec{T}$, making up an angle α with the plane, generate a glide reflection plane parallel to the generating plane and removed from it in the direction of translation by $(T/2)\sin\alpha$.

Theorem 10: A symmetry axis with a rotation angle α and a translation $\vec{T}$ perpendicular to the axis generate a similar symmetry axis parallel to the first one , removed from it by $(T/2)\sin\alpha$ and located on a perpendicular to translation T in its middle.

Theorem 11: A screw symmetry axis with the angle of rotation α and displacement T_1 and a translation $\vec{T}$ perpendicular to the given axis generate a screw axis having the same angle and the same displacement. It is parallel to the given one, spaced from it by $(T/2)\sin(\alpha/2)$ and located in the middle of a perpendicular to the translation $\vec{T}$.

Theorem 12: A symmetry axis with the angle of rotation α and a translation $\vec{T}$ making up an angle β with the axis generate a screw symmetry axis.

Theorem 13: A screw symmetry axis with a rotation angle α and displacement T_1 and a translation $\vec{T}$ making an angle β with the axis, generate a screw symmetry axis having the same angle of rotation.

Theorem 14: An inversion-rotation axis with a rotation angle α and a translation $\vec{T}$ perpendicular to it generate the same inversion-rotation axis parallel to the generating one.

Theorem 15: An inversion-rotation axis with a rotation angle α and a translation $\vec{T}$ that makes an angle β with the axis generate an inversion axis having the same rotation α parallel to the given one.

4.8.2 The International (Hermann-Mauguin) symbols

Symbols for point symmetry elements

The following symbols are used for point symmetry elements.

a) n is an n-fold rotation axis, $n = 1, 2, 3, 4, 6$
b) $\overline{n}$ is an n-fold roto-inversion axis, say the combination of a rotation with an inversion, $\overline{n} = \overline{1}, \overline{3}, \overline{4}, \overline{6}$
$\overline{1}$ is the inversion point; the symbol $\overline{2}$ corresponds to m and is not used

c) m is a mirror plane
d) nm is an n-fold axis with n symmetry planes passing through it, e.g. $3m, 4m$
e) n/m is an n-fold axis with a symmetry plane perpendicular to it, e.g. $3/m$
if n is even, there is also a centre of symmetry (Theorem 3)
f) $n2$ is an n-fold axis with n two-fold axes perpendicular to it, e.g. $32, 42$
g) $\frac{n}{m}\, m$ or n/mmm is an n-fold axis with planes parallel and perpendicular to it

Table 4.7: The 32 crystallographic point groups divided according to the crystallographic The second and third column give the Schönflies and the International (Hermann-Mauguin) symbols, respectively.

System	Schoenflies	International	
Triclinic	C_1, C_i	$1, \overline{1}$	
Monoclinic	C_2, C_{1h}, C_{2h}	$2, m, 2/m$	
Orthorhombic	D_2, C_{2v}, D_{2h}	$222, mm2, mmm$	
Tetragonal	C_4, S_4, C_{4h}	$4, \overline{4}, 4/m$	
	$D_4, D_{2d}, C_{4v}, D_{4h}$	$422, \overline{4}2m, 4mm, 4/mmm$	
Trigonal	$C_3, C_{3i}, D_3, C_{3v}, D_{3d}$	$3, \overline{3}, 32, 3m, \overline{3}$	
Hexagonal	C_6, C_{3h}, C_{6h}	$6, \overline{6}, 6/m$	
	$D_6, D_{3h}, C_{6v}, D_{6h}$	$622, \overline{6}2m, 6mm, 6/m, mm$	
Cubic	T	23	Tetrahedron proper rot.
	T_d	$\overline{4}3m$	Tetrahedron symmetry
	$T_h = T \otimes C_i$	$m\overline{3}$	Full tetrahedral symmetry
	O	43	Octahedron proper rot.
	$O_h = O \otimes C_i$	$m\overline{3}m$	Full octahedral symmetry

General rules for the point-group symbols

The International symbols for point groups are sequences of symbols of symmetry elements, organised according to the following rules.

1. The symbol of a point-group contains at most three symbols of symmetry elements.
Only the "generating" symmetry elements are written.
The whole set of symmetry elements can be obtained from the listed "generating" elements by the theorems of the previous Subsection.
2. The order in which the symmetry elements are written is of the utmost importance: the meaning of a symbol of a symmetry element depends on its position in the group symbol.

The Schönflies notation and Hermann-Mauguin symbols of the 32 crystallographic point-groups are listed in Table 4.7.

Example: the point groups of the cubic system

The cubic crystal system supports five point-symmetry groups, which are listed in the bottom part of Table 4.7.

In the Schöenflies notation one distinguishes the three tetrahedral groups T, T_h, T_d and the two octahedral groups O, O_h.

In the International notation, the five groups are uniquely characterised by 3 or $\overline{3}$ in the second position, corresponding to the four 3-fold rotation axes typical of the cubic system.
The first element represent symmetry elements parallel to the coordinate axes (parallel to the cubic edges).

The third element, when necessary, represents symmetry elements diagonal with respect to the coordinate axes.
Considering in more detail the five groups:

a) 23 means the presence of two-fold axes parallel to the cube edges
b) $\bar{4}3m$ means the presence of 4-fold roto-inversion axes parallel to teh cube edges and of mirror planes diagonal with respect to teh cube edges
c) $m\bar{3}$ means the presence of mirror planes parallel to the cube edges
d) 43 means the presence of 4-fold axes paralle to the cube edges
e) $m\bar{3}m$ means the presence of mirror planes both parallel and diagonal with respect to the cube edges

General rules for the space-group symbols

The Bravais-lattice symbol is added in front of the symmetry element symbols:

P = primitive	A = A-face centred	R = rhombohedral
F = all-face centred	B = B-face centred	H = hexagonal
I = body centred	C = C-face centred	

Some symmetry elements can be added or modified with respect to the point group symbol, to take into account glide-planes and screw axes.

1. Screw axes are represented by n_k, where n is the order of the rotation axis and $1 \leq k < n$: when the axis rotates by $2\pi/n$, at the same time it moves by k/n of its translation vector.
2. A glide plane combines a mirror reflection with a translation of usually 1/2 of the lattice translation parallel to the plane.
   Glide planes along the $\vec{a} = \vec{a}_1$, $\vec{b} = \vec{a}_2$ and $\vec{c} = \vec{a}_3$ are represented by a, b and c, respectively.
   Double glide planes (two simultaneous translations) are represented by e.
   Diagonal glide planes are represented by n.
   A diagonal glide plane with gliding vector = 1/4 of the diagonal translation is represented by d (from the diamond structure).

Example: The point-group symbol is the same for the fcc and the diamond structures (copper and silicon, for example), say $m\bar{3}m$ (Schönflies O_h).
The space groups are different: $Fm\bar{3}m$ for fcc, $Fd\bar{3}m$ for diamond: the d in the diamond symbol means the presence of a diagonal glide plane. Schönflies distinguishes the two space groups by a conventional apex without any physical meaning: O_h^5 and O_h^7).

4.9 Bibliography of Chapter 4

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions).
  Chapter 7 (Classification of Bravais lattices and crystal structures).
- P. Atkins and J. De Paula: *Physical Chemistry*, 9th ed., W.H. Freeman 2009.
  Chapter 11 (Molecular symmetry).
- P. Atkins: *Chimica Fisica* [in Italian], Zanichelli 1981.
  Cap. 16 (La simmetria : rappresentazione e conseguenze).
- C. Giacovazzo editor: *Fundamentals of Crystallography*, 2002.
  Chapter 1 (Applications of symmetry groups to crystals)
- *International Tables for Crystallography*, Volume A: *Space-Group Symmetry*, ed. by T. Hahn, Springer 2005.
- D. C. Harris and M. D. Bertolucci: *Symmetry and Spectroscopy*, Oxford University Press 1978, reprinted by Dover Publications 2014.
  Chapter 1 (A chemist's view of group theory).

- L. Landau and E. Lifshitz: *Theoretical Physics* (various editions in different languages).
  Vol. 3: *Quantum Mechanics*, Chapter 12 (Point symmetry of molecules).
  Vol. 5: *Statistical Physics*, Chapter 13 (Symmetry of crystals).
- J. C. Slater: *Quantum theory of matter*, McGraw-Hill 1978.
  Chapter 25 (Group theory and wave functions symmetry).
- M. Tinkham: *Group theory and quantum mechanics*, Mc Graw Hill 1964, Dover 1992.
  Chapter 2 (Abstract group theory).
- B.K. Vainshstein: *Modern Crystallography*, Springer 1981.
  Vol. 1, Chapter 2 (Fundamentals of the theory of symmetry).
- Z. Dauter and M. Jaskolki: *How to read volume A of the International Tables for Crystallography: and introduction for nonspecialists*, Journal of Applied Crystallography, **43**, 1150-1170 (2010).

Chapter 5

Representations of symmetry groups

In the previous Chapter 4 we have seen how the group theory applied to the symmetry operations allows an exhaustive classification of molecular and crystalline structures.
The possible applications of symmetry groups are actually much larger than these simple classifications. This Chapter 5 is dedicated to an introduction to these further applications.

To fully exploit the potential of group theory, one needs to find mathematical representations of the abstract groups, in our case of the symmetry groups. Symmetry group representations consist in isomorphisms or homomorphisms between the groups of symmetry transformations and suitable mathematical groups; the most effective representations are given by groups of matrices defined in suitable vector spaces (Section 5.1).
The most elementary representations are based on the coordinate vector space and consist in three-dimensional matrices (Section 5.2). Some simple examples show that the dimensionality of the matrices can in some cases be reduced by a careful choice of the orientation of the basis vectors. These considerations lead to the concept if Irreducible Representations (IR), whose basic theory is introduced in Section 5.3.
Some examples of irreducible representations are given in Section 5.4 and the generalisation of the theory of irreducible representations from the three-dimensional coordinate space to function spaces is made in Section 5.5.
The irreducible representations are particularly useful for treating the properties of physical systems which are invariant with respect to symmetry transformations. An important case is that of the Hamiltonian function, whose invariance properties with respect to symmetry transformations can be exploited to simplify a number of quantum mechanical problems (Section 5.6).
Important applications of irreducible representations concern the electronic structure of molecules (Section 5.7) and molecular complexes (Section 5.8) as well as the evaluations of superposition integrals (Section 5.9) and the selection rules for electronic transitions (Section 5.10).

5.1 General properties of group representations

A representation of a group $\mathcal{G}$ (in our case a group of symmetry transformations) is given by a set Γ of non-singular square matrices $\mathbf{D}$, say matrices with non-zero determinant, $|\mathbf{D}| \neq 0$.
Each matrix of the set Γ is associated to one or more elements of the group $\mathcal{G}$ in such a way that the matrix multiplication reproduces the composition law in $\mathcal{G}$:

$$\mathbf{D}(g_i)\,\mathbf{D}(g_j) \;=\; \mathbf{D}(g_i g_j)\,. \tag{5.1}$$

The set Γ of matrices has thus the properties of a group. The group Γ of matrices is said to be

- isomorphic to the group $\mathcal{G}$ of symmetry transformations if each matrix $\mathbf{D}$ is associated to only one element g of $\mathcal{G}$ and vice-versa; in this case, the representation is said to be true or faithful;

- homomorphic to the group $\mathcal{G}$ if each matrix $\mathbf{D}$ is associated to more than one element g of $\mathcal{G}$; in this case, the representation is said to be unfaithful.

The matrix representation depends on the linear vector space on which the matrices $\mathbf{D}$ are defined: it can simply be the three-dimensional euclidean space of the (x, y, z) coordinates, but more generally it can be a space of functions of the coordinates. In some cases, even one-dimensional spaces are considered.

For an isomorphic representation

- the number of matrices, say the order of the group Γ, is equal to the order n of the group $\mathcal{G}$,
- the dimensionality of the square matrices (say the number of rows or of columns) is equal to the dimensionality of the linear vector space; it is called the dimensionality of the representation.

The representations of a given dimensionality are not univocally determined; by changing the basis of the linear vector space, the representation matrices change. Two representations $\mathbf{D}$ and $\mathbf{D}'$ are said to be equivalent if all their elements are connected by the same similarity transformation $\mathbf{S}$

$$\mathbf{D}' = \mathbf{S}^{-1}\,\mathbf{D}\,\mathbf{S}\,, \tag{5.2}$$

where $\mathbf{S}$ is an invertible matrix.
Particularly interesting are the similarity transformations (5.2) that lead to fully diagonalised matrices or matrices diagonalised by blocks, as will be seen in Section 5.3.

5.2 Representation in the three-dimensional Euclidean space

The most trivial representation of a finite point group is based on the transformation of a three-dimensional vector in the real space, already introduced in equation (4.13).
Let us choose an orthogonal reference frame (x, y, z), which corresponds to three basis vectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{5.3}$$

An orthogonal transformation modifies the coordinates of a point from $(x_1 y_1 z_1)$ to $(x_2 y_2 z_2)$. In vectorial notation, the transformation is represented by a 3×3 orthogonal matrix $\mathbf{D}$, according to the equation

$$\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} d_{xx} & d_{xy} & d_{xz} \\ d_{yx} & d_{yy} & d_{yz} \\ d_{zx} & d_{zy} & d_{zz} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \tag{5.4}$$

The isomorphism associates the n transformations of a point symmetry group $\mathcal{G}$ to the n 3×3 matrices $\mathbf{D}$ of the group Γ with the following properties.

a) To the composition of symmetry operations in $\mathcal{G}$ there corresponds the multiplication of matrices in Γ.

b) The identity element E (or e) is represented in the group Γ by the diagonal unit matrix $\mathbf{D}(e) = \mathbf{I}$, whose elements are

$$D_{ij} = \delta_{ij}\,, \quad \text{say} \quad D_{ii} = 1\,, \quad D_{ij} = 0 \;(i \neq j)\,. \tag{5.5}$$

c) The inverse of an element g in the group $\mathcal{G}$ is represented, in the group Γ, by the inverse matrix

$$\mathbf{D}(g^{-1}) = [\mathbf{D}(g)]^{-1}\,, \qquad \mathbf{D}^{-1} = \frac{\mathrm{Adj}(\mathbf{D})}{|\,\mathbf{D}\,|} \tag{5.6}$$

where Adj means adjoint matrix and $|\,\mathbf{D}\,|$ is the determinant.
A general demonstration of (5.6) can be found in specialised textbook, such as Byron and Fuller: *Mathematics of classical and quantum Physics*. A tutorial introduction can be found here in § 5.11.1.
For orthogonal transformations in the Euclidean space, the orthogonal inverse matrix is equal to the transposed matrix, $\mathbf{D}^{-1} = \mathbf{D}^T$.

d) For orthogonal transformations, the following property of the determinant holds

$$|\,\mathbf{D}\,| = \begin{cases} +1 & \text{for proper rotations (congruence)} \\ -1 & \text{for improper rotations (reflections)} \end{cases} \tag{5.7}$$

Since the determinant of a product of matrices is the product of the determinants,

a) a sequence of proper rotations always corresponds to a proper rotation,
b) a sequence of improper rotations (reflections, inversions) corresponds
   - to a proper rotation for an even number of improper rotations,
   - to an improper rotation for an odd number of improper rotations.

e) In the language of group representation, the *trace* of the matrix $\mathbf{D}(g)$ that represents a symmetry transformation g is called the *character* $\chi(g)$ of the symmetry transformation:

$$\chi(g) \;=\; \sum\nolimits_i D_{ii}(g)\,. \tag{5.8}$$

As it will be shown in the following, the characters play a fundamental role in the theory of group representation.

5.2.1 Example: coordinate representation of the C_{3v} group

Let us consider the coordinate representation of the C_{3v} group (for which a physical example is the ammonia molecule NH_3, Fig. 5.1 left).
We look for the six matrices that represent the transformations of the coordinates (x, y, z) of any point as a consequence of the six symmetry operations of the C_{3v} group, according to (5.4). In principle, we can arbitrarily choose the orientation of the three basis vectors with respect to the symmetry elements of the C_{3v} group; the six transformations matrices of (5.4) are different for different choices of the orientation of the basis vectors.
It is however evident that one choice of the basis vectors is particularly convenient, say the choice where one axis, e.g. the z axis, corresponds to the C_3 rotation axis. The representation of the C_{3v} group with respect to this basis system is connected to the representation with respect to any other basis system by a similarity transformation (5.2).
For more concreteness, let us choose the reference frame (x, y, z) not only with the z axis corresponding to the C_3 rotation axis but also with the y axis contained in the σ_v' mirror plane. The two σ_v'' and σ_v''' mirror planes form angles of $\pm\pi/6$ with the x axis (Fig. 5.1 centre and right).

Let us now determine the six matrices $\mathbf{D}$ that represent the six point symmetry transformations g of the group C_{3v} singled out in § 4.3. For each matrix, we will calculate the determinant and the trace, say the character of the corresponding symmetry transformation.

Identity e (E)

The identity transformation transforms $x_1 \to x_2, y_1 \to y_2, z_1 \to z_2$. It is represented by a diagonal matrix:

$$\mathbf{D}(e) \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.9}$$

The determinant is $|\mathbf{D}(e)| = 1$. The character is $\chi(e) = 3$.

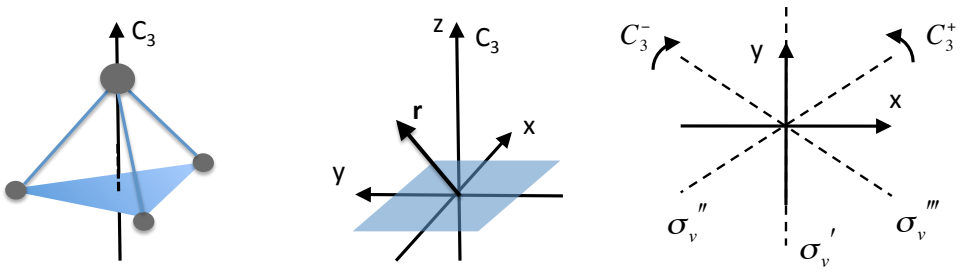

Figure 5.1: Right: the ammonia molecule NH_3 (stereo projection). Centre: the xyz reference system chosen for the coordinate representation of the C_{3v} group and the generic $\vec{r}$ position whose transformations are considered. Right: the xy axes and the symmetry planes.

Rotations $C_3^+ = C_3^1$ and $C_3^- = C_3^2$

With the peculiar choice of the z axis parallel to C_3, the z coordinate is unchanged by rotations. The transformations of the coordinates x, y for the anticlockwise or clockwise rotation of a vector by an angle θ are, respectively:

$$\begin{cases} x_2 = x_1 \cos\theta - y_1 \sin\theta \\ y_2 = x_1 \sin\theta + y_1 \cos\theta \end{cases} \qquad \begin{cases} x_2 = x_1 \cos\theta + y_1 \sin\theta \\ y_2 = -x_1 \sin\theta + y_1 \cos\theta \end{cases} \tag{5.10}$$

For $C_3^+ = C_3^1$, the angle is $\theta = 2\pi/3$, for $C_3^- = C_3^2$, the angle is $\theta = 4\pi/3$, so that

$$\mathbf{D}(C_3^1) = \begin{pmatrix} -1/2 & -\sqrt{3}/2 & 0 \\ +\sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad \mathbf{D}(C_3^2) = \begin{pmatrix} -1/2 & +\sqrt{3}/2 & 0 \\ -\sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.11}$$

The determinants ares $|\mathbf{D}(C_3^1)| = |\mathbf{D}(C_3^2)| = +1$, equal for the two rotations.
The character $\chi(C_3^1) = \chi(C_3^2) = 0$ is the same for the two rotations, which belong to the same class.

Reflections σ_v

With the peculiar choice of the z axis parallel to C_3, the z coordinate is unchanged by mirror transformations. The evaluation of the matrix $\mathbf{D}(\sigma_v')$ representing the reflection with respect to the plane containing the y axis is straightforward ($y_2 = y_1, x_2 = -x_1$):

$$\mathbf{D}(\sigma_v') = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.12}$$

For the reflections σ_v'' and σ_v''', it is convenient to consider two new frames (x'', y'') and (x''', y'''), rotated by $2\pi/3$ and $-2\pi/3$, respectively, with respect to the original frame (x, y) (Fig. 5.2). With respect to these new frames, the matrices $\mathbf{D}''(\sigma_v'')$ and $\mathbf{D}'''(\sigma_v''')$ are equal to $\mathbf{D}(\sigma_v')$ of (5.12). The matrices in the original frame (x, y) are obtained by the products

$$\mathbf{D}(\sigma_v'') = \mathbf{R}^{-1}\, \mathbf{D}''(\sigma_v'')\, \mathbf{R}; \qquad \mathbf{D}(\sigma_v''') = \mathbf{R}\, \mathbf{D}'''(\sigma_v''')\, \mathbf{R}^{-1}, \tag{5.13}$$

where $\mathbf{R}$ is the matrix of counter-clockwise rotation by $2\pi/3$. The final result is

$$\mathbf{D}(\sigma_v'') = \begin{pmatrix} 1/2 & \sqrt{3}/2 & 0 \\ +\sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \qquad \mathbf{D}(\sigma_v''') = \begin{pmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ -\sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.14}$$

The determinants are $|\mathbf{D}(\sigma)| = -1$, equal for the three mirror transformations.
The character $\chi(\sigma_v') = \chi(\sigma_v'') = \chi(\sigma_v''') = 1$ is the same for the three mirror transformations, which belong to the same class.

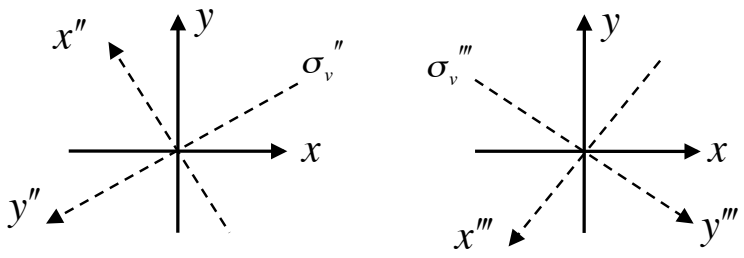

Figure 5.2: The new reference systems $(x'','')$ and (x''', y''') used to calculate the matrices of the σ_v'' and σ_v''' transformations.

Comments

1. One can easily check that the six matrices fulfil the requirements of a faithful representation, since each matrix **D** is associated to only one element of the symmetry group and vice-versa.
2. Transformations of the same class share the same character:

$$\chi(e) = 3\,, \quad \chi(C_3) = 0\,, \quad \chi(\sigma_v) = 1\,. \tag{5.15}$$

3. The representation of the C_{3v} group here considered has dimensionality 3 (three-dimensional coordinate basis and 3×3 matrices). Other representations of the C_{3v} group are possible, of the same dimensionality as well as of smaller or larger dimensionalities.
4. The determinant of the transformation matrices is $+1$ for identity and rotations, say for the operations of direct congruence (proper isometric correspondence), it is -1 for the mirror reflections, say for the operations of opposite congruence (improper isometric correspondence).

5.2.2 Diagonal structure of the matrices of the C_{3v} group

In the three-dimensional representation (x, y, z) of the C_{3v} group, with the particular choice of the z axis coinciding with the C_3 rotation axis, all the six matrices are "blocked": two non-zero square blocks can be singled out along the main diagonal, the remaining elements being zero:

$$\mathbf{D} \;=\; \left(\begin{array}{cc|c} D_{xx} & D_{xy} & 0 \\ D_{yx} & D_{yy} & 0 \\ \hline 0 & 0 & 1 \end{array}\right) \tag{5.16}$$

As a consequence, one can consider two other independent representations of the C_{3v} group:

1. *A two-dimensional representation in the xy basis.*
   The six two-dimensional matrices are different for different transformations. The group of matrices is isomorphic to the group of symmetry transformations (faithful representation). The transformations mix up the x and y coordinates.
   Again, transformations of the same class share the same character, which is however different from the character of the 3-dimensional representation:

$$\chi(e) = 2\,, \quad \chi(C_3) = -1\,, \quad \chi(\sigma_v) = 0\,. \tag{5.17}$$

2. *A one-dimensional representation in the z basis.*
   All the six 1×1 matrices are equal to 1. The z coordinate is unaffected by any symmetry transformation of the group C_{3v}. The group of matrices is homomorphic to the group of symmetry transformations. All symmetry transformations have the same character

$$\chi(e) = \chi(C_3) = \chi(\sigma_v) = 1. \tag{5.18}$$

This apparently trivial representation, named "identical representation", is clearly unfaithful; however, as we will see later on, it is useful when treating physical properties that are invariant with respect to the symmetry transformations.

Let us stress again that the particular diagonal structure (5.16) of the representation matrices depends on the choice of the basis vectors (or better of their orientation with respect to the symmetry elements of the system).

The question naturally arises whether this matrix structure (5.16) can be further "reduced", leading to fully diagonal matrices, by a suitable choice of basis vectors. If not, the two-dimensional representation in the basis xy should be considered as "irreducible". The problem of how to decide whether a representation is reducible or not will be solved in the next Section 5.3. We anticipate here that the two-dimensional representation of the C_{3v} group cannot be further decomposed and is thus irreducible.

Other irreducible representations of the C_{3v} group?

Another *one-dimensional representation* of the group C_{3v}, independent of the previous ones, can be obtained by considering the determinants of the six matrices:

$$|\mathbf{D}(E)| = |\mathbf{D}(C_3^1)| = |\mathbf{D}(C_3^2)| = 1\,, \qquad |\mathbf{D}(\sigma_v')| = |\mathbf{D}(\sigma_v'')| = |\mathbf{D}(\sigma_v''')| = -1\,. \tag{5.19}$$

This one-dimensional representation is formed by two one-dimensional matrices, $\mathbf{D}_1 = 1$ and $\mathbf{D}_2 = -1$, each one of them corresponding to three different symmetry transformations. One can easily check that the requirements of group representations are fulfilled: the multiplication of the one-dimensional matrices corresponds to the composition of the symmetry elements, the identity element is $\mathbf{D}_1 = 1$, the inverse elements are $[\mathbf{D}_1]^{-1} = \mathbf{D}_1$ and $[\mathbf{D}_2]^{-1} = \mathbf{D}_2$ (make reference to the multiplication table 4.4).
It is however evident that this representation cannot be simply connected to the coordinate transformations. A possible interpretation will be given later on in terms of functions of coordinates.
In the homorphism between the group C_{3v} and this representation, the unit matrix $\mathbf{D}_1$ corresponds to the invariant subgroup $\{E, C_3^1, C_3^2\}$, the other matrix $\mathbf{D}_2$ to its coset $\{\sigma_v', \sigma_v'', \sigma_v'''\}$. The two matrices $\mathbf{D}_1$ and $\mathbf{D}_2$ form a faithful representation of the factor group of the invariant subgroup.

A new question arises: how many independent irreducible representations characterise a given symmetry group? The problem will be solved in the next Section 5.3 too. We anticipate here that the three representations found up to now (one two-dimensional and two one-dimensional) are the only possible irreducible representations for the C_{3v} group.

Problem: Consider the matrices $|\mathbf{D}(E)| = 1$, $|\mathbf{D}(C_3^1)| = |\mathbf{D}(C_3^2)| = -1$ and $|\mathbf{D}(\sigma_v')| = |\mathbf{D}(\sigma_v'')| = |\mathbf{D}(\sigma_v''')| = 1$ and verify that they don't are a representation of the C_{3v} group.

5.2.3 Example: coordinate representation of the C_6 group

As a further example, let us now study the cyclic group C_6, already considered in Section 4.3.
One can easily see that, if again the z axis is chosen parallel to the C_6 axis, the six 3×3 matrices representing the 6 symmetry transformations are

$$\mathbf{D}(C_6^k) = \left(\begin{array}{cc|c} \cos(k\pi/3) & -\sin(k\pi/3) & 0 \\ \sin(k\pi/3) & \cos(k\pi/3) & 0 \\ \hline 0 & 0 & 1 \end{array}\right) \tag{5.20}$$

where $k = 1, 2, 3, 4, 5, 6$.
The determinant of all the six matrices is $+1$ (proper rotations).
The characters of the six transformations (each of which forms a separate class) are however different, varying from $\chi = 3$ for $C_6^6 = e$ to $\chi = 2$ for C_6^1 and C_6^5 to $\chi = 0$ for C_6^2 and C_6^4 to $\chi = -1$ for C_6^3.

The matrices (5.20) are blocked, so that we can distinguish a two-dimensional representation xy and a one-dimensional representation z as for the C_{3v} group.

The question again arises whether this matrix structure can be further "reduced" or is instead "irreducible". We anticipate here that the two-dimensional representation of the C_6 group can be further decomposed into two one-dimensional representations. In Section 5.5 we will find that there are actually six independent irreducible representation of the C_6 group; to find them, we have however to consider a vector space of functions.

Problem: Consider the two-dimensional representation xy of the C_6 group, evaluate the character of the different transformation and discuss the results.

5.3 Irreducible representations: theory

In the previous Section, the concept of irreducible representation arose from the analysis of some working examples. Let us now give a more rigorous foundation.

5.3.1 Similarity transformations and irreducible representations

1.
The dimensionality of a given representation depends on the dimensionality of the linear vectorial space where the matrices are defined. Within a given vectorial space, similarity transformations (5.2) lead to equivalent representations, since they leave the matrix equations unchanged.
Let $\mathbf{D}$ and $\mathbf{D}'$ be the representation matrices before and after a similarity transformation $\mathbf{S}$. It is easy to show that

$$\mathbf{D}'(g_1)\,\mathbf{D}'(g_2) = \left[\mathbf{S}^{-1}\,\mathbf{D}(g_1)\,\mathbf{S}\right]\,\left[\mathbf{S}^{-1}\,\mathbf{D}(g_2)\,\mathbf{S}\right] = \left[\mathbf{S}^{-1}\,\mathbf{D}(g_1)\,\mathbf{D}(g_2)\,\mathbf{S}\right] = \mathbf{D}'(g_1 g_2)\,. \tag{5.21}$$

Here and in the following, similarity transformations are represented by unitary matrices $\mathbf{S}^{-1} = \mathbf{S}^\dagger$ in order to guarantee the possibility of operating in complex spaces.

2.
Given any two representations of a group, $\mathbf{D}^{(1)}(g_i)$ of dimensionality ℓ_1 and $\mathbf{D}^{(2)}(g_i)$ of dimensionality ℓ_2, it is always possible to combine the two set of matrices to obtain a new representation of dimensionality $\ell = \ell_1 + \ell_2$:

$$\left(\begin{array}{c|c} \mathbf{D}^{(1)}(g_i) & 0 \\ \hline 0 & \mathbf{D}^{(2)}(g_i) \end{array}\right) \;=\; \mathbf{D}(g_i)\,. \tag{5.22}$$

By similarity transformations it is possible to obtain an equivalent representation of dimensionality ℓ where the matrix elements are scrambled so that the block structure of (5.22) is lost.

3.
We are actually interested in the reverse operation, say in the decomposition of a given representation in all its possible irreducible component representations, by means of a suitable similarity transformation that acts on all the matrices of the group.
Otherwise stated, we are interested in the search of all the irreducible representations of a given group.
This task is facilitated by a number of theorems and rules which are considered below. (The demonstrations can be found in specialised textbooks, for example, in M. Tinkham: *Group theory and quantum mechanics*).

5.3.2 General orthogonality theorem

One can demonstrate that every matrix representation is equivalent, through a similarity transformation, to a representation by unitary matrices, say matrices such that $\mathbf{D}^{-1} = \mathbf{D}^\dagger$ (Schur-Auerbach theorem).

As a consequence, through a quite long procedure, one can further demonstrate the so-called orthogonality theorem, which assert that for every pair of *inequivalent* irreducible unitary representations $\mathbf{D}^{(\alpha)}$ and $\mathbf{D}^{(\beta)}$ of a group

$$\boxed{\sum_g \left[D^{(\alpha)}_{\mu\nu}(g)\right] \left[D^{(\beta)}_{\mu'\nu'}(g)\right]^* = \frac{n}{\ell_\alpha}\,\delta_{\alpha\beta}\,\delta_{\mu\mu'}\,\delta_{\nu\nu'}} \tag{5.23}$$

where

g labels the n elements of the symmetry group,
α and β label the two irreducible representations,
μ and ν label the rows and columns of the matrices,
ℓ_α is the dimensionality of the α irreducible representation.

Let us now consider some consequences of the orthogonality theorem.

Dimensionality of irreducible representations

To grasp the meaning of the theorem (5.23), it is convenient to consider an abstract n-dimensional vector space, whose dimension is equal to the number n of group elements.
Within this n-dimensional space, for each irreducible representation $\mathbf{D}^{(\alpha)}$ of dimensionality ℓ_α one can define ℓ_α^2 n-dimensional vectors $D^{(\alpha)}_{\mu\nu} = \{D^{(\alpha)}_{\mu\nu}(g_1), D^{(\alpha)}_{\mu\nu}(g_2), \dots D^{(\alpha)}_{\mu\nu}(g_n)\}$, one for each pair of values $\mu\nu$, whose n components are the $D_{\mu\nu}$ values corresponding to the n symmetry group elements.
The theorem (5.23) expresses the orthogonality of the vectors corresponding to the different ℓ_α^2 pairs $\mu\nu$ for a given representation ($\delta_{\mu\mu'}\,\delta_{\nu\nu'}$) or to different representations within the n dimensional space ($\delta_{\alpha\beta}$).
The total number of vectors whose orthogonality is guaranteed by (5.23) is $\sum_\alpha \ell_\alpha^2$, where the sum is over the possible irreducible representation.
Since the maximum number of mutually orthogonal vectors in an n-dimensional space is just n, the following relation holds between the representations dimensionality and the number of symmetry operations:

$$\sum\nolimits_\alpha \ell_\alpha^2 \;\le\; n\,. \tag{5.24}$$

Actually, one can demonstrate (see e.g. Tinkham, pagg. 30-31) that

$$\boxed{\sum\nolimits_\alpha \ell_\alpha^2 \;=\; n} \tag{5.25}$$

This relation is a constraint on the number and dimensionality of the irreducible representations of a group.

5.3.3 Characters

As already noted above, the traces of the matrices are named characters in the theory of group representations. The character of the matrix representing a given group element in a given representation, be it reducible or irreducible, depends on the dimensionality ℓ of the representation,

$$\chi(g) \;=\; \sum_{\mu=1}^{\ell} D_{\mu\mu}(g)\,, \tag{5.26}$$

and is invariant with respect to similarity transformations.
Symmetry operations within a class are connected by similarity transformations, expressed by the conjugation relation (Section 4.2). As a consequence, all elements of a class share the same character: we can thus speak of the *character of a class*.

A sum rule for characters

From the orthogonality theorem (5.23), one can see that, if α and β are irreducible representations,

$$\sum_g \left[\chi^{(\alpha)}(g)\right] \left[\chi^{(\beta)}(g)\right]^* \tag{5.27}$$

$$= \sum_g \left[\sum_{\mu=1}^{\ell_\alpha} D^{(\alpha)}_{\mu\mu}(g)\right] \left[\sum_{\mu'=1}^{\ell_\beta} D^{(\beta)}_{\mu'\mu'}(g)\right]^* = \frac{n}{\ell_\alpha}\,\delta_{\alpha\beta} \sum_{\mu\mu'} \delta_{\mu\mu'} = \frac{n}{\ell_\alpha}\,\delta_{\alpha\beta}\,\ell_\alpha\,, \tag{5.28}$$

whence

$$\boxed{\sum_{g=g_1}^{g_n} \chi^{(\alpha)}(g) \left[\chi^{(\beta)}(g)\right]^* = n\,\delta_{\alpha\beta}} \tag{5.29}$$

Equation (5.29) represents an orthogonality relation for the characters of irreducible representations.
For $\alpha = \beta$, say for a given irreducible representation, (5.29) asserts that the sum of the squares of the characters is qual to the number n of group elements:

$$\sum_{g=g_1}^{g_n} |\chi(g)|^2 = n\,, \tag{5.30}$$

For $\alpha \neq \beta$, say for two different irreducible representations, the sum is zero.

Note: The matrix elements $D_{\mu\nu}$ and the characters χ can be complex numbers, whence the asterisk * in (5.23) and (5.29).

Irreducible representations and classes

Since all the group elements in a given class share the same character, we can substitute the sum over the n group elements in (5.29) with the sum over the h classes of the group:

$$\sum_{k=1}^{h} \chi^{(\alpha)}(k) \left[\chi^{(\beta)}(k)\right]^* n_k = n\,\delta_{\alpha\beta}\,, \tag{5.31}$$

where k labels the classes and n_k is the number of group elements in the class k (we are always considering only inequivalent irreducible representations).

To better grasp the meaning of (5.31), it is convenient to consider an abstract h-dimensional vector space, whose dimension is equal to the number of classes of the group.
Within this space, for each irreducible representation one can define an h-dimensional vector $\chi^{(\alpha)} = \{\chi^{(\alpha)}(1), \chi^{(\alpha)}(2), \ldots \chi^{(\alpha)}(h)\}$. Eq. (5.31) expresses the orthogonality of the vectors corresponding to different irreducible representations (I.R.)
Since the maximum number of orthogonal vectors in an h-dimensional space is just h, the number of inequivalent I.R. cannot exceed the number h of classes.
Actually, one can demonstrate that

Number of irreducible representations = Number of classes.

Example 1: The symmetry group C_{3v}, of order $n = 3$, contains three classes: $\{e\}$, $\{C_3{}^+, C_3{}^-\}$, $\{\sigma'_v, \sigma''_v, \sigma'''_v\}$. The number of irreducible representations (I.R.) of C_{3v} is three.

Example 2: The symmetry group C_6, of order $n = 6$, is cyclic and thus commutative; each element is a class by itself, so the group contains six classes and the number of I.R. is six.

5.3.4 Table of characters

The properties of the irreducible representations of a given symmetry group are summarised in the *table of characters*, which lists the characters of the symmetry classes in the different irreducible representations of the group. The tables of characters play a fundamental role in the applications of the symmetry group theory which will be considered in this chapter.

The tables of characters of the point symmetry groups can be found in a number of textbooks (e.g. Landau-Lifshitz or Tinkham) or specialised monographs. The procedures for finding the I.R. of each symmetry group are in some cases far from simple, but are greatly facilitated by the sum rules introduced above.

To understand the structure of character tables, let us refer to the concrete case of the C_{3v} group, whose characters table is reproduced in Table 5.1.

1. Each line of the table is dedicated to a different irreducible representation.
2. The first column (not necessarily present) specifies the dimension of the irreducible representation.
3. The second column contains the conventional symbols of the irreducible representations. For example:
   - The symbols A and B characterise one-dimensional I.R., symmetric and antisymmetric, respectively, with respect to rotation. A_1 and A_2 refer to symmetry and antisymmetry with respect to mirror reflections, respectively.
   - The symbol E characterises two-dimensional I.R. (don't confuse with the symbol $E = e$ of the identity transformation).
   - The symbol T (not present in Table 5.1) characterises the three-dimensional I.R.
4. The third column (actually a set of columns) lists the characters of the different classes of symmetry transformations.
5. The last column lists some of the coordinates, or combinations of coordinates, or functions of the coordinates, which transform according to the given representation.
   In Table 5.1, R_x, R_y, R_z are the components of an axial vector.

Table 5.1: Character table of the irreducible representations (I.R.) of the symmetry group C_{3v}. The three classes contain one (e), two ($2C_3$) and three ($3\sigma_v$) elements, respectively.

		Classes			
		e	$2C_3$	$3\sigma_v$	
one-dim.	A_1	1	1	1	$z, x^2+y^2+z^2$
one-dim.	A_2	1	1	-1	R_z
two-dim.	E	2	-1	0	$(x,y),(xz,yz),(R_x,R_y)$

The A_1 and E irreducible representations appearing in Table 5.1 have been obtained in Section 5.2 from the 3×3 reducible representation in the coordinate basis.
The A_2 irreducible representation, which was obtained again in Section 5.2 from the determinants of the 3×3 matrices, can be understood by the following example: consider a body spinning around the z axis. The z component of the angular velocity vector $\vec{R}$ (axial vector) is unchanged by the e and C_3 symmetry transformations, while it is reversed by the σ_v symmetry transformations.

Note 1: The order of appearance of the different columns of a character table can be different in different presentations.

Note 2: In Section 4.3, we have seen that isomorphic symmetry groups have the same multiplication table. This is the case, for example, of the D_2, C_{2h} and C_{2v} groups, which are isomorphic to the Klein four-group. Isomorphic group share the character table too.

5.3.5 Applications of the sum rule for characters

The sum rule (5.29) is useful to verify whether a given representation is irreducible or not and can be of help in finding the characters of an unknown irreducible representation.
Let us consider some examples referring to the C_{3v} group.

Example 1: Let us consider the 3×3 representation of the C_{3v} group in the coordinate basis introduced in Section 5.2. The characters are: $\chi(e) = 3\,, \chi(C_3) = 0\,, \chi(\sigma) = 1\,$. The sum rule (5.29) gives

$$\sum_g |\,\chi(g)\,|^2 \;=\; 9 + 0 + 3 = 12\,.$$

The sum is different from $n = 6$. The 3×3 representation is thus reducible.

Example 2: In the coordinate basis we have seen that the 3×3 representation of the C_{3v} group, with the choice of the z axis corresponding to the C_3 axis, is naturally decomposed into two representations of smaller dimensionality; the characters of these two representations correspond to the characters of the A_1 and E irreducible representations of Table 5.1.
We can easily check that the sum rule (5.29) is satisfied for these two representations:

$$\begin{aligned} A_1: &\quad \textstyle\sum_g \left|\,\chi^{(A_1)}(g)\,\right|^2 \;=\; 1 + 2 + 3 = 6 \\ E: &\quad \textstyle\sum_g \left|\,\chi^{(E)}(g)\,\right|^2 \;=\; 4 + 2 + 0 = 6 \end{aligned}$$

Example 3: The group C_{3v} has three classes, so we expect three irreducible representations (I.R.). As we have seen, only two I.R. of C_{3v} can be easily obtained in the coordinate basis, A_1 and E. Suppose we don't know the third I.R.; let's call it β, for the moment. To fulfil (5.25), say

$$\ell_{A_1}^2 + \ell_E^2 + \ell_\beta^2 = 1 + 4 + \ell_\beta^2 = 6$$

the dimensionality of the β I.R. must be $\ell_\beta = 1$.
We can also find the characters of the β I.R. as follows. Since β is one-dimensional, $\chi^{(\beta)}(e) = 1$. The remaining two characters $\chi^{(\beta)}(C_3)$ and $\chi^{(\beta)}(\sigma_v)$ can be determined by calculating the sum rules (5.29) for the pairs of I.R. (A_1, β) and (E, β) and solving the system:

$$\begin{cases} 1\times 1 & +2[1\times\chi^{(\beta)}(C_3)] & +3[1\times\chi^{(\beta)}(\sigma_v)] & = 0 \\ 2\times 1 & +2[-1\times\chi^{(\beta)}(C_3)] & +0 & = 0 \end{cases} \tag{5.32}$$

The system is solved for $\chi^{(\beta)}(C_3) = 1$ and $\chi^{(\beta)}(\sigma_v) = -1$. The third I.R. β corresponds to the A_2 I.R. listed in Table 5.1. The sum rule (5.29) is verified:

$$A_2: \quad \sum_g \left|\,\chi^{(A_2)}(g)\,\right|^2 \;=\; 1 + 2 + 3 = 6\,.$$

Actually, we had already seen that the third one-dimensional I.R. β is given by the determinants (5.19) of the 3×3 matrices: the characters are +1 for proper rotations, −1 for mirror inversions.

5.3.6 Further consequences of the sum rule for characters

We have seen that the number of *inequivalent irreducible representations* of a group is equal to the number of classes of the group, and that the dimensionalities ℓ of the irreducible representations are connected to the order of the group through (5.25).
On the other hand, in principle there is no limit to the dimensionality of a *reducible representation*. It can thus happen that an irreducible representation is contained more than once within a reducible representation.
One can evaluate how many times a given irreducible representation is contained in a given reducible representation, once the characters of the two representations are known, by a suitable application of the sum rule (5.29).

Let us consider a reducible representation, which can be decomposed into several different irreducible representation and let $n^{(\beta)}$ be the number of identical irreducible representations β contained in the initial reducible representation. Since the matrices in the reducible representation can be blocked, the character of each transformation in the reducible representation can be expressed as the sum of the characters of the different irreducible representations:

$$\chi(g) = \sum_\beta n^{(\beta)} \chi^{(\beta)}(g) \,. \tag{5.33}$$

Our goal is to evaluate the numbers $n^{(\beta)}$. To this purpose, let us multiply both sides of (5.33) by the character of a generic irreducible representation α, sum over all symmetry transformations g, and make use of (5.29):

$$\begin{aligned} \sum_g \left\{ \chi(g) \left[\chi^{(\alpha)}(g) \right]^* \right\} &= \sum_g \left\{ \sum_\beta n^{(\beta)} \chi^{(\beta)}(g) \left[\chi^{(\alpha)}(g) \right]^* \right\} \\ &= \sum_\beta n^{(\beta)} \left\{ \sum_g \chi^{(\beta)}(g) \left[\chi^{(\alpha)}(g) \right]^* \right\} \\ &= \sum_\beta n^{(\beta)} n \, \delta_{\alpha\beta} \end{aligned} \tag{5.34}$$

where as usual n is the order of the group.
One can conclude that

$$\sum_g \left\{ \chi(g) \left[\chi^{(\alpha)}(g) \right]^* \right\} = n \, n^{(\alpha)} \,, \tag{5.35}$$

whence the reduction formula

$$n^{(\alpha)} = \frac{1}{n} \sum_g \left\{ \chi(g) \left[\chi^{(\alpha)}(g) \right]^* \right\} . \tag{5.36}$$

Problem: Consider the reducible 3×3 representation of the C_{3v} group, whose characters are listed in (5.15) and the three irreducible representations listed in Table 5.1. Making use of (5.36), verify that $n^{(A_1)} = 1$, $n^{(A_2)} = 0$, $n^{(E)} = 1$.

5.4 Irreducible representations: examples

In Section 5.2, we have considered the representations of some symmetry groups in the three-dimensional coordinate space. We have however found that the coordinate space is generally unsuited for finding *all* the irreducible representations of a given group.
A more general and effective approach is based on N-dimensional vector spaces, typically function spaces. This Section 5.4 contains a tutorial introduction to the representations in function spaces, focusing the attention on the pure rotation groups, beginning with the coordinate representation of the C_3 group, contunuing with the C_6 group to finish with the generic C_N group.

5.4.1 The C_3 group

The C_3 group contains three symmetry transformations: $C_3 = \{e, C_3^1, C_3^2\}$. It is a subgroup of C_{3v}. The multiplication table of C_3 has been given in Table 4.2 of Chapter 4.
The C_3 group is cyclic, the generator element is C_3^1. Being cyclic, it is commutative and each element forms a class by itself. There are thus three classes and three irreducible representations.

Let us first consider the representations in the coordinate basis (x, y, z) with z parallel to the C_3 axis. The three transformations of the C_3 group are represented by the matrices (5.9) and (5.11), already found when considering the representations of the C_{3v} group. The three matrices are blocked: as for the C_{3v} group, there is again the one-dimensional irreducible representation A_1 in the z basis. Contrary to the case of the C_{3v} group, however, the two-dimensional representation in (x, y) is further reducible for the C_3 group, which should have three one-dimensional irreducible

representations. By the way, by applying the sum rule (5.29) to the characters (5.17) of the two-dimensional representation, one finds

$$\sum\nolimits_g |\,\chi(g)\,|^2 \;=\; 4+2=6\neq n=3\,. \tag{5.37}$$

The problem thus arises of decomposing the two-dimensional reducible (x,y) representation in two one-dimensional irreducible representations. This decomposition cannot be achieved in the real Euclidean space.
It is convenient to consider the (xy) plane as the complex plane and introduce a new pair of independent complex coordinates, linear combinations of x and y:

$$x+iy=r\,\mathrm{e}^{i\phi}\,,\qquad x-iy=r\,\mathrm{e}^{-i\phi}\,. \tag{5.38}$$

Since r is constant under rotation (which is the only allowed type of symmetry transformations of the C_3 group) we can drop it and consider only the angular coordinate ϕ.
The two-dimensional coordinate basis (x,y) is substituted by a new two-dimensional basis represented by the two *orthonormal functions* $(e^{i\phi},e^{-i\phi})$. In the coordinate basis (x,y), the symmetry transformations of the group are represented by matrices that mix up the x and y coordinates. We want now to find the matrices representing the symmetry transformations in the new basis of functions $(e^{i\phi},e^{-i\phi})$.

The effect of the rotation C_3^1, say of the rotation of $2\pi/3$, on the basis of functions $(\mathrm{e}^{+i\phi},\,\mathrm{e}^{-i\phi})$ is obtained by the substitution $\phi\to\phi+2\pi/3$. In matrix notation,

$$\begin{pmatrix} \mathrm{e}^{i(\phi+2\pi/3)} \\ \mathrm{e}^{-i(\phi+2\pi/3)} \end{pmatrix} = \begin{pmatrix} e^{i2\pi/3} & 0 \\ 0 & e^{-i2\pi/3} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{i\phi} \\ \mathrm{e}^{-i\phi} \end{pmatrix} = \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon^2 \end{pmatrix} \begin{pmatrix} \mathrm{e}^{i\phi} \\ \mathrm{e}^{-i\phi} \end{pmatrix} \tag{5.39}$$

where $\epsilon=\mathrm{e}^{i2\pi/3}$ and $\epsilon^2=\mathrm{e}^{i4\pi/3}=\mathrm{e}^{-i2\pi/3}$.

The effect of the rotation C_3^2, say of the rotation of $4\pi/3$, on the basis of functions $(\mathrm{e}^{+i\phi},\,\mathrm{e}^{-i\phi})$ is obtained by the substitution $\phi\to\phi+4\pi/3$. In matrix notation,

$$\begin{pmatrix} \mathrm{e}^{i(\phi+4\pi/3)} \\ \mathrm{e}^{-i(\phi+4\pi/3)} \end{pmatrix} = \begin{pmatrix} \epsilon^2 & 0 \\ 0 & \epsilon \end{pmatrix} \begin{pmatrix} \mathrm{e}^{i\phi} \\ \mathrm{e}^{-i\phi} \end{pmatrix} \tag{5.40}$$

where again $\epsilon=\mathrm{e}^{i2\pi/3}=e^{-i4\pi/3}$ and $\epsilon^2=\mathrm{e}^{i4\pi/3}$.

Summarising, the three transformation matrices representing the C_3 group in the two-dimensional basis of functions $(\mathrm{e}^{+i\phi},\,\mathrm{e}^{-i\phi})$ are

$$\mathbf{D}(e)=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\qquad \mathbf{D}(C_3^1)=\begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon^2 \end{pmatrix},\qquad \mathbf{D}(C_3^2)=\begin{pmatrix} \epsilon^2 & 0 \\ 0 & \epsilon \end{pmatrix} \tag{5.41}$$

The three matrices of (5.41), including $\mathbf{D}(C_3^1)$ and $\mathbf{D}(C_3^2)$, are blocked, and give thus rise to two one-dimenional irreducible representations, as expected from the general theory.
Globally, the three-dimensional cyclic abelian C_3 group has three one-dimensional irreducible representations. The character table is given in Table 5.2.

Table 5.2: Character table of the irreducible representations (I.R.) of the symmetry group C_3. All I.R. are one-dimensional.

C_3	$e=C_3^3$	C_3^1	C_3^2
A_1	1	1	1
$E\,\{$	1	ϵ	ϵ^2
	1	ϵ^2	ϵ

Comments

a) The matrix elements $\epsilon = \mathrm{e}^{i2\pi/3}$ and $\epsilon^2 = [\mathrm{e}^{i2\pi/3}]^2 = \mathrm{e}^{-i2\pi/3}$ are complex numbers. They are also the characters of the rotations C_3^1 and C_3^2 in the one-dimensional representations based on the functions $e^{i\phi}$ and $e^{-i\phi}$ (Table 5.2).

b) One can easily verify that the sum rule (5.29) is satisfied by the three irreducible representations of Table 5.2.

c) The two basis functions $e^{i\phi}$ and $e^{-i\phi}$ defined in (5.38) are complex conjugate. In the character tables, the two corresponding one-dimensional irreducible representations are globally classified as one doubly degenerate two-dimensional E representation. This choice is connected to the fact that the (x, y) coordinates are mixed by the C_3^1 or C_3^2 rotation. We will see in Section 5.6.3 that the dimensionality of irreducible representations corresponds to the degeneracy of quantum energy levels; besides, complex conjugation of the eigenfunctions of the Schrödinger equation corresponds to time inversion symmetry; time inversion symmetry is not considered in the C_N group, so that the two functions $e^{i\phi}$ and $e^{-i\phi}$ give rise to distinct irreducible representations.

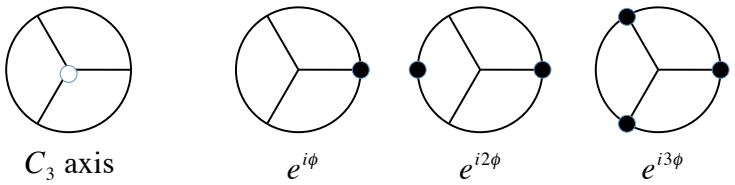

Figure 5.3: Schematic representation of the periodicity of the three functions $e^{im\phi}$, $m = 1, 2, 3$.

Basis functions

In the procedure depicted above, starting from the Euclidean space (x, y, z) and mixing the x and y coordinates to form the complex functions $(e^{i\phi}, e^{-i\phi})$ of the rotation coordinate ϕ, we have found three one-dimensional irreducible representations of the group C_3. This procedure appears somewhat unsatisfactory, since it mixes a coordinate basis z with two complex bases $e^{\pm i\phi}$.

Let us now introduce a more general approach, taking always as example the group C_3.

We know, from the general theory of irreducible representations, that there are three independent one-dimensional representations of the group C_3. We consider then a three-dimensional function space and choose three orthonormal functions as a basis forf the three irreducible representations. It is convenient to choose the three functions of the angular coordinate ϕ:

$$\mathrm{e}^{im\phi} \quad (m = 1, 2, 3)\,. \tag{5.42}$$

The three functions (5.42) are periodic, with angular period $2\pi/m$ (Fig. 5.3).

Within the basis represented by the three functions (5.42), the three symmetry operations $C_3^s (s = 1, 2, 3)$ are represented by the three 3×3 *reducible* matrices that operate on the column vector

$[e^{i\phi}/e^{i2\phi}/e^{i3\phi}]$:

$$\mathbf{D}(C_3^1) = \begin{pmatrix} (\mathrm{e}^{i2\pi/3})^1 & 0 & 0 \\ 0 & (\mathrm{e}^{i2\pi/3})^2 & 0 \\ 0 & 0 & (\mathrm{e}^{i2\pi/3})^3 \end{pmatrix},$$

$$\mathbf{D}(C_3^2) = \begin{pmatrix} (\mathrm{e}^{i4\pi/3})^1 & 0 & 0 \\ 0 & (\mathrm{e}^{i4\pi/3})^2 & 0 \\ 0 & 0 & (\mathrm{e}^{i4\pi/3})^3 \end{pmatrix} \tag{5.43}$$

$$\mathbf{D}(C_3^3) = \begin{pmatrix} \mathrm{e}^0 & 0 & 0 \\ 0 & \mathrm{e}^0 & 0 \\ 0 & 0 & \mathrm{e}^0 \end{pmatrix},$$

The row index corresponds to the parameter m (first row $m = 1$, second row $m = 2$, third row $m = 3$).
The three matrices (5.43) are diagonal; the three diagonal elements of each matrix correspond to the three 1×1 irreducible representations of the same symmetry transformation.

Table 5.3: The three basis functions $e^{i\phi}$ and the corresponding one-dimensional I.R. of each symmetry transformation.

Basis	C_3^1	C_3^2	$C_3^3 = e$	I.R.
$e^{i\phi}$	$e^{i2\pi/3}$	$e^{i4\pi/3}$	$e^{i6\pi/3} = 1$	} E
$e^{i2\phi}$	$e^{i4\pi/3}$	$e^{i8\pi/3}$	$e^{i12\pi/3} = 1$	
$e^{i3\phi}$	$e^{i6\pi/3} = 1$	$e^{i12\pi/3} = 1$	$e^{i18\pi/3} = 1$	A_1

Each one of the three functions (5.42) is a basis of a one-dimensional irreducible representation. For a given representation, say for a given value of m,

$$\mathbf{D}(C_3^s) = (\mathrm{e}^{i2\pi s/3})^m . \tag{5.44}$$

The situation can be clarified by looking at Table 5.3. By setting $\epsilon = e^{i2\pi/3}$, the character Table 5.2 is recovered.

Comments

a) The function $e^{i3\phi}$ ($m = 3$) is the basis of the identical I.R. A_1: the function has periodicity $2\pi/3$ and is not modified by the symmetry transformations of the group C_3.
The effect of the symmetry transformations on the other basis functions is a multiplication by a phase factor.

b) The representation based on the function $e^{i\phi}$ is the group of the complex roots of order 3 of the number one.

c) The one-dimensional representation of the three symmetry transformations based on the function $e^{im\phi}$ can be obtained even if the basis function is changed to $e^{i(m\pm3)\phi}$. There are thus infinite equivalent irreducible representations based on different functions and infinite possible sets of three orthonormal basis functions.
One frequently adopted choice is $m = -1, 0, +1$.

.

5.4.2 The C_6 group

The advantage of representations based on bases of functions rather than on the three-dimensional coordinate basis is further highlighted by the example of the C_6 group, already considered in

Section 5.2. In the coordinate basis we found the 3×3 representation (5.20), which is composed by a one-dimensional z representation and a two-dimensional xy representation. However, since the group C_6 is cyclic and thus commutative, we know from Section 5.3 that there are six independent one-dimensional irreducible representations.
To single out the six one-dimensional irreducible representations, it is convenient to operate in a six-dimensional function space spanned by the six orthonormal functions of the angular coordinate ϕ

$$\mathrm{e}^{im\phi} \quad (m = 1, 2, \ldots, 6)\,. \tag{5.45}$$

The six functions (5.45) are periodic, with angular period $2\pi/m$.
In the basis (5.45), the six group elements (independent C_N^s rotations) are represented by six 6×6 matrices, analogous to the 3×3 matrices (5.43) of the C_3 group. For the generic rotation C_6^s, the matrix is

$$\mathbf{D}(C_6^s) = \begin{pmatrix} (\mathrm{e}^{is2\pi/6})^1 & 0 & 0 & 0 & 0 & 0 \\ 0 & (\mathrm{e}^{is2\pi/6})^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & (\mathrm{e}^{is2\pi/6})^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & (\mathrm{e}^{is2\pi/6})^4 & 0 & 0 \\ 0 & 0 & 0 & 0 & (\mathrm{e}^{is2\pi/6})^5 & 0 \\ 0 & 0 & 0 & 0 & 0 & (\mathrm{e}^{is2\pi/6})^6 \end{pmatrix}. \tag{5.46}$$

and operates on a six-component column vector $\{e^{i\phi}, e^{i2\phi}, e^{i3\phi}, e^{i4\phi}, e^{i5\phi}, e^{i6\phi}\}$.
All the $\mathbf{D}(C_6^s)$ matrices are diagonal, so that the 6×6 representation is reducible into 6 one-dimensional I.R. Each one of the functions (5.45) is the basis of a one-dimensional I.R.
The characters of the six I.R. are different. The character table is shown in Tab. 5.4, where $\omega = \exp(i2\pi/6)$.

In particular, the function $e^{i6\phi}$ $(m = 6)$ is the basis of the identical I.R. A_1: the function has periodicity $2\pi/6$ and is not modified by the symmetry transformations of the group C_6. The effect of the symmetry transformations on the other basis functions is a multiplication by a phase factor.
The one-dimensional representation of the six symmetry transformations based on the function $e^{im\phi}$ can be obtained even if the basis function is changed to $e^{i(m+6)\phi}$.

Table 5.4: Character table of the irreducible representations (I.R.) of the symmetry group C_6, where $\omega = \exp(i2\pi/6)$. All I.R. are one-dimensional. The two I.R. labeled E' are equivalent, and are considered as doubly degenerate two-dimensional representations; the same for E''.

C_6	$e = C_6^6$	C_6^1	C_6^2	C_6^3	C_6^4	C_6^5
A_1	1	1	1	1	1	1
E'	1	ω	ω^2	ω^3	ω^4	ω^5
E''	1	ω^2	ω^4	1	ω^2	ω^4
B	1	-1	1	-1	1	-1
E''	1	ω^4	ω^2	1	ω^4	ω^2
E'	1	ω^5	ω^4	ω^3	ω^2	ω

Problem 1: Verify the character table 5.4 of the group C_6. Verify the sum rule (5.29).

Problem 2: It is instructive to compare the group C_6 with the group C_{6v}, in analogy with the comparison of C_3 with C_{3v}. The classes of the C_{6v} group are six, two containing one element, two containing two elements, two containing three elements:

$$\{e\}\,,\ \{C_6^3\}\,,\ \{C_6^2, C_6^4\}\,,\ \{C_6^1, C_6^5\}\,,\ \{3\sigma_v\}\,,\ \{3\sigma_d\}\,. \tag{5.47}$$

The conjugate rotations C_6^m and C_6^{6-m} are grouped in the same class.

5.4.3 The C_N group and its generalisations

The approach depicted in the two examples of the C_3 and C_6 groups can be generalised to any C_N group:

a) The N possible rotations are C_N^s, where $s = 1, \dots N$.
C_N^1 is the generator of the cyclic group and $C_N^N = C_N^0 = e$ is the identity.
b) We operate in an N-dimensional function space spanned by N orthonormal basis functions

$$\mathrm{e}^{im\phi} \quad (m = 1, \dots N)\,. \tag{5.48}$$

c) The N matrices representing the N symmetry operations are diagonal; for a given rotation C_N^s, the matrix is diagonal and the matrix elements (along the diagonal) are

$$\left(\mathrm{e}^{i2\pi s/N}\right)^1\,,\ \left(\mathrm{e}^{i2\pi s/N}\right)^2\,,\cdots\,,\ \left(\mathrm{e}^{i2\pi s/N}\right)^N = \left(\mathrm{e}^{i2\pi s/N}\right)^0\,. \tag{5.49}$$

d) For a given basis function $e^{im\phi}$ with a given value of m, the C_N^s transformation is represented by the complex number

$$\mathbf{D}(C_N^s) = (\mathrm{e}^{i2\pi s/N})^m\,. \tag{5.50}$$

e) The function $e^{iN\phi}$ $(m = N)$ is the basis of the identical I.R. A_1: the function has periodicity $2\pi/N$ and is not modified by the symmetry transformations of the group C_N.
The effect of the symmetry transformations on the other basis functions is a multiplication by a phase factor.
f) A basis set of dimension smaller than N would not give all the irreducible representations. A basis set of dimension larger than N would be redundant.

Alternative basis sets.

The basis set (5.48) is not the only one possible.
One can verify that the complex number representing the rotation C_N^s with respect to the basis function $e^{i(m)\phi}$, where $m = 1, \dots N$, is equal to the complex number representing the same rotation with respect to the function $e^{i(m\pm N)\phi}$:

$$\mathbf{D}(C_N^s) = \left(e^{i2\pi s/N}\right)^{(m\pm N)} = \left(e^{i2\pi s/N}\right)^m\,. \tag{5.51}$$

There are thus infinite possible choices of the N-dimensional basis set.
A frequently adopted basis set is obtained by considering the N values of m centred on $m = 0$. More precisely, the new basis set is

$$\mathrm{e}^{im\phi} \quad \text{where} \quad \begin{cases} -N/2 < m \le N/2 & \text{for } N \text{ even} \\ -N/2 < m < N/2 & \text{for } N \text{ odd} \end{cases} \tag{5.52}$$

Example 1: For $N = 3$ the basis functions are $e^{-i\phi}, 1, e^{i\phi}$. Such a basis set contains a pair of complex conjugate basis functions $(e^{i\phi}, e^{-i\phi})$.

Example 2: For $N = 6$ the basis functions are $e^{-2\phi}, e^{-\phi}, 1, e^{\phi}, e^{2\phi}, e^{3\phi}$. Such a basis set contains two pairs of complex conjugate basis functions $(e^{i\phi}, e^{-i\phi})$ and $(e^{2i\phi}, e^{-2i\phi})$. Can you recognise the corresponding I.R. in the Table of characters 5.4?

General basis - Floquet theorem

The basis functions $e^{im\phi}$ of (5.48) or (5.52) don't represent the most general choice of basis set for the group C_N.
One can actually see that a basis of an I.R. of the group C_N is represented also by any function

$$e^{im\phi}\ e^{inN\phi}\,, \tag{5.53}$$

where n is an integer number.
Actually, a rotation of $2\pi s/N$ transforms the function (5.53) according to

$$\left(e^{im(\phi+2\pi s/N)}\, e^{inN(\phi+2\pi s/N)}\right) \;=\; \left[\left(e^{2\pi s/N}\right)^m \underbrace{e^{ins2\pi}}_{=1}\right] \left(e^{im\phi}\, e^{inN\phi}\right) . \tag{5.54}$$

The representation of the $2\pi s/N$ rotation is thus given again by (5.50)

$$\mathbf{D}(C_N^s) = (\mathrm{e}^{i2\pi s/N})^m \tag{5.55}$$

for the new basis (5.53) too.
Notice that in (5.53) the first factor ($m \leq N$) is a low-frequency function, the second factor ($nN \geq N$) is a high frequency function.
The result can be further generalised. Any function

$$e^{im\phi} \sum\nolimits_n A_n e^{inN\phi} \tag{5.56}$$

is a basis of a one-dimensional I.R. of the C_N group. The sum in (5.56) represents a complex periodic function of period $2\pi/N$.

Application to the one-dimensional linear chain

An interesting application of the results found for the C_N group concerns the translational symmetry of crystals (Section 4.6).
Let us focus the attention on the one-dimensional case and consider a linear chain of N points with spatial period a. The translations are $T = sa$, where $s = 1, 2, ..., N$.
To avoid the difficulties connected with the finite size of real crystals and exploit the translational symmetry properties of infinite crystals, the periodic boundary conditions were introduced in § 2.5.1, which amount to impose that the value of any property is the same for $s = 0$ and for $s = N$.

For a one dimensional system, the periodic boundary condition formally corresponds to joining the two extrema of the linear chain, forming a circumference of length Na and radius $r = Na/2\pi$. We can now take advantage of the previous treatment of the C_N group by considering that any length on the circumference can be expressed as $x = r\phi$.
The N basis functions (5.48) of the rotation group can thus be expressed in terms of translational coordinates:

$$e^{im\phi} \;=\; e^{imx/r} \;=\; \exp\left(i\, m \frac{2\pi}{aN}\, x\right) \;=\; e^{ik_m x}\,, \tag{5.57}$$

where $k_m = (m/N)(2\pi/a)$ and $m = 1, 2, ..., N$.
In (5.57), the adimensional angular coordinate ϕ is substituted by the dimensional length coordinate x, and the adimensional parameter m is substituted by the wavenumber k_m, with the dimension of inverse length.
The N orthonormal functions

$$\exp(ik_m x)\,, \qquad k_m = (2\pi m/Na)\ [1, \ldots m, \ldots N] \tag{5.58}$$

are periodic in the x coordinate, with periodicity Na/m. The shortest period is $\Delta x = a$ for $m = N$, the longest period is $\Delta x = Na$ for $m = 1$.

The N functions (5.58) are the basis of an $N \times N$ reducible representation, where the transformation induced by the generic translation $T = sa$ is represented by

$$\begin{pmatrix} \exp[ik_1(x+T)] \\ \cdots \\ \exp[ik_N(x+T)] \end{pmatrix} = \begin{pmatrix} \exp[ik_1 T] & 0 & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & \exp[ik_N T] \end{pmatrix} \begin{pmatrix} \exp[ik_1 x] \\ \cdots \\ \exp[ik_N x] \end{pmatrix} \tag{5.59}$$

All matrices in the $N \times N$ representation are diagonal; the representation can be decomposed in N one-dimensional I.R., one for each of the basis functions (5.58).

For a given basis function, say for a given value of k_m in (5.57), the translation $T = sa$ is represented by the complex number

$$e^{ik_m T} \; : \tag{5.60}$$

say the basis function is multiplied by the phase factor $e^{ik_m T}$.
The function $\exp(ik_N x)$ is the basis of the identical I.R. A_1; one can easily verify that for any translation $T = sa$

$$\exp(ik_N T) \;=\; \exp\left(i\frac{2\pi}{a}sa\right) \;=\; \exp(i2\pi s) \;=\; 1\,. \tag{5.61}$$

The function $\exp(ik_N x)$ has periodicity $\Delta x = a$ and is not modified by any translation $T = sa$ (Fig. 5.4).

The quantities k_m are connected to the primitive vector $b = 2\pi/a$ of the reciprocal space by

$$k_m \;=\; \frac{m}{N}\,b\,. \tag{5.62}$$

Otherwise stated, the values k_m are confined within a primitive cell of the reciprocal space.

As it was observed for the basis set (5.48) of the C_N group, also for the linear chain with periodic boundary conditions the basis set (5.58) is not the only one possible. An alternative basis set is

$$\exp(ik_m x)\,, \qquad k_m = \frac{2\pi}{a}\,\frac{m}{N} \quad (N/2 < m \le N/2)\,, \tag{5.63}$$

for which the values k_m are confined within the first Brillouin Zone.

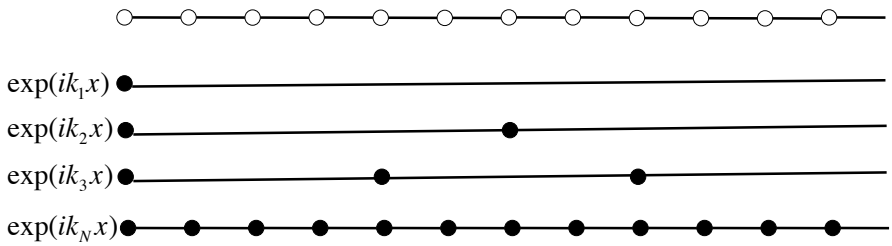

Figure 5.4: Schematic representation of the periodicity of the functions $e^{iK_m x}$, with respect to the periodicity of the lattice (open circles).

Bloch functions

For the linear chain of period a with periodic boundary conditions, the most general basis functions can be obtained by adapting the rotational equation (5.56) to the linear case; one obtains

$$\psi(x) \;=\; e^{ik_m x} \sum\nolimits_n A_n e^{inbx}\,, \tag{5.64}$$

where $b = 2\pi/a$ is the primitive vector of the reciprocal one-dimensional lattice.
According to (3.1) of Section 3.1, the sum in (5.64) is the Fourier expansion of a periodic function $u(x)$ of period a.
Functions like (5.64) are called Bloch functions. We will see in Section 9.2 that the eigenfunctions of the Schrödinger equation for an electron in a periodic crystalline potential are Bloch functions.

5.5 Group representations in the function space

The examples of the previous Section 5.4 (groups C_3, C_6, C_N) have shown that, in order to find all the irreducible representations of a symmetry group, it is convenient to work in an ℓ-dimensional subspace of the linear space of the functions of the coordinates.
In this Section 5.5 we introduce some general considerations on the group representation in function spaces, which are illustrated by a simple example relating to the C_{3v} group.
This Section is preliminary to the next Section 5.6, mainly dedicated to some applications of group theory to Quantum Mechanics.

5.5.1 General considerations

Let $\Phi_1(q), \Phi_2(q), ...\Phi_\ell(q)$ be an orthonormal basis set of ℓ linearly independent functions of coordinates (typically generalised coordinates in a suitable configurations space).

Transformations of components

Every function belonging to the vectorial space spanned by the Φ_i functions can be expressed as a linear combination of the ℓ basis functions:

$$\Psi \;=\; \sum_{i=1}^{\ell} \alpha_i \, \Phi_i \, . \tag{5.65}$$

Let us now consider a group of symmetry transformations $\mathcal{G} = \{g_1, \ldots, g_n\}$. Each one of the n transformations g of the symmetry group operates on the functions Ψ according to

$$\Psi' \;=\; P_g \, \Psi \;\; , \tag{5.66}$$

where P_g is the operator corresponding to g and

$$\Psi' \;=\; \sum_{i=1}^{\ell} \alpha'_i \, \Phi_i \tag{5.67}$$

is again expressed as a linear combination of the ℓ basis functions, with new coefficients α'.
Each one of the n operators P_g is represented, within the space spanned by the ℓ basis functions Φ_i, by an $\ell \times \ell$ matrix $\mathbf{D}(P_g)$.
Each matrix $\mathbf{D}(P_g)$ transforms the column vector of the components α_i of Ψ (5.65) into the column vector of the components α'_i of Ψ' (5.67):

$$\alpha'_i \;=\; \sum_{k=1}^{\ell} D_{ik}(P_g) \, \alpha_k \, . \tag{5.68}$$

This procedure formally corresponds to the transformations of (x, y, z) coordinates considered in Section 5.2.

Transformations of basis functions

In Section 5.4 we considered the effect of symmetry transformations directly on the basis functions rather than on the components.
To give a formal support to that approach, let us study what happens when the generic function Ψ of (5.65) corresponds to one of the basis functions Φ_ν, say $\alpha_i = \delta_{i\nu}$. The operator P_g acting on a basis function Φ_ν gives rise to a new function Φ', which can obviously be expressed as a linear combination of all the basis functions

$$P_g \, \Phi_\nu \;=\; \Phi' \;=\; \sum_{i=1}^{\ell} \beta'_i \, \Phi_i \tag{5.69}$$

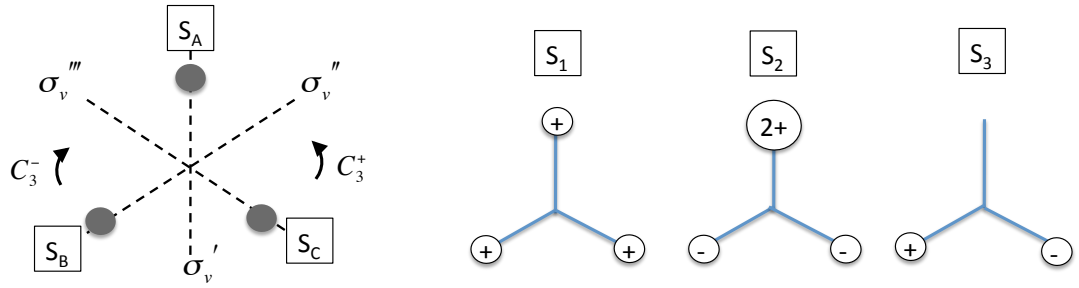

Figure 5.5: Left: the three $1s$ orbitals S_a, S_b, S_c of the H atoms in the basal plane of the NH_3 molecule. Right: The three linear combinations $S_1, S_2 S_3$ (5.77) of the three atomic orbitals.

The coefficients β'_i of the transformed function can be expressed in terms of the coefficients α_k of the original function according to (5.68) and taking into account that the original function is the basis Φ_ν, so that $\alpha_k = \delta_{k\nu}$:

$$\Phi' = P_g\,\Phi_\nu \;=\; \sum_{i=1}^{\ell}\left[\sum_{k=1}^{\ell} D_{ik}(P_g)\,\alpha_k\right]\Phi_i \;=\; \sum_{i=1}^{\ell} D_{i\nu}(P_g)\,\Phi_i\,. \tag{5.70}$$

If we now let the index ν vary from 1 to ℓ, we can see that the matrix $\mathbf{D}(P_g)$ transforms the column vector of the basis functions Φ_ν ($\nu = 1 \ldots \ell$) into the column vector of the transformed basis functions $\Phi'_\nu = P_g\Phi_\nu$ ($\nu = 1 \ldots \ell$). Otherwise stated, we can evaluate the matrices $\mathbf{D}(P_g)$ by considering the effect of P_g on the basis functions (as was done in Section 5.4).

Note: One can observe that in (5.68) the sum is over the second index of the matrix, while in (5.70) the sum is over the first index. This fact formally corresponds to the two different but equivalent description of a symmetry operations: a rotation of an object with respect to a fixed reference or a rotation of the reference with respect to the fixed object.

In conclusion, we consider groups which are isomorphic to the coordinate transformation groups, and whose elements are transformation operators which operate on functions rather than on coordinates.

Note: Only for commutative groups, such as the cyclic rotation groups considered in Section 5.4, are the I.R. necessarily one-dimensional.

5.5.2 An orbital-based representation of the C_{3v} group

To give a simple example of the representation of a non-commutative group in the function space, let us consider again the C_{3v} group and focus our attention on the NH_3 molecule for concreteness. A convenient set of basis functions is given by the $1s$ atomic orbitals of the nitrogen and hydrogen atoms. Let S_n, S_a, S_b, S_c be the $1s$ orbitals of the N atom and of the three H atoms, respectively (Fig. 5.5, left).
The basis is 4-dimensional, and the symmetry transformations are represented by 4×4 matrices:

$$\begin{pmatrix} S'_n \\ S'_a \\ S'_b \\ S'_c \end{pmatrix} = \begin{pmatrix} D_{nn} & D_{na} & D_{nb} & D_{nc} \\ D_{an} & D_{aa} & D_{ab} & D_{ac} \\ D_{bn} & D_{ba} & D_{bb} & D_{bc} \\ D_{cn} & D_{ca} & D_{cb} & D_{cc} \end{pmatrix} \begin{pmatrix} S_n \\ S_a \\ S_b \\ S_c \end{pmatrix} \tag{5.71}$$

The matrix representing the identity e is diagonal

$$\mathbf{D}(e) \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.72}$$

Both the $C_3^1 = C_3^+$ and $C_3^2 = C_3^-$ rotations leave unchanged the N orbital and permute the H orbitals, $(S'_n S'_a S'_b S'_c) = (S_n S_b S_c S_a)$ and $(S'_n S'_a S'_b S'_c) = (S_n S_c S_b S_a)$, respectively. The transformation matrices are

$$\mathbf{D}(C_3^1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \mathbf{D}(C_3^2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{5.73}$$

The transformation matrices for the mirror reflections are

$$\mathbf{D}(\sigma'_v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathbf{D}(\sigma''_v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \mathbf{D}(\sigma'''_v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.74}$$

The characters of this 4×4 representation are

$$\chi(e) = 4\,, \quad \chi(C_3) = 1\,, \quad \chi(\sigma_v) = 2\,. \tag{5.75}$$

The character of each transformation corresponds to the number of basis functions (orbitals) which are left unchanged by the transformation.

The 4×4 representation is clearly reducible: the sum rule (5.1) gives $\sum_g |\chi(g)|^2 = 30$.

First decomposition of the 4×4 representation

To find the irreducible representations of the C_{3v} group in the given function basis, one can start by observing that all matrices of the 4×4 representation are blocked, so that one can easily separate two independent representations:

1. A one-dimensional I.R. of type A_1 in the S_n basis. Actually, the $1s$ orbital of the N atom is invariant with respect to any transformation of the group.
2. A three-dimensional representation in the $(S_a S_b S_c)$ basis, where the symmetry transformations mix up the basis functions. The characters in this representations are

$$\chi(e) = 3\,, \quad \chi(C_3) = 0\,, \quad \chi(\sigma_v) = 1\,. \tag{5.76}$$

This three-dimensional representation is reducible, since the sum rule (5.1) gives $\sum_g |\chi(g)|^2 = 12$.

Alternative 3-dimensional representations

To further decompose the three-dimensional representation in the $(S_a S_b S_c)$ basis, it is necessary to find a different basis, where the transformation matrices are conveniently blocked. To this effect, we search for suitable linear combinations of the $S_a S_b S_c$ basis functions. One such linear combination is the following one (Fig. 5.5, right):

$$\begin{aligned} S_1 &= S_a + S_b + S_c \\ S_2 &= 2S_a - S_b - S_c \\ S_3 &= S_b - S_c \end{aligned} \tag{5.77}$$

Let us build up the transformation matrices in this new representation.
The matrix representing identity is diagonal

$$\mathbf{D}(e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.78}$$

Let us now consider the two rotations C_3^1 and C_3^2. The S_1 function is invariant with respect to the rotations, while the S_2 and S_3 functions are mixed up. The new S_2' and S_3' functions obtained by the group rotations are linear combinations of the old S_2, S_3 functions. One can easily verify that the transformation matrices are

$$\mathbf{D}(C_3^1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & +3/2 \\ 0 & -1/2 & -1/2 \end{pmatrix} \; ; \quad \mathbf{D}(C_3^2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -3/2 \\ 0 & +1/2 & -1/2 \end{pmatrix} \tag{5.79}$$

Let us finally consider the mirror reflections. The S_1 function is invariant with respect to all reflections. The S_2 and S_3 functions are instead modified, and in two of the three transformations they are mixed up. One can easily verify that the transformation matrices are

$$\mathbf{D}(\sigma_v') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} ; \mathbf{D}(\sigma_v'') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -3/2 \\ 0 & -1/2 & +1/2 \end{pmatrix} ; \mathbf{D}(\sigma_v''') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & +3/2 \\ 0 & +1/2 & +1/2 \end{pmatrix} \tag{5.80}$$

The characters in this new 3×3 representations are equal to the characters (5.76) (the linear combination of basis functions doesn't modify the characters).

Irreducible representations

The advantage of the new 3-dimensional representation based on the S_1, S_2, S_3 functions is that the corresponding matrices are "blocked" One can then separate two independent representations:

1. The identical one-dimensional I.R. A_1 in the S_1 basis. Actually, the $S_1 = S_a + S_b + S_c$ combination of H orbitals is invariant with respect to any transformation of the group. The same invariance is exhibited by the $1s, 2s, 2p_z$ orbitals of nitrogen.
2. A two-dimensional I.R. E in the $S_2 S_3$ basis. The characters are

   $$\chi(e) = 2\,, \quad \chi(C_3) = -1\,, \quad \chi(\sigma_v) = 0\,. \tag{5.81}$$

   The group transformations mix up the (S_2, S_3) functions, as well as the $(2p_x, 2p_y)$ orbitals of nitrogen.

In conclusion, the 4×4 initial representation of the C_{3v} group has been reduced into three I.R.: two equivalent identical representations A_1, and a two-dimensional representation E.
This is thus a good example for checking the sum rule (5.36). Considering the characters (5.75) of the 4×4 representation, one can calculate by (5.36) the numbers $n^{(\alpha)}$ for the three irreducible representations A_1, A_2, E of the group given in Table 5.1. One easily finds:

for $\alpha = A_1$, $n^{(\alpha)} = 2$
for $\alpha = A_2$, $n^{(\alpha)} = 0$
for $\alpha = E$, $n^{(\alpha)} = 1$

5.6 Group representations and Quantum Mechanics

In the following, we focus our attention on some applications of the theory of group representations to quantum mechanical problems.

5.6.1 Hamiltonian operator and symmetry operators

A basic problem of the quantum theory for atoms, molecules and solids is the solution of the stationary Schrödinger equation, $H\Psi = E\Psi$, in order to find the energy levels of the system and the corresponding eigenstates.
Only for very simple systems can the Schrödinger equation be exactly solved (harmonic oscillator, single particle in a central field). In most cases, the Hamiltonian function is the sum of many contributions to the kinetic and potential energy coming from the many particles involved (see below, Section 6.1). The solution is necessarily approximate. Symmetry considerations can simplify the problem in exact way, before the approximations are introduced.

Operators of the symmetry transformations

As a first step, since we are considering group representations with respect to basis functions, it is convenient to describe the symmetry transformations of systems by operators acting in the functions space.
Let us consider a group $\mathcal{G} = \{g_1, g_2, \ldots, g_n\}$ of symmetry transformations. To each transformation g one can associate an operator P_g; one can then introduce the eigenvalue equation

$$P_g \, \Psi(\vec{r}) \;=\; p_g \, \Psi(\vec{r}) \,. \tag{5.82}$$

The results of previous Sections 5.4 and 5.5 can be recast in the language of operators (with a slight modification of symbols). Let us give some simple examples.

Example 1: **Rotation operator**

Let R_N^s be the operator corresponding to a rotation C_N^s of $2\pi s/N$ around a given axis (where as usual s is an integer number, $s = 1 \cdots N$) and let

$$R_N^s \, \Psi(\phi) \;=\; r_N^s \, \Psi(\phi) \tag{5.83}$$

be the corresponding eigenvalue equation.
Since $R_N^s \, \Psi(\phi) = \Psi(\phi - 2\pi s/N)$, so that $\Psi(\phi - 2\pi s/N) = r_N^s \Psi(\phi)$, and since the two functions $\Psi(\phi)$ and $\Psi(\phi - 2\pi s/N)$ are normalised to one, it is easily seen that $|r_N^s|^2 = 1$. The eigenvalues of the rotation operator R_N^s correspond to the matrix elements (5.49)

$$(r_N^s)^m \;=\; \left(e^{i2\pi s/N}\right)^m \,, \quad (m = 1 \cdots N) \tag{5.84}$$

Example 2: **Translation operator**

Let $T_{\vec{n}}$ be the operator corresponding to a translation $\vec{n} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$ of a crystal lattice and let

$$T_{\vec{n}} \, \Psi(\vec{r}) \;=\; t_{\vec{n}} \, \Psi(\vec{r}) \tag{5.85}$$

be the corresponding eigenvalue equation.
Since $T_{\vec{n}} \, \Psi(\vec{r}) = \Psi(\vec{r} - \vec{n})$, so that $\Psi(\vec{r} - \vec{n}) = t_{\vec{n}} \Psi(\vec{r})$, and since the two functions $\Psi(\vec{r})$ and $\Psi(\vec{r} - \vec{n})$ are normalized to one, it is easily seen that $|t_{\vec{n}}|^2 = 1$.
The eigenvalues of the translation operator $T_{\vec{n}}$ are of the form (see 5.59)

$$t_{\vec{n}}^{\vec{k}} \;=\; e^{i\vec{k}\cdot\vec{n}} \,. \tag{5.86}$$

Example 3: **Inversion operator**

Let J be the inversion operator (corresponding to the inversion transformation I) and let

$$J \, \Psi(x) \;=\; j \, \Psi(x) \tag{5.87}$$

be the corresponding eigenvalue equation.
Since $J\Psi(x) = \Psi(-x) = j\Psi(x)$ and $J^2\Psi(x) = \Psi(x)$, the eigenvalues of J are $j = \pm 1$, corresponding to even and odd eigenfunctions.

Invariance of the Hamiltonian

In Quantum Mechanics, one is interested in the symmetry transformations that leave invariant the Hamiltonian operator of a given system. As a simple example, let us consider the familiar form of the Hamiltonian in the single-particle approximation

$$H = -\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r}) \,. \tag{5.88}$$

One can check that the kinetic energy operator

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \tag{5.89}$$

is invariant with respect to the symmetry operations of the rotation-inversion group $O(3)$ and with respect to translations.
The symmetry of H is thus governed by the symmetry of the potential energy $V(\vec{r})$.

Example 1: If the potential energy V of a system only depends on the scalar distance r from a central point, then every rotation or inversion of the system leaves the potential unchanged. It can be instructive to apply the known point symmetry transformations to the Coulomb potential

$$V(r) = \frac{1}{4\pi\epsilon_0}\frac{q}{r} = \frac{1}{4\pi\epsilon_0}\frac{q}{\sqrt{x^2+y^2+z^2}} \tag{5.90}$$

Example 2: The periodic potential of an electron in a crystal $V(\vec{r}) = V(\vec{r}+\vec{T})$ is invariant with respect to crystal translations $\vec{T}$.

Commutation of the Hamiltonian with the symmetry operators

If the Hamiltonian operator is invariant with respect to the symmetry transformations of a group $\mathcal{G}$, then for every $g \in \mathcal{G}$ one has

$$P_g^{-1} H P_g = H \qquad \Rightarrow \qquad P_g H = H P_g\,, \tag{5.91}$$

say the two operators P_g and H commute: $[P_g, H] = 0$.
According to a general theorem on commuting operators, if Ψ is an eigenfunction of H with eigenvalue E, then $P_g\Psi$ is eigenfunction of H too, with the same eigenvalue E.
Let us consider the stationary Schrödinger equation

$$H\,\Psi_m \;=\; E_m\,\Psi_m \tag{5.92}$$

and apply the symmetry operator P_g to both the Hamiltonian operator H and to the eigenfunction Ψ_m:

$$P_g\,H\,P_g^{-1}\,(P_g\,\Psi_m) \;=\; E_m\,(P_g\,\Psi_m)\,; \tag{5.93}$$

taking into account (5.91), one finds

$$H\,(P_g\,\Psi_m) \;=\; E_m\,(P_g\,\Psi_m)\,. \tag{5.94}$$

In conclusion *the energy eigenvalues E_m are invariant with respect to the symmetry transformations.*

5.6.2 Group representations and degeneracy of energy eigenvalues

Eigenfunctions of the Schrödinger equation

It is of interest to study how the symmetry operators P_g $(g = g_1, \ldots, g_n)$ can modify the eigenfunctions Ψ_m belonging to a given eigenvalue E_m. We distinguish the two cases of non-degenerate levels and of degenerate levels.

1. If the eigenvalue E_m is *non degenerate*, there is only one eigenfunction Ψ_m, defined to within a phase factor.
   A symmetry operator P_g can only modify the phase factor of the eigenfunction; there is no possibility of mixing different eigenfunctions by any symmetry transformation.
   The eigenfunction Ψ_m of a non-degenerate level is a basis of a *one-dimensional irreducible representation* of the symmetry group.

2. It the eigenvalue E_m is *degenerate*, its eigenfunctions belong to a subspace of dimensionality ℓ_m of the Hilbert space of the system, where one can choose an ℓ_m-dimensional basis of orthonormal functions $\Phi_{m,i}$ $(i = 1 \ldots \ell_m)$.
   A symmetry operator P_g transforms an eigenfunction into another eigenfunction belonging to

the degeneracy subspace: the different possible eigenfunctions are mixed up by the symmetry transformations.
Every set of ℓ orthonormal eigenfunctions forming a basis for the eigenfunctions sub-space represents also a basis of an irreducible representation of the symmetry group. The multiplicity of the degenerate level corresponds to the dimensionality of the irreducible representation. Note that the number n of symmetry operations g (order of the group) can be (and generally is) larger than the degeneracy ℓ_m of the eigenvalue.

There is thus a connection between the degeneracy of an energy level and the dimensionality of the irreducible representations of the symmetry group of a quantum systems. To each energy level of a quantum system it corresponds an I.R. of its symmetry group.
A system has degenerate energy levels if its symmetry group possesses irreducible representations of dimensionality higher than one (with some exception already encountered when treating cyclic groups, which will be discussed below). The degree of degeneracy of a degenerate level must be equal to the dimensionality of one of the irreducible representations of the group.
For non linear molecules, the dimensionality of irreducible representations generally doesn't exceed three; as a consequence, the maximum degree of expected degeneracy of energy levels is generally three.

5.6.3 Hamiltonian group and symmetry breaking.

Up to now, we have considered only spatial symmetry transformations (proper and improper rotations, translations).
The Hamiltonian function can be invariant also with respect to symmetry transformations of different kind, such as time reversal, charge conjugation, permutation of particles and so on. The group of all transformations with respect to which the Hamiltonian is invariant is named the *symmetry group of the Hamiltonian.*

We present here two particularly interesting examples where the symmetries of the Hamiltonian are not amenable solely to the point and translation symmetries up to now considered: the time reversal symmetry and the "umbrella inversion" of the ammonia molecule. We consider then the distinction between essential and accidental degeneracy.

Time reversal symmetry

A peculiar case was encountered when considering the cyclic Abelian point groups C_3 in Section 5.4. In the set of basis functions, the pairs of complex conjugate functions $e^{im\phi}, e^{-im\phi}$ were considered as bases of two different one-dimensional irreducible representations.

The situation is more complicate if the two functions $e^{im\phi}, e^{-im\phi}$ are obtained as eigen-functions of a Schrödinger equation; since they are complex conjugate, they correspond to the same doubly degenerate eigenvalue E_m.
We have already observed that the complex conjugation of the eigenfunctions of the Schrödinger equation is connected with the time reversal symmetry. Time inversion implies that $\vec{r}(-t) = \vec{r}(t)$ and $\vec{p}(-t) = -\vec{p}(t)$. Let us consider a time-independent Hamiltonian; the kinetic term depends on p^2, say it is invariant with respect to the inversion $\vec{p} \to -\vec{p}$, the potential term only depends on $\vec{r}$ and is again independent of time reversal.
The time-dependent Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) \;=\; H\, \Psi(\vec{r}, t)\,. \tag{5.95}$$

If we change t into $-t$ and take the complex conjugate of both sides, we obtain

$$i\hbar \frac{\partial}{\partial t} \Psi^*(\vec{r}, -t) \;=\; H\, \Psi^*(\vec{r}, -t)\,. \tag{5.96}$$

The two complex conjugate eigenfunctions

$$\Psi(\vec{r}, t) = \psi(\vec{r})\, e^{-iEt/\hbar}\,, \qquad \Psi^*(\vec{r}, t) = \psi^*(\vec{r})\, e^{+iEt/\hbar}\,. \tag{5.97}$$

share the same energy eigenvalue.

Actually, the time reversal symmetry transformations is not included in the point symmetry transformations of the C_3 group considered in Section 5.4. If the time-reversal transformation is added to the symmetry transformations of the C_3 group, the two functions $(e^{im\phi}, e^{-im\phi})$ are reciprocally transformed one into the other by time reversal and become the basis of a two-dimensional irreducible representation.
That's one of the reasons why the two one-dimensional irreducible representations of the group C_3 are labeled as a doubly degenerate two-dimensional representation in its table of characters, Table 5.2. A similar situation is encountered for the group C_6 (Table 5.4) and in general for the groups C_N.
Eliminating the time-reversal symmetry corresponds to breaking a higher symmetry and eliminating the degeneracy of energy levels.

Another important example of time reversal symmetry concerns the systems with translational symmetry. In Section 5.4 we have found that the translations group of a linear chain under periodic boundary conditions has only one-dimensional irreducible representations, whose more general basis functions are Bloch functions (5.64)

$$\psi(x) \;=\; e^{ik_m x} \sum\nolimits_n A_n e^{inbx} \,,$$

The Bloch functions (5.64) corresponding to pairs of opposite values $(k_m, -k_m)$ in the first Brillouin Zone share the same eigenvalue of the Hamiltonian. They correspond to waves traveling in opposite directions.

Note: While all symmetry operators up to now encountered are linear, the time reversal operator K is anti-linear. Actually, since $K\vec{p} = -\vec{p}$ but $K\vec{r} = \vec{r}$, time reversal changes the sign of the canonical commutation relation $[\vec{r}, \vec{p}]$. Otherwise stated, $Ki = -iK$, which is a property of anti-linear operators: $Kc = c^* K$.

"Umbrella inversion" of the ammonia molecule

Let us now consider the ammonia molecule NH_3 (Figs. 4.2 and 5.1). The Hamiltonian of the ammonia molecule is invariant not only with respect to the six point symmetry transformations of the C_{3v} group, previously considered (Section 4.3), but also with respect to the inversion transformation where the nitrogen atom is displaced to the symmetric site with respect to the basal plane of the molecule (Fig. 5.6).

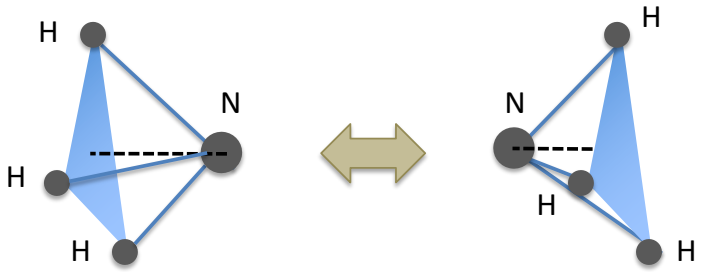

Figure 5.6: Ammonia molecule NH_3. The nitrogen atom can assume two different positions with respect to the basal plane of hydrogen atoms.

The two positions of the nitrogen atom correspond to the same energy and are equivalent. Actually, the nitrogen atom oscillates between the two positions, tunnelling the potential barrier, with a frequency of about 3×10^{10} Hz.
Considering the molecule with a well defined configuration, as was done in Section 4.3, say neglecting the tunnelling of the nitrogen atom with respect to the basal plane, means breaking the symmetry of the Hamiltonian with respect to the two equivalent positions of the nitrogen atom.

For NH_3 the symmetry breaking is effective only for obsrvation times shorter than about 10^{10} s. For comparison, let us consider two iso-structural molecules, such as PH_3 and PF_3, of larger masses. The inversion frequency decreases when the mass increases. For PF_3 the frequency is so low that it cannot be experimentally measured; symmetry breaking is always present.

Accidental degeneracy

In connection with the symmetry properties of a system, one can distinguish two types of degeneracy of the quantum energy levels:

1. *Essential degeneracy*, which arises from some kind of symmetry of the system, with respect to which the Hamiltonian is invariant.
2. *Accidental degeneracy*, which arises from special features of the system or from the particular analytic form of the potential energy (e.g. $\propto 1/r$ or $\propto r^2$).

Let's consider some examples.

Example 1: The hydrogen atom.
The electron is subject to a pure Coulomb potential $\propto 1/r$. The energy levels only depend on the principal quantum number n. For a given n, all states corresponding to different values of the azimuthal quantum number ℓ and to different values of the magnetic quantum number m have the same energy.
The degeneracy with respect to m is essential, since it is connected to the absence of preferred spatial direction.
The degeneracy with respect to ℓ is generally classified as accidental, since it depends on the peculiar form of the Coulomb potential $\propto 1/r$.

Example 2: The isotropic three-dimensional harmonic oscillator.
The space of the states of the three-dimensional harmonic oscillator is the tensor product of the three states associated with one-dimensional harmonic oscillators. The stationary Schrödinger equation is

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}\right)+\frac{1}{2}m\omega^2\left(x^2+y^2+z^2\right)\psi \;=\; E\psi \tag{5.98}$$

and the energy eigenvalues are

$$E_n \;=\; (n+3/2)\,\hbar\omega\,, \qquad \text{where } n=n_x+n_y+n_z\,. \tag{5.99}$$

The eigenvalue E_n is degenerate, the degree of accidental degeneracy being $(n+1)(n+2)/2$.

5.7 Molecular Orbitals and degeneracy removal

In this and in the following sections we analyse some situations where the theory of symmetry groups and of their irreducible representations is of help in simplifying the calculations of the structural and electronic properties of molecules and crystals.
As a first example, in this Section 5.7 we show how the group theory can be of help in determining some ground-state properties of the methane molecule CH_4 and in evaluating the effects of external perturbations.

5.7.1 Molecular Orbitals of the methane molecule

The electronic structure of molecules can be described in terms of Molecular Orbitals (MO), say electron wave-functions extending over the entire molecule (see Section 6.5 for details). In principle, the MOs can be obtained by solving the Schrödinger equation, and are thus eigenfunctions of the Hamiltonian, which can correspond to degenerate or non-degenerate eigenvalues. According to the considerations of previous sections, MOs are basis functions for irreducible representations of the symmetry group of the molecule.

An approximate procedure for calculating MOs is the linear combination of atomic orbitals (LCAO); the LCAO must in any case be consistent with the symmetry properties of the molecule.

The symmetry transformations of the methane molecule belong to the T_d group (the symmetry group of a tetrahedron, see Section 4.5); we will first study the character table of the T_d group, then we will show how symmetry considerations can give information on the MOs.

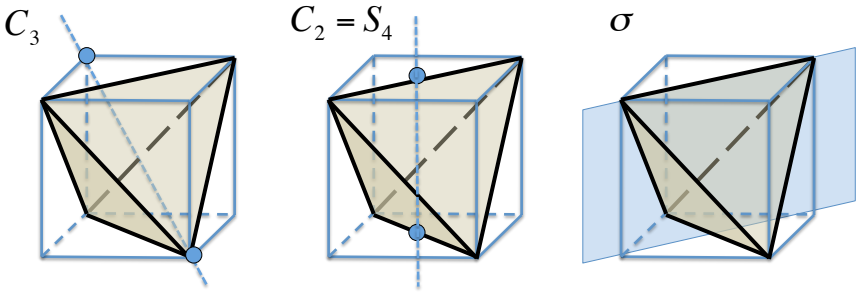

Figure 5.7: The symmetry elements of the T_d group (symmetry group of a tetrahedron): $4\,C_3$ axes, $3\,C_2 (= S_4)$ axes and $6\,\sigma$ reflection planes.

Properties of the T_d group

The symmetry elements of the T_d group are better identified by inscribing the tetrahedron in a cube (Fig. 5.7); they are:

- 4 C_3 axes through the main diagonals of the cube
- 3 C_2 axes, corresponding to S_4 axes, perpendicular to the cube faces
- 6 reflection planes σ containing the face diagonals of the cube

One can verify that the $n = 24$ symmetry operations (group elements) are grouped into five classes:

1) 1 identity e
2) 4 rotations C_3^1 and 4 rotations C_3^2 around the C_3 axes
3) 3 rotations C_2 around the C_2 axes
4) 6 reflections σ
5) 3 roto-reflections S_4^1 and 3 roto-reflections S_4^3 around the C_2 axes

The number of irreducible representations is five, equal to the number of classes (Section 5.3). The sum rule (5.25),

$$\sum_\alpha \ell_\alpha^2 \;=\; n\,,$$

can be satisfied only if there are two one-dimensional, one two-dimensional and two three-dimensional irreducible representations. This means that the maximum possible degeneracy of an energy level, for a system belonging to the discrete group T_d, is three.
The characters of the five irreducible representations of the T_d group are shown in Table 5.5.

LCAO and Molecular orbitals

The electronic configurations of the five atoms of the CH_4 molecule are

$$\mathrm{H}: \; 1s^1 \qquad\qquad \mathrm{C}: 1s^2\, 2s^2\, 2p^2$$

so that only s and p atomic orbitals can participate to the chemical bond.

It is our goal now to see how the LCAO of the $1s$ atomic orbitals of H and the $2s$ and $2p$ atomic orbitals of C can give rise to MOs consistent with the symmetry properties of the molecule.

Table 5.5: Character table of the symmetry group T_d. The irreducible representations are: two one-dimensional A, one two-dimensional E and two three-dimensional T. The last column of each row lists the atomic orbitals that are bases of the corresponding irreducible representation.

T_d	e	$8C_3$	$3C_2$	6σ	$6S_4$		orbitals
A_1	1	1	1	1	1	$x^2+y^2+z^2$	s
A_2	1	1	1	-1	-1		
E	2	-1	2	0	0	$(2z^2-x^2-y^2, x^2-y^2)$	$(d_{z^2}, d_{x^2-y^2})$
T_1	3	0	-1	-1	1	(R_x, R_y, R_z)	
T_2	3	0	-1	1	-1	(x,y,z); (xy, xz, yz)	$(p_x, p_y, p_z)(d_{xy}, d_{xz}, d_{yz})$

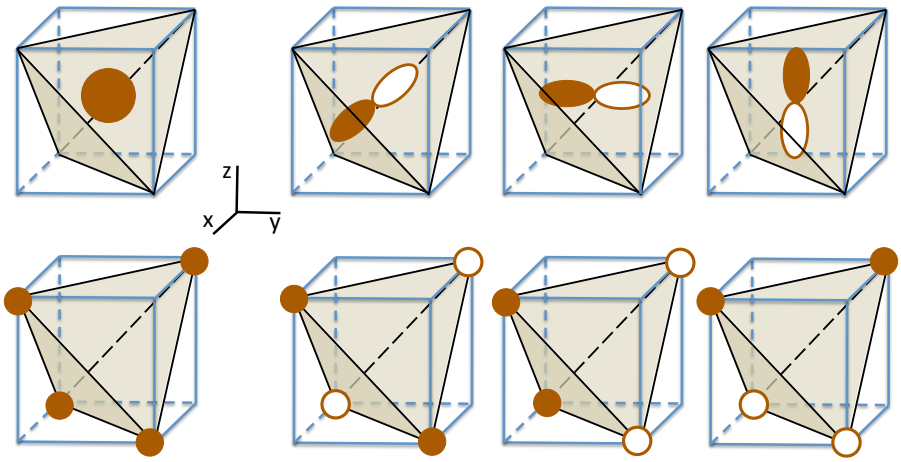

Figure 5.8: Methane molecule. Top row, from left to right: $2s, 2p_x, 2p_y, 2p_z$ atomic orbitals of carbon. Bottom row: linear combinations (5.101)-(5.104) of $1s$ atomic orbitals of hydrogen; from left to right $\Psi(A_1), \Psi(T_2)_1, \Psi(T_2)_2, \Psi(T_2)_3$.

A) The $2s$ orbital of C, which contains two electrons, is a basis of the A_1 irreducible representation: it is unaffected by any symmetry operation of the T_d group (Fig. 5.8, top left).

B) The p_x, p_y, p_z orbitals of C (Fig. 5.8, top right), which contain a total of two electrons, form a basis of the T_2 irreducible representation (see the last column of Table 5.5). The symmetry operations of the T_d group transform the p orbitals into linear combinations of p orbitals.

C) More attention deserve the four $1s$ orbitals of H.
By taking the four $1s$ orbitals of H as basis functions (see the example of the NH_3 moleule in Section 5.5), one obtains a 4×4 representation of the T_d group, whose characters can easily be calculated by counting the number of orbitals that are invariant for each one of the symmetry operations:

$$\chi(e) = 4\,,\ \chi(C_3) = 1\,,\ \chi(C_2) = 0\,,\ \chi(\sigma) = 2\,,\ \chi(S_4) = 0 \qquad (5.100)$$

This 4×4 representation is clearly reducible. Making use of the reduction formula (5.36)

$$n^{(\alpha)} = \frac{1}{n}\sum_g \left\{\chi(g)\left[\chi^{(\alpha)}(g)\right]^*\right\}\,,$$

and considering the characters of all the possible irreducible representations listed in Table 5.5, one can see that the 4×4 representation can be decomposed into the irreducible representations A_1 and T_2. Convenient basis functions for the A_1 and T_2 representations can be

obtained by the following linear combinations of the four H $1s$ orbitals ψ_i (Fig. 5.8, bottom row):

$$\Psi(A_1) \propto \psi_1 + \psi_2 + \psi_3 + \psi_4 \tag{5.101}$$
$$\Psi(T_2)_1 \propto \psi_1 - \psi_2 + \psi_3 - \psi_4 \tag{5.102}$$
$$\Psi(T_2)_2 \propto \psi_1 - \psi_2 - \psi_3 + \psi_4 \tag{5.103}$$
$$\Psi(T_2)_3 \propto \psi_1 + \psi_2 - \psi_3 - \psi_4 \tag{5.104}$$

The atomic orbitals of the C and H atoms give thus rise to bases of the A_1 and T_2 irreducible representations of the T_d group.
As a consequence, the MOs of the methane molecule, obtained by LCAO of the atomic orbitals of the C and H atoms, must be basis functions of the same two irreducible representations A_1 and T_2:

a) one bonding MO a_1, basis of the A_1 irreducible representation, formed by the LCAO of the orbitals $2s$ of C and $\Psi(A_1)$ from the four $1s$ orbitals of H atoms;
b) three bonding MOs t_2, bases of the T_2 irreducible representation, formed by the LCAO of the orbitals $2p$ of C and $\Psi(T_2)$ from the four $1s$ orbitals of H atoms.

The two MOs correspond to two different bonding energy levels: the lowest one (a_1) is non-degenerate and can accommodate two electrons, the highest one (t_2) is three-fold degenerate and can accommodate six electrons. Note that the group theory gives information on the number and degeneracy of the energy levels, not on their energy values.

Note: A different approach to the electronic structure of molecules, the method of valence orbitals based on the concept of hybridisation, will be considered in § 6.5.

5.7.2 Effect of external perturbations on degenerate levels

It is well known that the degeneracy of an energy level can be reduced or removed by a suitable external perturbation. The theory of irreducible representations allows one to solve the following problem: can a given external perturbation reduce or remove the degeneracy, by partially or totally splitting the unperturbed degenerate energy level?

Unperturbed Hamiltonian

Let us consider a system whose symmetry operations form the group $\mathcal{G}$.
The unperturbed Hamiltonian H_0, say the Hamiltonian of the systems in the absence of any external perturbation, is invariant with respect to all symmetry operations of the group $\mathcal{G}$.
Let us suppose that the unperturbed Hamiltonian has at least one degenerate energy level E_m, with degeneracy ℓ_m. The basis functions of the subspace belonging to the degenerate level E_m are a basis of an irreducible representation of dimensionality ℓ_m of the symmetry group $\mathcal{G}$ of the system.

Perturbation effects

Let us now introduce an external perturbation, described by a potential energy V. The total Hamiltonian of the perturbed system is $H = H_0 + V$.
The perturbation V can induce a splitting of the degenerate level E_m if the symmetry of V is lower than the symmetry of H_0, say if the symmetry operations that leave V unchanged represent a proper subgroup $\mathcal{S}$ of the group $\mathcal{G}$. The total Hamiltonian H has then the symmetry of V, lower than the symmetry of H_0.
In such a case, the elements of the subgroup $\mathcal{S}$ are still represented, in the basis of the eigenfunctions of E_m, by $\ell_m \times \ell_m$ matrices. It can happen, however, that these matrices are blocked, so that the $\ell_m \times \ell_m$ representation of the subgroup $\mathcal{S}$ is reducible. Its reduction gives rise to irreducible representations of smaller dimensionalities, say to the splitting of the original degenerate energy level E_m into levels of lower degeneracy.

The determination of the new representations can be accomplished by comparing the tables of characters of the groups $\mathcal{G}$ and $\mathcal{S}$ (see the following example).

Example: lower-symmetry perturbation to the T_d group

Let us again consider the methane molecule CH_4. The previous study has led us to find the two molecular orbitals a_1 and t_2.
Let us here focus our attention on the three-fold degenerate level t_2, whose wavefunctions are the basis of the three-dimensional irreducible representation T_2 (last row in Table 5.5).

Let us submit the system to an external perturbation (for example an electric or magnetic field) of C_{3v} symmetry. One can verify that C_{3v} is a subgroup of T_d, since all its six symmetry operations $(e, 2C_3, 3\sigma)$ belong also to T_d. The symmetry of the system is lowered from T_d to C_{3v}.
The three wavefunctions of the three-fold degenerate level t_2 of the unperturbed system are a basis also of a 3×3 representation of the C_{3v} group. The characters of the three classes are the same as for the T_d group (see Table 5.6).

Table 5.6: Characters of the 3×3 irreducible representation T_2 of the symmetry group T_d (from last line of Table 5.5) and characters of the 3×3 reducible representation of the group C_{3v} in the same basis set.

T_d	e	$8C_3$	$3C_2$	6σ	$6S_4$
T_2	**3**	**0**	-1	**1**	-1

C_{3v}	e	$2C_3$	3σ
	3	**0**	**1**

It is immediate to verify that the 3×3 representation of the C_{3v} group based on the eigenfunctions of the t_2 level is reducible, since $\sum |\chi|^2 = 12 \neq g = 6$.
By means of the reduction formula (5.36), one can see that the 3×3 reducible representation can be decomposed into a one-dimensional irreducible representation A_1 and a two-dimensional irreducible representation E. Otherwise stated, the three-fold degenerate level t_2 is split into a non-degenerate level and a two-fold degenerate level.

Note again that the group theory cannot give information on the energy values of the new energy levels.

5.8 Crystal field splitting

Another interesting application of group theory is the splitting of the d energy levels of transition metal atoms in complexes where the metal atom is coordinated to four or six other atoms (typically oxygens) in tetrahedral or octahedral coordination.
In this case, the starting system is an atom, and we have thus the opportunity of exploring some aspects of the quite complex treatment of spherical symmetry.

5.8.1 The group $SO(3)$ of proper rotations

Let us first consider a free atom with perfect spherical symmetry and focus, for the moment, only on the proper rotations. The group of all proper rotations in the three-dimensional space is $SO(3)$ (Section 4.5). The group in infinite: there are

a) infinite rotation axes
b) infinite values of rotation angles α around each axis.

All rotations by the same angle α around all possible axes belong to the same class, because any two rotation axes are connected by the rotation around a third axis (conjugation relation). The number of classes is thus infinite, since infinite are the values of α; as a consequence, infinite is the number of irreducible representations.
The most convenient basis functions are the spherical harmonics, say the eigenfunctions of the angular momentum

$$Y_L^M(\theta,\phi) \;\propto\; P_L^M(\theta)\, e^{iM\phi}\,, \tag{5.105}$$

where $P_\ell^M(\theta)$ are Legendre polynomials. The spherical coordinates refer to a well defined orientation of the cartesian xyz axes: ϕ is the angle of rotation around the z axis.
Each value of the azimuthal quantum number L corresponds to a different value of total orbital angular momentum, say to a different energy level (with the exception of the hydrogen atom, where the energy of levels only depends on the principal quantum number n). To each value of L there correspond $2L+1$ values of the magnetic quantum number M. The energy eigenvalue is thus $(2L+1)$-fold degenerate.

For each value of L, the corresponding spherical harmonics $Y_L^M(\theta,\phi)$ are the basis of a $(2L+1)$-dimensional irreducible representations of the infinite $SO(3)$ group. Since the values of L are unlimited, there are infinite possible irreducible representations of increasing dimensionality.
The representation in the basis of $(2L+1)$ spherical harmonics corresponding to a given value of L is reducible, since the rotations mix up the spherical harmonics.
A particular case is the one of rotations around the z axis. Let us consider the effect of a rotation by an angle α around the z axis. One can show that

$$P_\alpha\, Y_L^M(\theta,\phi) \;=\; Y_L^M(\theta,\phi-\alpha) \;=\; e^{-iM\alpha}\, Y_L^M(\theta,\phi)\,. \tag{5.106}$$

Each rotation by an angle α is represented by a $(2L+2)\times(2L+1)$ *diagonal* matrix

$$\mathbf{D}^{(L)}(\alpha) \;=\; \begin{pmatrix} e^{-iL\alpha} & 0 & \dots & 0 \\ 0 & e^{-i(L-1)\alpha} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{-iL\alpha} \end{pmatrix} \tag{5.107}$$

There are infinite such matrices for a given representation L, each one corresponding to a different α value.
The character of the class of an α rotations, say the trace of $\mathbf{D}^{(L)}(\alpha)$, is

$$\begin{aligned} \chi^{(L)}(\alpha) \;&=\; e^{-iL\alpha}\sum_{k=0}^{2L}(e^{i\alpha})^k \;=\; e^{-iL\alpha}\,\frac{e^{i(2L+1)\alpha}-1}{e^{i\alpha}-1} \\ &=\; \frac{e^{i(L+1/2)\alpha}-e^{-i(L+1/2)\alpha}}{e^{i\alpha/2}-e^{-i\alpha/2}} \;=\frac{\sin[(L+1/2)\alpha]}{\sin(\alpha/2)}\,. \end{aligned} \tag{5.108}$$

Since all rotations by the same angle α around different axes belong to the same class, the character evaluated from the diagonal matrix of the rotation around the z axis is the same for α rotations around every other axis.

Example 1: Titanium atom.

The electronic configuration is $[\mathrm{Ar}]\,3d^2\,4s^2$. The ground state term is ${}^3F_2$, which corresponds to $L=3$ and $S=1$. Since $L=3$, the d level is $2L+1=7$-fold degenerate and the I.R. based on the spherical harmonics is 7-dimensional. The characters of rotations are

$$\chi^{(L=3)}(\alpha) \;=\; \frac{\sin(7\alpha/2)}{\sin(\alpha/2)}\,. \tag{5.109}$$

Example 2: Iron atom.

The electronic configuration is $[\mathrm{Ar}]\,3d^6\,4s^2$. The ground state term is ${}^5D_4$, which corresponds

to $L = 2$ and $S = 2$. Since $L = 2$, the d level is $2L + 1 = 5$-fold degenerate and the I.R. based on the spherical harmonics is 5-dimensional. The characters of rotations are

$$\chi^{(L=3)}(\alpha) = \frac{\sin(5\alpha/2)}{\sin(\alpha/2)} . \tag{5.110}$$

(?) Plot the function $\chi^{(L)}(\alpha)$ in the interval $0 < \alpha \leq 2\pi$ for different values of L.

5.8.2 Symmetry breaking by external perturbation

External field of octahedral symmetry

When the atom is inserted in an external field of octahedral symmetry, the spherical symmetry is broken. The group $SO(3)$ of the proper rotations of the sphere is substituted by the group O of the proper rotations of the cube (or of the octahedron, see Section 4.5).
The group O is of order $n = 24$. The character table is shown in Table 5.7. There are five classes and five I.R. The classes $3C_2$ and $6C_4$ contain rotations around the same axes C_4.
At the bottom of the table, a line is added with the values of the rotation angle α for the five classes and the last two lines contain the characters of the five classes in the representations of the spherical harmonics for $L = 2$ (Iron) and $L = 3$ (Titanium).
These two representations are clearly reducible, as one can check by means of the sum rules for characters.

Table 5.7: Character table of the symmetry group O. The I.R. are: two one-dimensional A, one two-dimensional E and two three-dimensional T. The last column lists the atomic orbitals that are bases of the corresponding I.R.

O	e	$8C_3$	$3C_2$	$6C_2$	$6C_4$		orbitals
A_1	1	1	1	1	1		
A_2	1	1	1	-1	-1		
E	2	-1	2	0	0		$(d_{z^2}, d_{x^2-y^2})$
T_1	3	0	-1	-1	1	$(x, y, z), (R_x, R_y, R_z)$	
T_2	3	0	-1	1	-1	(xy, xz, yz)	$(p_x, p_y, p_z)(d_{xy}, d_{xz}, d_{yz})$
α	0	$2\pi/3$	π	π	$\pi/2$		
$\Gamma^{L=2}$	5	-1	1	1	-1		
$\Gamma^{L=3}$	7	1	-1	-1	-1		

The two representations $\Gamma^{L=2}$ and $\Gamma^{L=3}$ can be decomposed into I.R. by comparing their characters with the characters if the I.R. of the group O by means of the sum rule

$$n^\alpha = \frac{1}{24} \sum_{g=1}^{24} \chi^{(\Gamma)} \left[\chi^\alpha(g)\right]^* , \tag{5.111}$$

where α labels the I.R. of the group O.
The decomposition is given by the I.R. for which the sum is not zero. One finds:

$$\Gamma^{L=2} = E + T_2 , \qquad \Gamma^{L=3} = A_2 + T_1 + T_2 . \tag{5.112}$$

The 5-fold degenerate d level of Fe is split into a 2-fold degenerate level e_g (orbitals d_{z^2} and $d_{x^2-y^2}$ and a 3-fold degenerate level t_g (orbitals d_{xy}, d_{xz}, d_{yz}).
The 7-fold degenerate d level of Ti is split into a non degenerate level and two different 3-fold degenerate levels.

Breaking the octahedral symmetry

Let us now consider what happens to the Iron atom ($L = 2$) if the octahedron is deformed by an elongation along one of the C_3 axes. The O group is reduced to the lower symmetry D_3 group, whose order is $n = 6$.
As one can see from Table 5.8, only a limited number of the rotations of the O group are present in the D_3 group.
The two-dimensional representation E is an I.R. for the D_3 group too.
The three-dimensional representation T_2 is not an I.R. for D_3. One can verify that it can be decomposed as $T_2 = E + A_1$.
The initial 5-fold degenerate d level of the free atom is thus split into two 2-fold degenerate e levels, whose eigenfunctions are basis of E representations, and a non-degenerate a_1 level, whose eigenfunction is basis of the A_1 representation.

Table 5.8: *Top:* I.R. of the group O based on the eigenfunctions of the $L = 2$ angular momentum. *Bottom:* I.R. of the D_3 group.

O	e	$8C_3$	$3C_2$	$6C_2$	$6C_4$
E	2	-1	2	0	0
T_2	3	0	-1	1	-1

D_3	e	$2C_3$	=	$3C_2$	=
A_1	1	1		1	
A_2	1	1		-1	
E	2	-1		0	

The effect of inversion symmetry

Up to now, we have considered only proper rotations groups, $SO(3)$, O and D_3. For a free atom or an atom in octahedral symmetry the improper rotations should be considered too.
The full symmetry groups $O(3)$ and O_h can be obtained by the direct product of the proper symmetry groups with the inversion group $C_i = \{e, i\}$:

$$S(3) = SO(3) \otimes C_1\,, \qquad O_h = O \otimes C_i\,. \tag{5.113}$$

Table 5.9: Character table of the C_i group. The two I.R. are labeled as g (*gerade*, German for even) and u (*ungerade*, German for odd), according to the parity of the basis functions; alternatively, they can be labeled $+$ and $-$.

C_i	e	i	
g	1	1	even function
u	1	-1	odd function

The group C_i has two elements, two classes and two one-dimensional I.R. The characters are shown in Table 5.9. Moreover, the two elements of the group C_i commute with all the proper rotations of the $SO(3)$ and O groups.

One can demonstrate the following general properties of the direct product of two groups:

a) the direct product of two I.R. of the component groups forms an I.R. of the direct product group;

b) all the I.R. of the direct product group are direct products of the I.R. of the component groups, so that the number of I.R. of the direct product group is equal to the product of the number of I.R. of the component groups;
c) the number of classes of the direct product group is equal to the product of the number of classes of the component groups;
d) the characters of the I.R. of the product group are equal to the product of the characters of the component groups.

Table 5.10: Character tables of the O and C_i groups, whose direct product is the O_h group.

O	e	$8C_3$	$3C_2$	$6C_2$	$6C_4$
A_1	1	1	1	1	1
A_2	1	1	1	-1	-1
E	2	-1	2	0	0
T_1	3	0	-1	-1	1
T_2	3	0	-1	1	-1

$\otimes$

C_i	e	i
g	1	1
u	1	-1

Let us consider the direct product of the discrete groups $O_h = O \otimes C_i$ (Table 5.10).
The group O_h has:

$24 \times 2 = 48$ elements
$5 \times 2 = 10$ classes
$5 \times 2 = 10$ irreducible representations.

The basis functions of the I.R. of the O_h group are obtained as products of the basis functions of the O group by the even and odd basis functions of the C_i group.
The characters are the products of characters:

$$\chi_{O_h}(g) = \chi_O(g)\ \chi_{C_i}(g)\,. \tag{5.114}$$

The 5×5 character table of the O group becomes a 10×10 character table of the O_h group. A schematic representation is given in Table 5.11.

Table 5.11: Character table of the O_h group. The last line lists the characters of the ten classes in the representation of the spherical harmonics for $L = 2$ and $\chi(i) = 1$.

O_h	e	$8C_3$	$3C_2$	$6C_2$	$6C_4$	i	$8S_6$	$3\sigma_h$	$6\sigma_v$	$6S_4$
A_{1g}	1	1	1	1	1	1	1	1	1	1
A_{2g}	1	1	1	-1	-1	1	1	1	-1	-1
E_g	2	-1	2	0	0	2	-1	2	0	0
T_{1g}	3	0	-1	-1	1	3	0	-1	-1	1
T_{2g}	3	0	-1	1	-1	3	0	-1	1	-1
A_{1u}	1	1	1	1	1	- 1	- 1	- 1	- 1	- 1
A_{2u}	1	1	1	-1	-1	- 1	- 1	- 1	1	1
E_u	2	-1	2	0	0	-2	1	-2	0	0
T_{1u}	3	0	-1	-1	1	-3	0	1	1	-1
T_{2u}	3	0	-1	1	-1	-3	0	1	-1	1
$\Gamma^{L=2}$	5	-1	1	1	-1	5	-1	1	1	-1

The left half and the right half parts of the table correspond to the eg and ig transformations, where g is the generic transformation of the group O.

The upper half and the lower half of the table correspond to basis functions that are equal to the product of the O basis functions by even (g) or odd (u) basis functions of C_i, respectively.
The upper left quadrant corresponds to the character table of the O group. The same characters χ appear in the two upper quadrants and in the left lower quadrant of the table, while opposite characters $-\chi$ appear in the lower right quadrant.

Let us come back to the problem of the free atom and the effect of its insertion in a octahedral field, considering now the effect of inversion symmetry too.
The parity of a one-electron wavefunction is $(-1)^\ell$, where ℓ is the one-electron azimuthal quantum number. For a many-electron atom, the parity is $\prod_k (-1)^{\ell_k} = \chi(i)$, where k is the number of electrons.
The character of a symmetry transformation of a free atom (proper or improper rotation) is thus

$$\chi^{(L)}(\alpha) \cdot \chi(i) \;=\; \chi^{(L)}(\alpha) \cdot \prod_k (-1)^{\ell_k} \,. \tag{5.115}$$

For the $3d$ electrons of Fe, $\ell = 2$, so that $\chi(i) = 1$ (even functions). The corresponding characters for the symmetry transformations of the O_h group are listed in the last line of Table 5.11. As one can see, the decomposition of the reducible representation $\Gamma^{L=2}$ in terms of the I.R. of the group O_h leads to the same splitting of the $3d$ level as the decomposition in terms of the I.R. of the O group of proper rotations previously considered.
In general: g terms split into g terms, u terms split into u terms.

5.9 Evaluation of integrals

An important application of the theory if irreducible group representations concerns the possibility of evaluating, from symmetry considerations, whether the integral of a function or of a product of functions can be different from zero or is necessarily zero.

5.9.1 Invariance of integrals

Let us consider a generic function $\Psi(q)$ defined in the configuration space. By q we mean here the set of all generalised coordinates necessary to define the configuration of a system: $q = \{q_1, q_2, \ldots\}$. In most cases, the functions $\Psi(q)$ we are interested in are eigenfunctions of the Schrödinger equation. We are interested in determining whether the integral

$$I \;=\; \int \Psi(q)\, dq \,, \tag{5.116}$$

extended over the entire configuration space, is necessarily zero or can be different from zero.
In a sense, we want to generalise the rule that the integrals of one-dimensional odd functions are zero.

Starting point is the fact that the integral I of any function extended over all the configuration space must be invariant with respect to all symmetry transformations of the system.

Let us focus the attention of the basis functions of irreducible representations; for concreteness, let $\Psi^\alpha_\nu(q)$ be one of the basis functions of the I.R. α.
The invariance of the integral of $\Psi^\alpha_\nu(q)$ with respect to a symmetry transformation g means that:

$$\int \Psi^{(\alpha)}_\nu(q)\, dq \;=\; \int P_g\, \Psi^{(\alpha)}_\nu(q)\, dq \;=\; \int dq \sum_{i=1}^{\ell} D^{(\alpha)}_{i\nu}(g)\, \Psi^{(\alpha)}_i(q) \tag{5.117}$$

where in the rightmost member the matrix elements of P_g have been explicitly expressed according to (5.70).

Let us now sum the first and last members of (5.117) over all the n elements g of the symmetry group; the leftmost member is simply multiplied by n, the rightmost member can be readjusted so that

$$n \int \Psi_\nu^{(\alpha)}(q)\, dq \;=\; \sum_{i=1}^{\ell} \int dq\, \Psi_i^{(\alpha)}(q) \sum_{g=1}^{n} D_{i\nu}^{(\alpha)}(g)\,. \tag{5.118}$$

It is convenient now to recall the general orthogonality theorem (5.23)

$$\sum_{g=1}^{n} \left[D_{i\nu}^{(\alpha)}(g)\right] \left[D_{j\mu}^{(\beta)}(g)\right]^* \;=\; \frac{n}{\ell_\alpha}\, \delta_{\alpha\beta}\, \delta_{ij}\, \delta_{\nu\mu}\,. \tag{5.119}$$

If $\beta \neq \alpha$, the sum in (5.119) is zero. In particular, if one chooses $\beta = A_1$ (identical representation), its representation matrices are one-dimensional and for all group elements $D^{(\beta)}(P_g) = 1$, so that (5.119) collapses to

$$\sum_{g=1}^{n} D_{i\nu}^{(\alpha)}(g) \;=\; 0\,. \tag{5.120}$$

Since we assumed $\alpha \neq \beta = A_1$, this result means that for $\alpha \neq A_1$ the integral (5.118) is zero.
To summarise:

1. If $\Psi_\nu^{(\alpha)}(q)$ in (5.117) belongs to the basis of an irreducible representation different from A_1, say if $\Psi_\nu^{(\alpha)}(q)$ is not invariant with respect to *all* symmetry transformations of the group, then the integral is zero, as a consequence of (5.118) and (5.120):

$$I \;=\; \int \Psi_\nu^{(\alpha)}(q)\, dq = 0 \qquad (\alpha \neq A_1) \tag{5.121}$$

   Example: The $S_2 = 2S_a - S_b - S_c$ and $S_3 = S_b - S_c$ functions (5.77) for the ammonia molecule are a basis of the I.R. E of the C_{3v} group, but not of the I.R. A_1. Their integral is zero (see also the schematic picture of the S_2 and S_3 function in Fig. 5.5).

2. Only if $\Psi_\nu^{(\alpha)}(q) = \Psi^{(\alpha)}(q)$ is a basis function of the identical I.R. A_1, say if $\Psi^{(\alpha)}(q) = \Psi^{(A_1)}(q)$ is invariant with respect to all symmetry transformations of the group, *can* the integral be different from zero:

$$I \;=\; \int \Psi^{(A_1)}(q)\, dq \text{ can be } \neq 0\,. \tag{5.122}$$

   Example: The S_n function and the $S_1 = S_a + S_b + S_c$ function (5.77), introduced in Section 5.5 when treating the orbital-based representations of the C_{3v} group (ammonia molecule) are both basis functions of the I.R. A_1 of C_{3v}. Their integral is different from zero (see also the schematic picture of the S_1 function in Fig. 5.5).

   Note: Be careful: the condition dictated by the symmetry properties is necessary but not sufficient. The integral could be zero or negligible for reasons not connected to the considered symmetry.

In the following, we consider two common applications of the invariance of integrals:

a) superposition integrals (interesting for example for chemical bonds),
b) matrix elements (interesting for selection rules in radiative transitions).

5.9.2 Superposition integrals

Let us now consider the integral of the product of two functions $\Psi_i(q)$ and $\Psi_j(q)$:

$$J \;=\; \int \Psi_i(q)\,\Psi_j(q)\,dq\,. \tag{5.123}$$

A typical example is the superposition integral of two atomic orbitals of a molecule.
We are interested in establishing the condition for the integral *not to be zero*. (Only atomic orbitals with superposition integral different from zero can participate to chemical bonds).

According to the previous considerations, necessary condition for the integral not to be zero,

$$J \;=\; \int \Psi_i(q)\,\Psi_j(q)\,dq \neq 0\,, \tag{5.124}$$

is that the product of the two functions $\Psi_i(q)\,\Psi_j(q)$ is a one-dimensional basis of the irreducible representation A_1 of the symmetry group of the system under consideration. Otherwise stated, the product $\Psi_i(q)\,\Psi_j(q)$ must be invariant with respect to all the transformations of the group.

Note: Again, the condition dictated by the symmetry properties is necessary but not sufficient. The integral could be zero or negligible for reasons not connected to symmetry, e.g. the negligible spatial superposition of the two functions.

Direct product of irreducible representations

To solve the problem, say to find the condition for the integral (5.124) not to be zero, it is convenient to start from a more general issue: how can we infer the symmetry properties of the product $\Psi_i(q)\,\Psi_j(q)$ of any two functions from the symmetry properties of the single functions?
To that purpose, one introduces the concept of direct product of irreducible representations.
Let the two functions $\Psi_i^{(\alpha)}(q)$ and $\Psi_j^{(\beta)}(q)$ belong to the basis sets of two I.R. α and β, respectively, of the same symmetry group:

$$\Psi_i^{(\alpha)}(q) \in \text{basis set } \{\Psi_1^{(\alpha)}, ...\Psi_{\ell_\alpha}^{(\alpha)}\} \text{ of an I.R. } \alpha$$
$$\Psi_j^{(\beta)}(q) \in \text{basis set } \{\Psi_1^{(\beta)}, ...\Psi_{\ell_\beta}^{(\beta)}\} \text{ of an I.R. } \beta$$

The two functions can be, for example, eigenfunctions of the energy levels E_α and E_β with degeneracies ℓ_α and ℓ_β, respectively.

The set of $\ell_\alpha \times \ell_\beta$ product functions $\Psi_i^{(\alpha)}(q)\,\Psi_j^{(\beta)}(q)$, $(i = 1...\ell_\alpha,\ j = 1...\ell_\beta)$ is the basis of a new representation that is called the *direct product representation* (DP) of the two irreducible representations α and β. In the DP representation, the n group elements are represented by n matrices $(\ell_\alpha\ell_\beta) \times (\ell_\alpha\ell_\beta)$.

The direct product of two I.R. is generally a *reducible* representation. One can demonstrate [Landau-Lifschiz, Quantum Mechanics, § 94] the following properties:

1. The character of any transformation in the direct product representation DP is the product of the characters of the transformation in the two component irreducible representations α and β:
$$\chi^{(DP)}(g) \;=\; \chi^{(\alpha)}(g)\,\chi^{(\beta)}(g)\,. \tag{5.125}$$
Actually, from
$$P(g)\,\Psi_\nu^{(\alpha)} \;=\; \sum_i D_{i\nu}^{(\alpha)}(g)\Psi_i^{(\alpha)}\,, \qquad P(g)\,\Psi_\mu^{(\beta)} \;=\; \sum_j D_{j\mu}^{(\beta)}(g)\Psi_j^{(\beta)}\,,$$
one finds
$$P(g)\,[\Psi_\nu^{(\alpha)}\,\Psi_\mu^{(\beta)}] \;=\; \sum_{ij} D_{i\nu}^{(\alpha)}(g)D_{j\mu}^{(\beta)}(g)\,\Psi_i^{(\alpha)}\Psi_j^{(\beta)}\,,$$

whence

$$\chi^{DP}(g) \;=\; \sum_{\nu,\mu} D^{(\alpha)}_{\nu\nu}(g)\, D^{(\beta)}_{\mu\mu}(g) \;=\; \sum_{\nu} D^{(\alpha)}_{\nu\nu}(g) \sum_{\mu} D^{(\beta)}_{\mu\mu}(g) \;=\; \chi^{(\alpha)}(g)\, \chi^{(\beta)}(g)\,.$$

2. As a consequence of (5.125) and of the sum rule (5.29), the direct product DP is by itself an I.R. only if $n = 1$ or $m = 1$, say if at least one of the two I.R. α or β is one-dimensional. In fact,

$$\sum_{g} |\chi^{DP}(g)|^2 \;=\; \sum_{g} |\chi^{(\alpha)}(g)|^2 |\chi^{(\beta)}(g)|^2 \;=\; n \tag{5.126}$$

only if $|\chi^{(\alpha)}(g)|^2 = 1$ or $|\chi^{(\beta)}(g)|^2 = 1$ for all g.

Note: Don't confuse the direct product of the irreducible representations, here introduced, with the direct product of two different groups. In Section 5.8, the direct product of different groups $O_h = O \otimes C_i$ and the corresponding I.R. were considered was considered.

Example: Let us consider the five I.R. representations of the T_d group listed in Table 5.5 on page 110. Let us consider two possible DPs, one (DP_a) product of E and T_2, the other (DP_b) product of T_2 by itself. The characters of the two direct products DP_a and DP_b are listed in Table 5.12, left and right, respectively. One can verify that both product representations, DP_a and DP_b, are reducible by comparing the sum of the squared characters with the order of the group $n = 24$.

Table 5.12: Characters of the direct product (DP) of two I.R. of the T_d group. Left: DP of E and T_2. Right: DP of T_2 by itself.

T_d	e	$8C_3$	$3C_2$	6σ	$6S_4$
E	2	-1	2	0	0
T_2	3	0	-1	1	-1
DP_a	6	0	-2	0	0

T_d	e	$8C_3$	$3C_2$	6σ	$6S_4$
T_2	3	0	-1	1	-1
T_2	3	0	-1	1	-1
DP_b	9	0	1	1	1

I.R. direct product and superposition integrals

Let us now come back to our initial problem: we know that necessary condition for the superposition integral (5.124) to be non-zero is that the product of the two functions $\Psi_i(q)\,\Psi_j(q)$ is a basis of the identical I.R. A_1.
This is possible only if the identical I.R. A_1 is a component of the DP representation based on the two functions $\Psi_i(q)\,\Psi_j(q)$.

One can demonstrate the two following relevant properties:

1. The DP of two *different* irreducible representations ($\alpha \neq \beta$) doesn't contain the irreducible representation A_1. This statement can be easily deduced from (5.36):

$$n^{(A_1)} \;=\; \frac{1}{n} \sum_{g=1}^{n} \left\{ \chi^{DP}(g) \left[\chi^{(A_1)}(g) \right]^* \right\}, \tag{5.127}$$

substituting $\chi^{(DP)}(g) = \chi^{(\alpha)}(g)\,\chi^{(\beta)}(g)$ and $\chi^{(A_1)}(g) = 1$ and applying the sum rule (5.29). If $\alpha \neq \beta$, the sum is zero, so that A_1 is not a component of the DP; there is no possibility that the product $\Psi^{(\alpha)}_i(q)\,\Psi^{(\beta)}_j(q)$ is a basis of A_1, and as a consequence the integral $J = 0$.

2. The DP of an irreducible representation by itself ($\alpha = \beta$) contains the irreducible representation A_1. In fact, making again use of (5.36) and (5.125) and of $\chi^{(A_1)}(g) = 1$ one finds:

$$n^{(A_1)} = \frac{1}{n}\sum_{g=1}^{n} |\chi^{(\alpha)}|^2 \left[\chi^{(A_1)}(g)\right]^* = \frac{1}{n}\sum_{g=1}^{n} |\chi^{(\alpha)}|^2 = 1\,. \tag{5.128}$$

If $\alpha = \beta$, the product $\Psi_i^{(\alpha)}(q)\,\Psi_j^{(\alpha)}(q)$ *can be* a basis of A_1, and it is possible that the integral $J \neq 0$.

Only in the second case ($\alpha = \beta$) it is thus possible that

$$J = \int \Psi_i^{(\alpha)}(q)\,\Psi_j^{(\alpha)}(q)\,dq \neq 0\,. \tag{5.129}$$

Otherwise stated, necessary (but not sufficient) condition for the superposition integral to be different from zero is that the two functions $\Psi_i(q)$ and $\Psi_j(q)$ belong to the basis of the same I.R., say have the same symmetry properties.

A simple procedure to check the above results is to consider the table of characters of the two irreducible representations α and β and add a bottom row containing the characters of their direct product (DP), obtained by (5.125). One then verifies whether the DP contains A_1 or not by means of (5.36).

Example 1: Let us consider again the two DPs formed by I.R. of the T_d group and listed in Table 5.12. It is easy to verify that the DP_a doesn't contain A_1: the basis functions of E and T_2 have different symmetry properties, and the integral of their product is necessarily zero. The DP_b instead contains A_1; the two factor functions belong to the same I.R. T_2 and the integral of their product can be different from zero.

Example 2: Let us again consider the ammonia molecule NH_3.

Both functions S_n and $S_1 = S_a + S_b + S_c$ introduced in Section 5.5 are bases of the same representation A_1: the integral of their product can be different from zero.

The functions S_n and $S_2 = 2S_a - S_b - S_c$ (or S_n and $S_3 = S_b - S_c$) belong instead to two different I.R., A_1 and E, respectively. The integral of their product is necessarily zero. Consider the table of characters of the two irreducible representations A_1 and E of the C_{3v} group, and add a bottom row containing the characters of their direct product (DP):

	e	$2C_3$	$3\sigma_v$
$\alpha = A_1$	1	1	1
$\beta = E$	2	−1	0
DP	2	−1	0

By (5.36) one easily verifies that the DP doesn't contain the A_1 unit representation: $n^{A_1} = 0$.

5.10 Matrix elements and selection rules

Let O be an operator. We want to know whether one of its matrix elements

$$O_{ij}^{(\alpha\beta)} = \int \Psi_i^{(\alpha)*}\, O\, \Psi_j^{(\beta)}\,dq\,. \tag{5.130}$$

is necessarily zero or can be non zero.

In (5.130), α and β label the energy levels, or equivalently the irreducible representations of the symmetry group; i and j label the eigenfunctions of an energy level, or equivalently the basis functions of the corresponding I.R.

The matrix element $O_{ij}^{(\alpha\beta)}$, say the integral of the right member of (5.130), can be different from zero only if the integrand is a basis of the identical I.R. A_1. We have then to take into account the symmetry properties not only of the two functions $\Psi_i^{(\alpha)*}$ and $\Psi_j^{(\beta)}$, but also of the operator O, in order to consider the direct product of their I.R.

5.10.1 Scalar and vector operators

Scalar operators

A scalar operator, such as a potential energy, is invariant with respect to all symmetry transformations. It is thus a basis of the identical I.R. A_1. The characters of the DP of the three I.R. are equal products of the characters of the I.R. of the two functions. One is thus reduced to the previous case of the integral of two functions.
The integral matrix element $O_{mn}^{(\alpha\beta)}$ can be non-zero only if $\alpha = \beta$, say if the two functions belong to the basis of the same I.R.

Vector operators: electric and magnetic dipole

In a number of problems we need to calculate integrals (5.130) where O is a vector operator.

A very important case is represented by the time-dependent perturbation theory applied to the processes of emission and absorption of electromagnetic radiation by atoms. To first-order approximation, the time-dependent perturbation theory leads to the so-called Golden Rule: the probability per unit time of transition from an initial stationary state $|\Psi_i\rangle$ to a final stationary state $|\Psi_f\rangle$ of the unperturbed Hamiltonian of the system, due to the presence of the time-dependent perturbation H_{int}, is

$$\begin{aligned} w_{fi} &= \frac{2\pi}{\hbar}\,\rho(E_f)\;|\langle\Psi_f|H_{\text{int}}|\Psi_i\rangle|^2 \\ &= \frac{2\pi}{\hbar}\,\rho(E_f)\;\left|\int \Psi_f(q)\,H_{\text{int}}\,\Psi_i(q)\,dq\right|^2 , \end{aligned} \tag{5.131}$$

where $\rho(E_f)$ is the density of final states; the second line of the equation is the expression of the Golden Rule in the coordinate representation.

Let us now consider two simple examples of interaction of electromagnetic radiation with matter.

Example 1: Electric-dipole transitions

To first-order in the interaction, the interaction Hamiltonian for one electron is $H_{\text{int}} = (e/m)\vec{p}\cdot\vec{A}$, where $\vec{A}$ is the vector potential of the electromagnetic field and $\vec{p}$ id the canonical momentum of the electron.
If the radiation wavelength is much larger than the atomic size, the electric dipole approximation holds, and the spatial variation of the vector potential can be neglected, so that $H_{\text{int}} \propto e\vec{p}\cdot\hat{\epsilon} \propto \omega^2 e\vec{r}\cdot\hat{\epsilon}$, where $\hat{\epsilon}$ is the polarisation unit vector of the electromagnetic field.
The dipole operators $e\vec{r}$ or $e\vec{p}$ are polar vectors, whose components transform like the coordinates (x, y, z). They change sign under inversion.

Example 2: Magnetic dipole transitions

In the dipole approximation, $H_{\text{int}} \propto \vec{\mu}\cdot\vec{B}$, where $\vec{\mu} = \mu_B(2\vec{S} + \vec{L})$, and μ_B is the Bohr magneton. The magnetic dipole operator $\vec{\mu}$ is an axial vector: it transforms as a polar vector under proper rotations but is invariant under inversion. The components of the axial vectors are labelled as (R_x, R_y, R_z).

5.10.2 Transformation properties of vector operators

The three cartesian components of the electric or magnetic dipole operators (more generally of polar or axial vectors) form the basis of a 3-dimensional representation of the symmetry group of a given system.
Polar and axial vectors are characterised by the same behaviour with respect to proper rotations, by a different behaviour with respect to improper rotations (inversion, mirror reflections).
Let us consider some general properties of the 3-dimensional representations in terms of a generic rotation axis ϕ, and in particular let us focus on groups with inversion symmetry, where improper rotations can be obtained as products of proper rotations and inversion. Our goal is to find the character as a function of the angle ϕ. The character is invariant with respect to the orientation

of the rotation axis; so it is convenient to choose a particularly simple configuration: one of the rotation axes coinciding with the z axis.

1. *Polar vectors* (V).
   Polar vectors change sign under inversion. The representation matrices for a proper rotation and for the product of a proper rotation with inversion are different, the characters of all elements differing by the sign.

$$\mathbf{D}^V(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{D}^V(\phi, i) = \begin{pmatrix} -\cos\phi & \sin\phi & 0 \\ -\sin\phi & -\cos\phi & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{5.132}$$

   The corresponding characters are

$$\chi^V(\phi) = 1 + 2\cos\phi\,, \qquad \chi^V(\phi, i) = -1 - 2\cos\phi \tag{5.133}$$

2. *Axial vectors* (A).
   Axial vectors are invariant under inversion. The representation matrices for a proper rotation and for the product of a proper rotation with inversion are equal.

$$\mathbf{D}^A(\phi) \;=\; \mathbf{D}^A(\phi, i) \;=\; \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.134}$$

   The corresponding characters are

$$\chi^A(\phi) \;=\; \chi^A(\phi, i) \;=\; 1 + 2\cos\phi\,. \tag{5.135}$$

Note: The reducibility of the 3×3 representations depends on the peculiar symmetry group to be considered (see below).

We will now see how these general properties apply to different symmetry groups; we will distinguish the groups without inversion symmetry (where both proper and improper rotation can be present) and the groups with inversion symmetry (where the improper rotations are the product of proper rotations and inversion).

Groups without inversion symmetry

For groups without inversion symmetry, the general expression for the characters of proper rotations is the same for transformations of polar and axial vectors: $\chi(\phi) = 1 + 2\cos\phi$.
For different groups, however, different are the number of symmetry axes and of the allowed rotation angles; in addition, mirror reflections can be present; as a consequence, different are the values of the characters and the dimensionality of the I.R. based on the components of polar and axial vectors.
The transformation properties of the vector operators and the associated irreducible representations can be obtained, for each group, by inspection of the character tables. Let us consider some examples.

- C_{3v} group (character table 5.1 on page 90).
  The group has only one C_3 axis with $\phi = 2\pi/3$. Mirror reflections represent improper rotations. The 3×3 representation is reducible.
  The z component of the polar vectors is a basis of the A_1 representation, the z component of the axial vectors is a basis of the A_2 representation, the difference in characters only concerning mirror reflections.
  The (x, y) components of both polar and axial vectors are a basis of the two-dimensional E representation (they are not affected by mirror reflections).

- D_3 group (character table 5.8, bottom, on page 115).
  The group D_3 is isomorphic to the group C_{3v}, and thus shares the same transformation properties of polar and axial vectors.

- T_d group (character table 5.5 on page 110).
  The group has four C_3 axes with $\phi = 2\pi/3$ and three C_2 axes with $\phi = \pi$. The 3×3 representation cannot be further reduced. Improper rotations are represented by mirror reflections and roto-reflections.
  The three components of polar vectors are a basis of the three-dimensional T_2 representation, the three components of axial vectors are a basis of the three-dimensional T_1 representation. Again, the characters of proper rotations are the same for polar and axial vectors, the characters of improper rotations are opposite.

Groups with inversion symmetry

Let us consider again the group of full cubic symmetry $O_h = O \otimes C_i$ (character table 5.11 on page 116).
Polar vectors change sign under inversion; their components are a basis of the T_u three-dimensional representation (lower half of the table). The characters of the improper rotations (lower right quadrant of the table) are opposite to the characters of the proper rotations (lower left quadrant of the table).
Axial vectors are invariant with respect to inversion; their components are a basis of the T_g three-dimensional representation (upper half of the table). The characters of the improper rotations (upper right quadrant of the table) are equal to the characters of the proper rotations (lower left quadrant of the table).

5.10.3 Selection rules

Let us come back to the original problem of evaluating whether the integral appearing in (5.131) can be different from zero.
The matrix element in (5.131) is the integral of the product of three functions $\Psi_f^{(\alpha)*}$, H_{int} and $\Psi_i^{(\beta)}$, each one belonging to a given irreducible representation of the symmetry group of the system under consideration.
Since H_{int} is a vector operator (electric or magnetic dipole), we have to focus on the transformation properties of its cartesian components. As we have seen, for high symmetry groups (such as T_d of O_h), symmetry transformations mix the three components, leading to irreducible three-dimensional representations T. For low-symmetry groups, (such as C_{3v} or D_3), the component along the higher-symmetry axis transforms independently with respect to the other two components, leading to a decomposition of the three-dimensional representation into A and E.

Necessary (although not sufficient) condition for the integral in (5.131) not to be zero is that the DP of the three I.R. based on the two functions and the operator contains the irreducible representation A_1. To check this condition, one can again make use of (5.125), extended now to three factors, to calculate the characters of the DP, and of (5.36) to evaluate whether A_1 is included in the DP.

This procedure leads to selection rules for electric dipole and magnetic dipole transitions.
A few examples can be of help.

Example 1: C_{3v} group

The characters of the C_{3v} group are reproduced in Table 5.13.
If the electric field is polarised along the z, it has no components in the x, y directions; the electric dipole operator transforms according to A_1. If the electric field is polarised in the x, y plane, the electric dipole operator transforms according to E. Same considerations for the magnetic dipole operator.

Table 5.13: Character table of the symmetry group C_{3v} ($n = 6$); the last two lines are two possible direct products.

	e	$2C_3$	$3\sigma_v$	
A_1	1	1	1	z
A_2	1	1	-1	R_z
E	2	-1	0	$(x, y), (R_x, R_y)$
$A_1 \times A_1 \times E$	2	-1	0	
$A_1 \times E \times E$	4	1	0	

As an example, the last two lines of Table 5.13 represent two possible direct products (DP) of irreducible representations, the first one corresponding to electric field polarised along the z direction, the second one corresponding to electric field polarised in the x, y plane:

1. The first DP doesn't contain A_1. Transitions from a basis state of the A_1 representation to a basis state of the E representation, induced by an electric field polarised in the z direction, are forbidden by symmetry in the C_{3v} group.
2. The second DP contains A_1. Transitions from a basis state of the A_1 representation to a basis state of the E representation, induced by an electric field polarised in the x, y plane, are not forbidden by symmetry in the C_{3v} group.

Example 2: T_d group

The characters of the T_d group are reproduced in Table 5.14.
The symmetry properties of the electric or magnetic dipole operators are independent of the polarisation direction. The vector components are basis of an irreducible three-dimensional representation (T_1 or T_2).

Table 5.14: Character table of the symmetry group T_d ($n = 24$). The last two lines are two possible direct products.

T_d	e	$8C_3$	$3C_2$	6σ	$6S_4$		orbitals
A_1	1	1	1	1	1		s
A_2	1	1	1	-1	-1		
E	2	-1	2	0	0		$(d_{z^2}, d_{x^2-y^2})$
T_1	3	0	-1	-1	1	(R_x, R_y, R_z)	
T_2	3	0	-1	1	-1	(x, y, z)	$(p_x, p_y, p_z)(d_{xy}, d_{xz}, d_{yz})$
$A_1 \times T_2 \times E$	6	0	-2	0	0		
$A_1 \times T_2 \times T_2$	9	0	1	1	1		

As an example, the last two lines of Table 5.14 represent two possible direct products (DP) of the irreducible representations corresponding to electric dipole transitions:

1. The first DP doesn't contain A_1. Transitions from a basis state of the A_1 representation to a basis state of the E representation, induced by an electric field, are forbidden by symmetry in the T_d group.
2. The second DP contains A_1. Transitions from a basis state of the A_1 representation to a basis state of the T_2 representation, induced by an electric field polarised in the x, y plane, are not forbidden by symmetry in the T_d group.

Example 3: O_h group

The group O_h is characterised by inversion symmetry. The character table 5.11 on page 116 can be schematized as in Table 5.15

Table 5.15: Schematic character table of the group with inversion symmetry: g are the elements of the corresponding group of proper rotations, e is the identity and i is the inversion transformation. Γ_g and Γ_u are bases of even and odd functions, respectively.

	eg	ig
Γ_g	χ	χ
Γ_u	χ	-χ

From the character table 5.7 on page 114, one can see that the components of polar and axial vectors are both basis of the T_1 representation in the proper rotations group O.
The transformation properties are however different under inversion transformations, so that the components of polar and axial vectors are basis of the different representations T_{1u} and T_{1g}, respectively, in the O_h group.
As a consequence, we can derive a general selection rule involving the parity of the basis functions. First we note that $\Gamma_g \times \Gamma_g = \Gamma_g$, $\Gamma_u \times \Gamma_u = \Gamma_g$ and $\Gamma_g \times \Gamma_u = \Gamma_u$.
The transition matrix elements of the electric dipole operator (belonging to Γ_u) are necessarily zero between two states of the same parity, since the direct product Γ_u doesn't contain A_1.
By converse, the transition matrix elements of the magnetic dipole operator can be non zero between states of the same parity.

Note: The above considerations can be applied to a free atom too. As already observed, the parity of a one electron wavefunction is $(-1)^\ell$, where ℓ is here the one-electron azimuthal quantum number. As a consequence of the symmetry properties of the electric dipole operator, electric dipole transitions are forbidden between states of the same parity, e.g. *even* $\ell \rightarrow$ *even* ℓ or *odd* $\ell \rightarrow$ *odd* ℓ.

5.11 Complements

5.11.1 Inversion of matrices

The inversion of a matrix $\mathbf{D}$ is performed according to the rule of equation (5.6), say

$$\mathbf{D}^{-1} = \frac{\mathrm{Adj}(\mathbf{D})}{|\,\mathbf{D}\,|} \tag{5.136}$$

where Adj means adjoint matrix and $|\,\mathbf{D}\,|$ is the determinant.
The adjoint of a matrix $\mathbf{D}$, say Adj($\mathbf{D}$), is the transposed of the matrix of the cofactors D_{ij} of each element d_{ij} of $\mathbf{D}$.
The meaning of (5.136) can be better grasped by some simple examples.

Example: two-dimensional case

For a two-dimensional matrix, the cofactor of an element d_{ij} is a single element, which shares with d_{ij} the diagonal, and has to be chosen with the proper sign.

$$\mathbf{D} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \Rightarrow \qquad \mathrm{Adj}(\mathbf{D}) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{5.137}$$

One can easily verify that

$$\mathbf{D}\mathbf{D}^{-1} = \mathbf{D}\frac{\mathrm{Adj}(\mathbf{D})}{|\,\mathbf{D}\,|} = \frac{\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}{ad-bc} = \frac{\begin{pmatrix} ad-bc & -ab+ba \\ cd-dc & -bc+ad \end{pmatrix}}{ad-bc} = \mathbf{I} \tag{5.138}$$

Example: three-dimensional case

For a three-dimensional case,

$$\begin{pmatrix} d_{xx} & d_{xy} & d_{xz} \\ d_{yx} & d_{yy} & d_{yz} \\ d_{zx} & d_{zy} & d_{zz} \end{pmatrix} \tag{5.139}$$

the cofactor D_{ij} of an element d_{ij} is the determinant of the two-dimensional matrix obtained by eliminating the row i and the column j, taken with the proper sign. For example:

$$D_{xx} = d_{yy}d_{zz} - d_{yz}d_{zy}\,, \qquad D_{xy} = -d_{yx}d_{zz} + d_{yz}d_{zx} \tag{5.140}$$

The determinant of the matrix is

$$|\mathbf{D}| = d_{xx}d_{yy}d_{zz} - d_{xx}d_{yz}d_{zy} - d_{xy}d_{yx}d_{zz} + d_{xy}d_{yz}d_{xx} + d_{xz}d_{yx}d_{zy} - d_{xz}d_{yy}d_{zx} \tag{5.141}$$

By a lengthy calculation, one can evaluate $\mathrm{Adj}(\mathbf{D})$ and verify that $\mathbf{D}\mathbf{D}^{-1} = \mathbf{I}$.

5.12 Bibliography of Chapter 5

- P. Atkins and J. De Paula: *Physical Chemistry*, 9th ed., W.H. Freeman 2009.
  Chapter 11 (molecular symmetry).
- P. Atkins: *Chimica Fisica* [in Italian], Zanichelli 1981.
  Cap. 16 (la simmetria : rappresentazione e conseguenze).
- F. W. Byron and R. W. Fuller: *Mathematics of classical and quantum Physics*, Dover 1992.
- D. C. Harris and M. D. Bertolucci: *Symmetry and Spectroscopy*, Oxford University Press 1978, reprinted by Dover Publications 2014.
- W. A. Harrison: *Solid State Theory*, Dover 1980.
  Part I, Chapters 4 and 5 (group theory and its applications)
- L. Landau and E. Lifshitz: *Theoretical Physics* (various editions in different languages).
  Vol. 3: *Quantum Mechanics*, Chapter 12 (point symmetry of molecules).
  Vol. 5: *Statistical Physics*, Chapter 13 (symmetry of crystals).
- A. Messiah: *Quantum Mechanics*, North Holland 1970.
  Cap. XV (Invariance and conservation theorems. Time reversal.)
- J. C. Slater: *Quantum theory of matter*, McGraw-Hill 1978.
  Chapter 25 (group theory and wave functions symmetry).
  [Italian edition: Zanichelli 1980]
- M. Tinkham: *Group theory and quantum mechanics*, Mc Graw Hill 1964, Dover 1992.
  Chapter 3 (theory of group representations).
- B.K. Vainshstein: *Modern Crystallography*, Springer 1981.
  Vol. 1, Chapter 2 (fundamentals of the theory of symmetry).

Chapter 6

The origin of structures

In Chapter 2 we introduced the basic ideas about solid aggregates of atoms – perfect crystals, real crystals, non-crystalline solids. The symmetry considerations of Chapter 4 allowed a rigorous classification of crystal structures .
In this Chapter 6 we try to understand the origin of the possible different structures in terms of physical interactions. At 1 bar pressure and at sufficiently low temperatures, all substances solidify (with the exception of helium). Why different substances assume different crystalline structures, with different degrees of packing and different physical properties? and why some substances easily solidify in non-crystalline form?
Different approaches are possible, of different degrees of sophistication, such as ab-initio calculations, chemical bond models, molecular dynamics simulations. Starting point for all the different approaches is the adiabatic approximation (Born-Oppenheimer, 1927), which is reviewed in Section 6.1
Some general properties of solids can be obtained by only considering pair interactions between neighbouring atoms (Section 6.2); this approach leads to a popular classification of structures in terms of chemical bonds: molecular crystals (Section 6.3), ionic crystals (Section 6.4), covalent molecules and solids (Section 6.5), metals (Section 6.6), hydrogen-bonded structures (Section 6.7).
The chemical bond models cannot account for all physical properties of solids; a thorough understanding the electronic structure and of the conduction properties of metals and insulators can be obtained only by a more refined approach, which is postponed to Chapter 9.
Section 6.8 is dedicated to an introduction to non-crystalline solids: basic structural properties and production techniques.
Ground state properties of electronic systems are nowadays frequently calculated within the framework of the Density Functional Theory (DFT). A short introduction to DFT is given in Section 6.9.

6.1 Born-Oppenheimer (adiabatic) approximation

The stationary Schrödinger equation for a many-atomic system, such as a molecule or a crystal, is

$$H\,\Psi_{\rm tot} \;=\; E\,\Psi_{\rm tot}\,, \tag{6.1}$$

where the total wavefunction $\Psi_{\rm tot}$ depends on the coordinates of the n electrons and of the N nuclei:

$$\Psi_{\rm tot} \;=\; \Psi_{\rm tot}(\vec{r}_1,\ldots,\vec{r}_n\,;\vec{R}_1\,\ldots,\vec{R}_N) \;=\; \Psi_{\rm tot}(r,R)\,. \tag{6.2}$$

In the last term of (6.2), r and R are short-hand notations for the full sets of electrons and nuclei coordinates, respectively. The wavefunction is defined in a $3N+3n$-fold configuration space.
Solving eq. (6.1) for a many-atomic system (or simply for a many-electron atom) is a typical many-body problem, which can be faced only by approximate methods.

In general, instead of considering separately all the electrons, it is convenient to distinguish the ions from the valence electrons:

a) ions = nuclei + core electrons (electrons which don't contribute to inter-atomic interaction),
b) valence electrons = external electrons, which contribute to inter-atomic interaction.

In such cases, R and N refer to ions, and r and n refer to valence electrons in (6.2).

Note: The distribution of valence electrons is peculiar of the different crystal types. It can be experimentally detected by refined analyses of high-accuracy X-ray diffraction spectra.

6.1.1 Total Hamiltonian

The total Hamiltonian (neglecting relativistic terms) is

$$H = T_n + V_{nn} + T_e + V_{ne} + V_{ee} \,. \tag{6.3}$$

The total kinetic energy operator of the ions, labelled by A, is

$$T_n = -\sum_A \frac{\hbar^2}{2M_A} \nabla_A^2 = -\sum_A \frac{\hbar^2}{2M_A} \left(\frac{\partial^2}{\partial x_A^2} + \frac{\partial^2}{\partial y_A^2} + \frac{\partial^2}{\partial z_A^2} \right) . \tag{6.4}$$

The total kinetic energy operator of valence electrons (labelled by i) is

$$T_e = -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 \,. \tag{6.5}$$

The total potential energy of the interaction between ions is

$$V_{nn} = \frac{1}{2} \sum_A \sum_{B \neq A} \frac{Z_A Z_B}{|\vec{R}_A - \vec{R}_B|} \,, \tag{6.6}$$

where Z_A and Z_B are the ionic charges.
The total potential energy for the interaction between valence electrons is

$$V_{ee} = \frac{1}{2} \sum_i \sum_{j \neq i} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \,. \tag{6.7}$$

The total potential energy of electrostatic interaction between valence electrons and ions is

$$V_{ne} = -\sum_{iA} \frac{e\, Z_A}{|\vec{r}_i - \vec{R}_A|} \,. \tag{6.8}$$

6.1.2 Separation of nuclear and electronic motion

In principle, the time-independent Schrödinger equation for the full system contains all the information on its stationary states. Its solution is however a prohibitive task.
The Born-Oppenheimer approximation consists in de-coupling the Schrödinger equation for electrons from the Schrödinger equation for nuclei.
The physical basis of such an approximation is the fact that the mass of nuclei is from about 2×10^3 to about 10^5 times larger than the electron mass. As a consequence, if the linear momenta of electrons and nuclei are of the same order of magnitude, the motion of electrons is much faster than the motion of nuclei, so that electrons can be considered as instantaneously following the nuclei.

Factorisation of the total wave-function

The Born-Oppenheimer approximation assumes that the total wave-function can be factorised:

$$\Psi_{\text{tot}} = \Psi_e(r\,;R) \;\; \Phi_n(R) \,, \tag{6.9}$$

where

a) the nuclear wave-function $\Phi_n(R)$ only depends on the nuclear coordinates R,
b) the electron wave-function $\Psi_e(r\,;R)$ depends directly on the electron coordinates r and parametrically on the nuclear coordinates R.

The electron wave-function can be a Slater determinant of spin-orbitals and is normalised within the electrons configuration space,

$$\int \Psi_e^*\,\Psi_e\,d\tau \;=\; 1\,. \tag{6.10}$$

6.1.3 Schrödinger equation for electrons

Let us consider the nuclei at fixed positions R. The Schrödinger equation for the electrons is

$$H_e\,\Psi_e(r\,;R) \;=\; [T_e + V_{ne} + V_{ee} + V_{nn}]\;\Psi_e(r\,;R) \;=\; E_e\,\Psi_e(r\,;R)\,. \tag{6.11}$$

Note that the Hamiltonian H_e in (6.13) includes the $V_{nn}(R)$ potential energy term of the nuclei.

By solving the electronic Schrödinger equation (6.13) for different values of R one can obtain the eigenvalues $E_e^{(\alpha)}(R)$ which characterise the stationary states of the electron system (α labels the ground state gs and the excited states).
The functions $E_e^{(gs)}(R)$ represent potential energy surface (PES) in the R space. The minimum of the ground-state function $E_e^{(gs)}(R)$ correspond to the equilibrium configuration R of the atomic aggregate. In the following, when referring to the ground state, the apex gs will be omitted.
The eigenvalues $E_e(R)$ can be obtained as

$$E_e \;=\; \frac{\int \Psi_e^*\,H_e\,\Psi_e\,d\tau}{\int \Psi_e^*\,\Psi_e\,d\tau} \;=\; E_e(R) \tag{6.12}$$

Example: the hydrogen molecular ion H_2^+

The simple case of the hydrogen molecular ion H_2^+, which is exactly solvable, can be of help for a better understanding.
The nuclear configuration is described by one parameter, the internuclear distance R. The V_{nn} potential energy corresponds to the Coulomb repulsion between the two nuclei, and increases when R decreases. There is no V_{ee} potential energy contribution, since there is only one electron. The electronic Schrödinger equation is

$$\left[\frac{\hbar^2}{2m}\nabla^2 + \frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{R} - \frac{1}{r_1} - \frac{1}{r_2}\right)\right]\;\Psi_e(r\,;R) \;=\; E_e(R)\,\Psi_e(r\,;R)\,. \tag{6.13}$$

The solution of the equation can be found in textbooks on Atomic Physics. Let here quote the main results concerning the lowest energy levels (Fig. 6.1).
It is convenient first to consider the values $E_e - V_{nn}$, say to exclude the nuclear repulsion potential energy. For large inter-nuclear distances R, the electron is bound to one of the two nuclei, and the value $E_e - V_{nn}$ corresponds to the binding energy of the H $1s$ level, -13.6 eV. For $R = 0$, the system corresponds to the He_2^+ ion, whose lowest energy levels are -54 eV (He^+ $1s$) and -13 eV (He^+ $2s$). In the first case, the purely electronic energy $E_e - V_{nn}$ is always decreasing when the distance R decreases, corresponding to an attractive force between the nuclei.
Summing up the contributions V_{nn}, one one obtains the values $E_e(R)$ of the two lowest molecular levels σ_g (bonding) and σ_u^* (antibonding) as a function of R; the first function $E_e(R)$ has a minimum position, corresponding to an equilibrium configuration of the nuclear system, the second one is always decreasing.

Note: This example shows that it can be sometimes convenient to first solve the Schrödinger equation (6.13) omitting the V_{nn} term, and to add it subsequently.

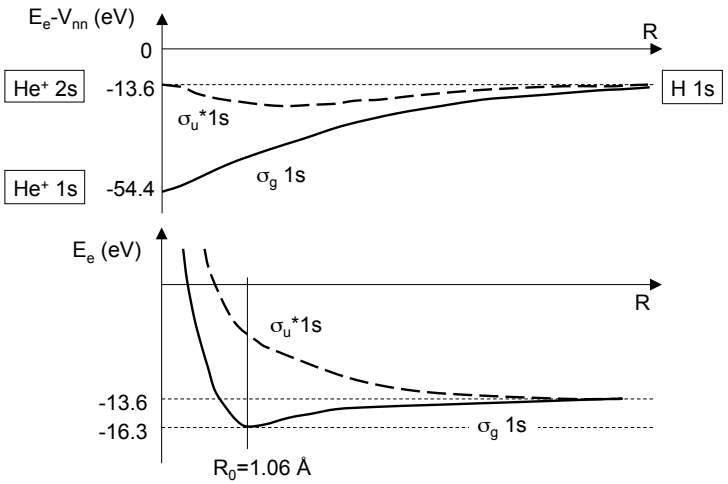

Figure 6.1: Purely electronic energy $E_e - V_{nn}$ (top) and total E_e energy (bottom) for the hydrogen molecular ion H_2^+.

6.1.4 Schrödinger equation for ions

Let us consider again the total Schrödinger equation

$$H\,\Psi_{\rm tot} \;=\; E\,\Psi_{\rm tot}\,, \tag{6.14}$$

where now the Hamiltonian includes the nuclear kinetic term, $H = H_e + T_n$, and the total wave-function is $\Psi_{\rm tot} = \Psi_e(r\,;R)\Phi_n(R)$.
The Schrödinger equation can thus be rewritten as

$$H_e\,(\Psi_e\Phi_n) + T_n\,(\Psi_e\Phi_n) \;=\; E\,(\Psi_e\Phi_n)\,, \tag{6.15}$$

say, taking into account the results obtained for the electronic Schrödinger equation,

$$E_e(R)\,\Psi_e\,\Phi_n + T_n\,(\Psi_e\Phi_n) \;=\; E\,\Psi_e\,\Phi_n\,. \tag{6.16}$$

In the first term on the left and in the term on the right, the electron wave-function Ψ_e can already be integrated over the electron configuration space.
Let us consider in more detail the T_n term:

$$\begin{aligned} T_n\,(\Psi_e\Phi_n) \;&=\; -\sum\nolimits_A \frac{\hbar^2}{2M_A}\,\left(\vec{\nabla}_A\cdot\vec{\nabla}_A\right)\,(\Psi_e\Phi_n) \\ &=\; -\sum_A \frac{\hbar^2}{2M_A}\,\left[\Psi_e\left(\nabla_A^2\Phi_n\right) + 2\left(\vec{\nabla}_A\Psi_e\right)\cdot\left(\vec{\nabla}_A\Phi_n\right) + \left(\nabla_A^2\Psi_e\right)\Phi_n\right]\,. \end{aligned} \tag{6.17}$$

The Born-Oppenheimer approximation consists in considering negligible the second and third terms within the square parentheses, where the nuclear kinetic operators act on the electron wavefunction. Within this approximation, we only consider the first term, where the electronic function Ψ_e is factorised.
When integrating the resulting Schrödinger equation over the electrons configuration space, the electron wave-function gives rise to the normalization value 1, while the nuclear wavefunction is unaffected. We finally obtain the Schrödinger equation for the nuclei in the Born-Oppenheimer approximation,

$$\left[-\sum_A \frac{\hbar^2}{2M_A}\nabla_A^2 + E_e(R)\right]\,\Phi_n \;=\; E\,\Phi_n \tag{6.18}$$

or

$$\left[-\sum_A \frac{\hbar^2}{2M_A}\nabla_A^2 + V(R)\right]\Phi_n \;=\; E\,\Phi_n\,, \tag{6.19}$$

where $V(R) = E_e(R)$.

6.1.5 The potential energy surface (PES)

The function $V(R) = E_e(R)$ appearing in (6.18) and (6.19) is the ground-state energy of a system of interacting electrons moving in the field of fixed ("clamped") nuclei. It is referred to as clamped ions energy or potential energy surface (PES) or Born-Oppenheimer energy surface.
The ground state properties of electrons are important to determine:

a) the equilibrium structure and the phase stability of the ion system (at $T = 0$), which correspond to the absolute minimum of the Born-Oppenheimer energy surface with respect to the set of nuclear coordinates at different temperatures and pressures. Partial minima correspond to metastable equilibrium ground states.

b) the finite temperature dynamical properties of ions (atomic vibrations), which depend on the behaviour of the Born-Oppenheimer energy surface in the vicinity of the equilibrium configuration, say on its derivatives of second and higher order evaluated at the equilibrium configuration.

In this Chapter, we are interested in the ground state equilibrium properties of ions. The dynamical properties of ions will be considered in Chapter 7.

The solution of the electronic Schrödinger equation gives also the excited states of the electronic system, which are necessary to explain the electronic spectroscopies and the transport phenomena.

PES calculation

The PES can in principle be obtained by solving the electronic Schrödinger equation (6.13) for a convenient number of ionic configurations R. An exact solution of the Schrödinger equation is out of question but for the simplest systems (harmonic oscillator, single particle in a central field, squared-walled box). Numerical solutions are unmanageable even for relatively small molecules, in view of the large number of variables involved to get a sufficiently accurate picture of the wave-function in the configuration space.
The problem is generally solved according to the following guidelines.

1. The problem is simplified by taking into account the symmetry of the system. Symmetry considerations don't introduce approximations.
2. The approximation is made of treating each electron as moving independently in an average field from the others. The ground state energy of the system is obtained self-consistently by variational methods. Typical examples are based on the Hartree-Foch-Slater method or the more recent Density Functional Theory (DFT).

In many cases, the equilibrium configuration is known from experiments, e.g. x-rays or neutron diffraction. The ability of reproducing the equilibrium configuration is a test of the theory, which can then give further information on, for example, the vibrational dynamics or the mechanisms of phase transitions.

6.1.6 Adiabatic approximation

The Born-Oppenheimer approximation is sometimes named adiabatic approximation. The meaning for that can be better grasped if one assumes that the massive ions can be treated as slow-moving classical particles. Let us suppose that, for a given set R of ion coordinates, the electrons system is in a given eigenstate $\Psi_e(r\,;R)$ of the Hamiltonian $H_e(R)$. When the values of the ion coordinates

are modified, the Hamiltonian $H_e(R)$ and the wave-function $\Psi_e(r\,;R)$ are in turn modified. Within the adiabatic approximation, the electron wavefunction $\Psi_e(r\,;R)$ evolves remaining at any time an eigenstate of $H_e(R)$.
The adiabatic approximation is valid if the nuclear motion is sufficiently slow with respect to the relaxation time of the electron cloud.

Note: The term "adiabatic" is generally used in Thermodynamics to describe processes without exchange of heat between system and environment. The quantum mechanical adiabaticity considered here is closer to the thermodynamical concept of a quasi-static process, and has no relation with heat exchange.

Non-adiabatic effects

Correction to the Born-Oppenheimer approximation arise because the electronic wave-functions depend parametrically on the ion coordinates. These corrections are accounted for by the second and third terms in the last line of (6.17):

a) $\vec{\nabla}_A \Psi_e$, the "non-adiabatic coupling", can be non-negligible
b) $\nabla^2_A \Psi_e$, the diagonal correction, is usually negligible

The magnitude of non-adiabatic corrections thus depends on the gradient of the electronic wave-functions when the nuclear configuration changes. When the adiabatic approximation is not valid, the electron wave-function is a linear combination of eigenstates of the evolving Hamiltonian $H_e(R)$. If the motion of ions is sufficiently fast (say if the energy of ionic excitations is sufficiently high) and if electron levels are sufficiently close, the non adiabatic coupling can cause the transition of the electron system from the ground state to an excited state.

6.2 Atomic aggregates

In this Section 6.2, some general considerations are presented concerning the atomic aggregates.

6.2.1 Pair interaction

The structure of atomic aggregates, be it crystalline or non-crystalline, results from the collective interaction of the atoms, as depicted by the Schrödinger equation (6.1). In some cases, however, the collective interaction can be approximately considered as the sum of single pair interactions.

Let us consider a two-atomic system. The Born-Oppenheimer surface $V(R)$ collapses to a potential energy curve $V(r)$ defined in a one-dimensional space, where now r is the inter-nuclear distance. The typical shape of the $V(r)$ function is depicted in Fig. 6.2 (left).
The minimum of the function $V(r)$ (equilibrium position) is determined by the balance between

a) the attractive part, whose strength can be very different for different atomic species,
b) the repulsive part, due to the Coulomb repulsion of the positive ions and to the Pauli exclusion principle which opposes to the superposition of electron clouds.

The simple two-atomic model can give important insights on the origin of many crystalline and non crystalline structures, provided we can consider the structure as generated by the simple sum of individual pair interactions.

Atomic vibrations

An atomic system cannot be at rest in the position corresponding to the the minimum of the $V(r)$ function (or at the minimum of the Born-Oppenheimer surface) due to the momentum-position uncertainty principle.
As effect of the uncertainty principle, the actual ground state energy is larger than the minimum of $V(r)$, the difference corresponding to the zero-point energy of harmonic oscillations.

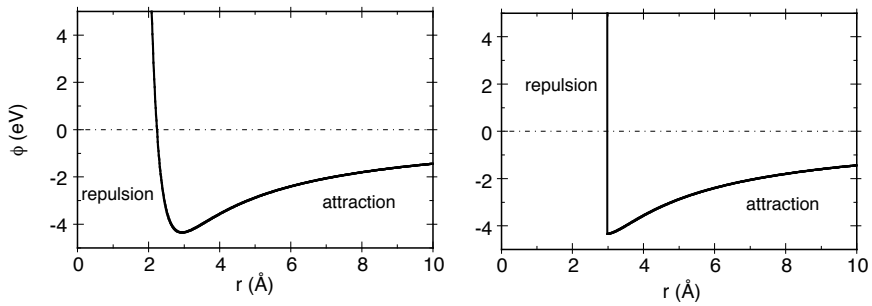

Figure 6.2: Left: typical behaviour of the interaction potential energy. Right: the potential energy of the hard-sphere model with smooth attraction.

A regular crystalline structure describes the average position of atoms. The actual instantaneous structure is missing translational symmetry event at $T = 0$.
A striking examples of the zero-point effect is represented by helium, which cannot solidify at at any temperature at the standard pressure of 1 bar.

6.2.2 Hard-sphere model and closest packing

As a simple example of the two-atoms interaction model, let us consider an ideal system composed of hard spheres, with an isotropic attractive force smoothly depending on r and a step-like repulsive force (Fig. 6.2, right).

Two-dimensional case with translational symmetry

In two dimensions, the resulting structure is a regular closest packing with hexagonal symmetry (Fig. 6.3, top left panel). Axes C_6 perpendicular to the plane pass through the centres of the spheres, axes C_3 pass through the centres of the voids regions.

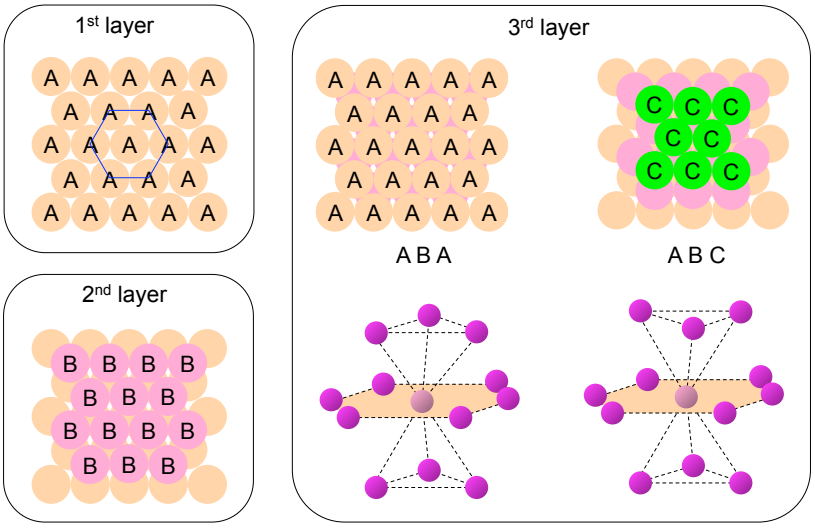

Figure 6.3: Closest packing of hard spheres in crystals. Top left: two-dimensional planar packing. Bottom left: second layer. Right: two possibilities for the third layer (top) and the corresponding 3-d arrangement of atoms.

Three-dimensional case with translational symmetry

In three dimensions one can again obtain a closest-packing of spheres.
A second layer ...BBBB... of spheres can be put on top of the first layer ...AAAA... in correspondence of half of the triangular void regions (Fig. 6.3, bottom left panel).
As for the third layer, two possibility exist, giving rise to two different structures (Fig. 6.3, right panel):

1. ABABAB stacking of layers, where the third-layer spheres are on the vertical of the first-layer spheres. This stacking corresponds to the hcp (hexagonal close packed) structure, the layer planes being perpendicular to the c axis.
2. ABCABC stacking of layers, where the third-layer spheres are on the vertical of voids of both the first and second layer. This stacking corresponding to the fcc (face centred cubic) structure, the layer planes being perpendicular to the [111] direction.

The closest packing corresponds to a coordination number z=12.
The hard spheres model shows how a closest-packed regular crystalline structure can be generated, but has no general validity:

a. The closest-packed stacking of layers can be affected by lack of regularity along the direction perpendicular to the layers, for example ABCABABCABC or ABABABCABAB. Such stacking faults are quite frequent in real closest-packed crystals.
b. Non closest-packed structures are frequently encountered, due to the lack of isotropy of the interatomic interaction (e.g. covalent bonds) or to other steric constraints (e.g. ionic bonds). Such problem will be considered in next Sections.

Three-dimensional case with no translational symmetry

Closest packing requires that each atom is surrounded by z=12 atoms at the same distance.
If a cluster of 13 atoms is considered (one central atom and 12 peripheral atoms), there are three different ways of obtaining closest-packing:

1. The 12 atoms are at the corners of a cuboctrahedron (Fig. 6.4, left). The numbers of atoms in the horizontal planes is 3–6–3.
   This arrangement is found in the fcc structure considered above.
2. The 12 atoms are at the corners of a twinned cuboctrahedron (obtained from the cuboctahedron by rotatong the upper half of 60° with respect to the bottom half). The number of atoms in the horizontal planes is 3–6–3.
   This arrangement is found for in hcp structure considered above.
3. The 12 atoms are at the corners of an icosahedron (Fig. 6.4, right). The numbers of atoms in the horizontal planes is 1–5–5–1.
   This arrangement is incompatible with long-range translational order, due to the presence of pentagonal symmetry.

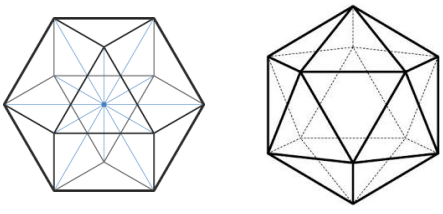

Figure 6.4: Cuboctahedron (left) and icosahedron (right).

The icosahedral structure is preferred by a cluster of only 13 atoms, in view of the lower surface energy. In fact, each surface atom of the icosahedron is coordinated to 5 other atoms, while in the cuboctahedron it is coordinated to only 4 other atoms.
Icosahedral ordering has been originally proposed to explain the phenomenon of supercooling in liquid metals (Frank, 1952). Recently, experimental evidence of icosahedral ordering has been found in some liquid metals.
Icosahedral ordering is encountered in solid nano-clusters, even for a number of atoms larger than 13. When the nano-cluster size increases, other arrangements, such as the cuboctahedron, are preferred.

6.2.3 Thermodynamic and statistical considerations

A general approach to the formation of crystal structures can be based on thermodynamic considerations.
From a thermodynamic point of view, the stable equilibrium configuration (at constant volume) is characterised by

1. the minimum of the Gibbs free energy $= U + pV - TS$ for system at fixed p and T;
2. the minimum of the Helmholtz free energy $F = U - TS$ for systems at fixed V and T.

At $T = 0$, the corresponding condition for stable equilibrium is the minimum of the enthalpy $H = U + pV$ or of the internal energy U (which includes the zero point energy).

The large variety of possible inter-atomic interactions gives rise to different equilibrium structures for different systems. Thermodynamic considerations lead to the conclusion that equilibrium structures at $T = 0$ should be crystalline. Actually, the minimum of energy of the system as a whole corresponds to the minimum of energy of any of its constituent parts. The possibility that any part has the same equilibrium structure is guaranteed only by the existence of long-range crystalline order.

Non-crystalline systems can be found at very low temperatures too, corresponding however to non-equilibrium states (glasses, liquid crystals, etc).

Macroscopic hints on microscopic structure

The difference of some macroscopic quantities of different crystals is an important clue about the binding properties at the atomic level. As an example, in Table 6.1 three parameters are compared for some selected crystals:

1. The cohesive energy, say the energy that is necessary to decompose the crystal into neutral free atoms at rest at infinite distance.
2. The melting point, say the temperature of the solid to liquid transition at the pressure of 1 bar.
3. The isothermal compressibility $\chi_T = -(1/V)\ (\partial V/\partial p)_T = 1/B$ (where V is the volume and B is the bulk modulus).

6.2.4 Classifications of chemical bonds

The origin of the different structures can be to a good extent explained in terms of the inter-atomic interactions, which are in turn based on the spatial arrangement of electrons in the real space. This approach leads to a quite popular classification of bonds:

1. Van der Waals bond
2. Ionic bond
3. Covalent bond
4. Hydrogen bond
5. Metallic bond

Table 6.1: Some macroscopic quantities for selected crystals.

	Cohesive energy (eV/atom)	Melting point (K)	Compressibility (10^{-11} m^2/N)
Ge	3.85	1211.	1.29
Cu	3.49	1358.	0.73
NaCl	6.37	1074.	4.16
Ne	0.02	24.56	100.
Kr	0.116	115.8	56.

This classifications has some advantages :

- consistency with the electronic densities revealed by X-ray diffraction experiments
- easy connection with the theories of molecular structure
- phenomenological understanding of the origin of different crystal structures (close-packed, NaCl, diamond, etc.)
- intuitive understanding of the relative easiness of some structures to solidify in non-crystalline form
- intuitive understanding of some macroscopic properties, e.g. why molecular, ionic, covalent and H-bonded crystals are electric insulators at $T = 0$ K while metals are good electric conductors
- usefulness for lattice-dynamical models

At least two limits of the classifications should be evidenced:

- there is not always a clear-cut separation between the different types of bond
- a full understanding of the electron conduction properties and of the distinction between conductors and insulators can be obtained only by a reciprocal space approach

6.3 Van der Waals bonds and molecular crystals

Molecular crystals are composed of neutral molecules with complete electron shells. They are held together by the weak dipole-dipole Van der Waals forces (after the name of the dutch physicist who first modelled the inter-molecular attractive forces in real gases, in 1873).

Examples of molecular crystals:

1. Solids of noble gases (Ne, Ar, Kr, Xe), where the molecules are monatomic. Atoms have no permanent electric dipoles: the attractive forces are due to fluctuating dipoles (Van der Waals - London forces). Helium cannot be solidified at ambient pressure.
2. Crystals of the H_2, O_2, N_2, Cl_2, CH_4 molecules.
3. Crystals of biological molecules (artificially produced in laboratory for protein crystallography experiments)

In non-monatomic molecular crystals, such as O_2 or N_2, one can distinguish

a) strong intra-molecular bonds, which give rise to high-frequency intra-molecular vibrations;
b) weak inter-molecular bonds, which give rise to low-frequency inter-molecular vibrations.

Molecular crystals are characterised by low melting points.

The electric dipole-dipole Van der Waals forces can be further classified as

1. forces between two permanent dipoles (Keesom forces)
2. forces between a permanent dipole and an induced dipole (Debye forces)

3. forces between two induced dipoles (London forces)

In the following we consider the London forces responsible for the crystallisation of noble gases. Note however that Van der Waals forces are always present, although their effect can be negligible with respect to other types of bond (ionic, covalent, metallic).

6.3.1 London attractive potential

To grasp the physical origin of the dipole-induced attractive potential and its dependence on distance, let us consider a simple one-dimensional model representing a couple of atoms [taken from Kittel].

Atomic oscillator model

Let us consider a nucleus of charge Q and its electron cloud, supposed to be spherical with radius r and homogeneous. A displacement x of the C.M. of the electron cloud from the nucleus gives rise to a restoring force

$$|F| \;=\; \frac{1}{4\pi\epsilon_0}\,\frac{Qq_x}{x^2} \;=\; \frac{1}{4\pi\epsilon_0}\,\frac{Q^2x^3}{x^2r^3} \;=\; kx\,, \tag{6.20}$$

where $q_x = Q\,[4\pi x^3/3]/[4\pi r^3/3]$ is the negative charge within the sphere of radius r and $k = (1/4\pi\epsilon_0)(Q^2/r^3)$ (which is strictly valid for a homogeneous distribution of charge). The restoring force is thus harmonic and corresponds to a potential energy $V = kx^2/2$. A homogeneous distribution of charge is quite a crude approximation, but the oscillator model is a good approximation, although the force constant k should be connected to the charge distribution in a more complicated way.

From the classical point of view, random oscillations correspond to a fluctuating electric dipole. From the quantum point of view, the atom can be considered as a quantum harmonic oscillator, characterised by a finite value of zero-point energy; the non-zero minimum of energy supports the classical picture of a fluctuating dipole.

The interaction between two atoms

Let us now study the interaction between two neutral equal atoms due to the Coulomb interaction between their components (electrons and nuclei).

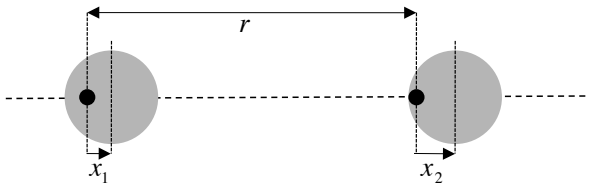

Figure 6.5: Schematic representation of the origin of London attractive potential between two neutral atoms (not to scale, in reality $|x_1|, |x_2| \ll R$).

Let R be the distance between the two nuclei, x_1 and x_2 the displacements of the electron clouds of atom 1 and atom 2 from the corresponding nuclei, respectively (Fig. 6.5). We assume $R \gg r$. If we neglect the interaction between the two atoms, the un-perturbed Hamiltonian is the sum of the independent Hamiltonians of the two atoms (harmonic oscillators)

$$\mathcal{H}_0 \;=\; \frac{p_1^2}{2m} + \frac{1}{2}kx_1^2 \;+\; \frac{p_2^2}{2m} + \frac{1}{2}kx_2^2\,. \tag{6.21}$$

(The force constant k depends on the actual electronic structure of atoms). The proper frequency of the independent harmonic oscillators is $\omega_0 \;=\; \sqrt{k/m}$.

The interaction between the two atoms, due to the Coulomb forces between the charges of different atoms, can described as a perturbation

$$\begin{aligned}\mathcal{H}_I &= \frac{Q^2}{4\pi\epsilon_0}\left[\frac{1}{R}+\frac{1}{R-x_1+x_2}-\frac{1}{R-x_1}-\frac{1}{R+x_2}\right] \\ &= \frac{Q^2}{4\pi\epsilon_0}\frac{1}{R}\left[1+\left(1-\frac{x_1}{R}+\frac{x_2}{R}\right)^{-1}-\left(1-\frac{x_1}{R}\right)^{-1}-\left(1+\frac{x_2}{R}\right)^{-1}\right] \end{aligned} \tag{6.22}$$

The first two terms in the square parentheses correspond to the repulsive contributions ion-ion and electrons-electrons, the last two terms correspond to the attractive contributions ion-electrons.
Since $|x_1|, |x_2| \ll r$, the terms in the parentheses of (6.22) can be expanded according to

$$(1+y)^\alpha = \sum_{n=0}^{\infty}\binom{\alpha}{n}y^n\,, \quad \binom{\alpha}{0}=1\,, \quad \binom{\alpha}{n}=\prod_{k=1}^{n}\frac{\alpha-k+1}{k}$$

where $\alpha=-1$. To second order, $(1+y)^{-1}=1-y+y^2$ and one obtains

$$\mathcal{H}_I \simeq \frac{Q^2}{4\pi\epsilon_0}\frac{1}{R}\left[-\frac{2x_1x_2}{R^2}\right] = -\frac{Q^2}{4\pi\epsilon_0}\frac{2x_1x_2}{R^3} = -2k'x_1x_2\,, \tag{6.23}$$

where $k' \propto 1/R^3$ and $k' \ll k$ (since $R \gg r$). According to (6.23), the attractive contributions prevail over the repulsive ones, and the interaction gives rise to a global attraction.
The total Hamiltonian $\mathcal{H}=\mathcal{H}_0+\mathcal{H}_I$ refers to two coupled harmonic oscillators, of coordinates x_1 and x_2, respectively.
One can diagonalize the total Hamiltonian $\mathcal{H}$ by the normal mode transformation, whereby the old coordinates and momenta (x_1, x_2, p_1, p_2) are expressed in terms of the new ones (x_s, x_a, p_s, p_a):

$$\begin{cases} x_1=(x_s+x_a)/\sqrt{2} \\ x_2=(x_s-x_a)/\sqrt{2} \end{cases} \qquad \begin{cases} p_1=(p_s+p_a)/\sqrt{2} \\ p_2=(p_s-p_a)/\sqrt{2} \end{cases} \tag{6.24}$$

The total Hamiltonian expressed in terms of the normal coordinates is the sum of the contributions of two independent normal modes

$$\mathcal{H} = \left[\frac{p_s^2}{2m}+\frac{1}{2}(k-2k')\,x_s^2\right] + \left[\frac{p_a^2}{2m}+\frac{1}{2}(k+2k')\,x_a^2\right]. \tag{6.25}$$

The frequencies of the two uncoupled oscillators are

$$\omega = \left[\frac{1}{m}(k\pm 2k')\right]^{1/2} = \left[\frac{k}{m}\left(1\pm\frac{2k'}{k}\right)\right]^{1/2} \simeq \omega_0\left[1 \pm \frac{k'}{k} - \frac{1}{8}\left(\frac{2k'}{k}\right)^2\right] \tag{6.26}$$

The zero-point energy of the system of the two atoms is the sum of the zero point energies of the two normal modes

$$\frac{1}{2}\hbar(\omega_s+\omega_a) = \hbar\omega_0 - \hbar\omega_0\frac{1}{8}\left(\frac{2k'}{k}\right)^2 = \hbar\omega_0 - \frac{A}{R^6}\,, \tag{6.27}$$

where the $1/R^6$ dependence is due to the fact that $k' \propto 1/R^3$.
The Coulomb interaction between the two atoms reduces the energy by a term proportional to $-1/R^6$, corresponding to an attractive potential energy.

6.3.2 Repulsive potential

The Coulomb repulsion between ions and between electrons of different atoms is included in the previous perturbative treatment, and is overcome by the attractive contribution.
There is however a further strong contribution to repulsion, due to the Pauli exclusion principle. The repulsive potential energy due to the Pauli principle has a much steeper slope as a function of R than the attractive potential energy.
For mathematical simplicity, one models the repulsive potential energy by a power law $1/R^{12}$, where the exponent 12 is chosen so as to be twice the exponent of the attractive component and facilitate some calculations.

6.3.3 Lennard-Jones potential 6-12

By summing the attractive and the repulsive contributions to the potential energy, one obtains the so-called Lennard-Jones (LJ) potential for a pair of atoms at distance r:

$$\phi_{ij}(r) = -\frac{A}{r^6} + \frac{B}{r^{12}} \tag{6.28}$$

A more compact expression is obtained by substituting $A = 4\epsilon\sigma^6$ and $B = 4\epsilon\sigma^{12}$,

$$\boxed{\phi_{ij}(r) = 4\epsilon\left[-\left(\frac{\sigma}{r}\right)^6 + \left(\frac{\sigma}{r}\right)^{12}\right]} \tag{6.29}$$

At the position of the minimum of the potential energy, say at the equilibrium position r_0, one has

$$\boxed{r_0 \;=\; \sqrt[6]{2}\,\sigma = 1.122\,\sigma}\,, \qquad \boxed{\phi_{ij}(r_0) \;=\; -\epsilon} \tag{6.30}$$

so that

1. $\epsilon = A^2/4B$ measures the strength of the attraction,
2. $\sigma = (B/A)^{1/6}$ measures the radius of the repulsive core: for $r = \sigma$ one finds $\phi = 0$.

The constants ϵ and σ are generally evaluated from measurements in the gas phase of the deviation from the ideal gas behaviour (virial coefficients, compressibility). Their values for He, Ne and Ar are listed in Table 6.2. The LJ potentials for He and Ne are compared in Fig. 6.6, left.

Table 6.2: Lennard-Jones parameters ϵ and σ for He, Ne and Ar, deduced from the properties of low-density gases. The corresponding theoretical nearest-neighbour equilibrium distances R_0 (eq. 6.36) and cohesive energies per atom u_0 (eq. 6.37) in the crystalline phases are compared with the experimental values $R_0^{\rm exp}$ and $u_0^{\rm exp}$.

	ϵ (eV)	σ (Å)	R_0 (Å)	$R_0^{\rm exp}$ (Å)	u_0 (eV)	$u_0^{\rm exp}$ (eV)
He	0.0009	2.56	2.79	–	-0.008	-
Ne	0.0031	2.74	2.99	3.13	-0.027	-0.02
Ar	0.0104	3.40	3.71	3.75	-0.089	-0.08

To better grasp the order of magnitude of the values of ϵ, it is useful to compare them with the thermal energy k_BT at different temperatures. As a rule of thumb, remember that $k_BT \simeq 0.025\,\text{eV}$ for T=300 K.
To evaluate the true binding energy of the atomic pair at $T = 0$ K, the zero-point energy of the relative atomic vibrations has to be added to the LJ potential energy. (Don't confuse the vibrations of the electronic clouds with respect to the nuclei, origin of London forces, with the relative vibrations of nuclei).

6.3.4 Total cohesive energy of a molecular crystal

In a crystal, each atom interacts with all the other atoms via the Lennard-Jones potential.
The total potential energy, neglecting surface effects, is

$$\Phi_{\rm tot} = \frac{1}{2}\,N \sum_{\vec{T}\neq 0} \phi(|\vec{T}|)\,, \tag{6.31}$$

where

a) the sum is over all lattice vectors $\vec{T}$ originating at a given atom,

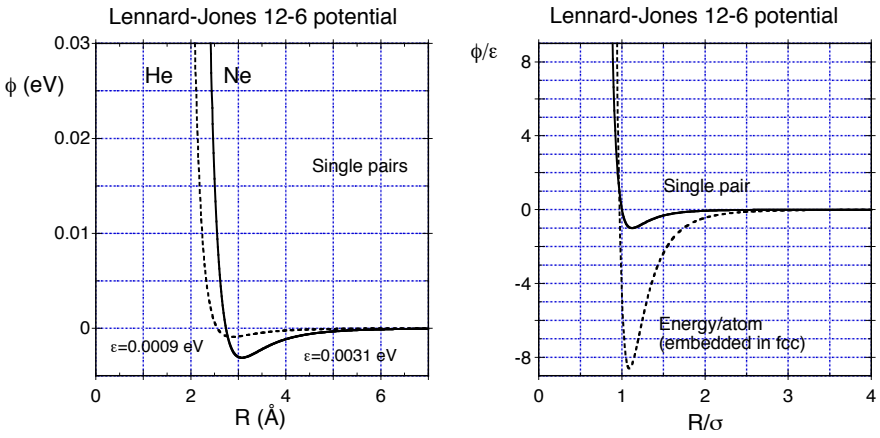

Figure 6.6: Lennard-Jones (LJ) potential. Left: comparison of the LJ single-bond potentials for He and Ne. Right: Comparison between the single-bond LJ potential and the energy per atom of a molecular crystal (where R is the nearest-neighbour distance); the quantities on the axes are expressed in reduced units.

b) N is the total number of atoms,
c) the factor $1/2$ is inserted to avoid counting twice the same distance.

The total energy (6.31) can be expressed in terms of the Lennard-Jones potential (6.29)

$$\Phi_{\text{tot}} = \frac{1}{2} N 4\epsilon \left[\sum\nolimits_j \left(\frac{\sigma}{r_j} \right)^{12} - \sum\nolimits_j \left(\frac{\sigma}{r_j} \right)^{6} \right] , \tag{6.32}$$

where r_j is the distance from the central atom to the j-th atom. It is convenient to express all the r_j values as a function of the nearest-neighbour distance R by introducing the adimensional parameters $p_j = r_j/R$, so that

$$\begin{aligned} \Phi_{\text{tot}} &= \frac{1}{2} N 4\epsilon \left[\sum\nolimits_j \left(\frac{\sigma}{p_j R} \right)^{12} - \sum\nolimits_j \left(\frac{\sigma}{p_j R} \right)^{6} \right] \\ &= \frac{1}{2} N 4\epsilon \left[A_{12} \left(\frac{\sigma}{R} \right)^{12} - A_6 \left(\frac{\sigma}{R} \right)^{6} \right] \end{aligned} \tag{6.33}$$

where the constants

$$A_{12} = \sum\nolimits_j \frac{1}{p_j^{12}} , \qquad A_6 = \sum\nolimits_j \frac{1}{p_j^{6}} \tag{6.34}$$

only depend on the geometrical structure of the crystal.
The London forces are isotropic and the noble gases Ne, Ar, Kr, Xe crystallize in the fcc close-packed structure.
For the fcc structures, A_{12}=12.13 and A_6=14.45. These values are only slightly larger than the coordination number 12, as a consequence of the short-range nature of the interaction.

The total cohesive energy per atom, expressed as a function of the nearest-neighbours distance R, is

$$\phi = \frac{\Phi_{\text{tot}}}{N} = 2\epsilon \left[A_{12} \left(\frac{\sigma}{R} \right)^{12} - A_6 \left(\frac{\sigma}{R} \right)^{6} \right] \tag{6.35}$$

The LJ single-bond potential and the cohesive energy per atom for fcc crystals are compared in Fig. 6.6 (right), where the reduced units r/σ (or R/σ) and ϕ/ϵ are used.

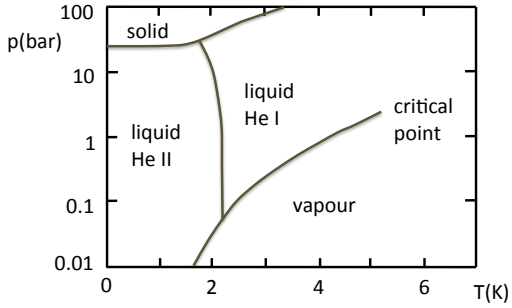

Figure 6.7: Phase diagram of ^{4}He.

The equilibrium nearest-neighbour distance of the LJ potential for the fcc structure is

$$R_0 \;=\; \sqrt[6]{\frac{2A_{12}}{A_6}}\,\sigma \;=\; 1.09\,\sigma \tag{6.36}$$

and the corresponding equilibrium cohesive energy per atom is

$$u_0 \;=\; -\frac{\epsilon A_6^2}{A_{12}} \;=\; -8.6\,\epsilon. \tag{6.37}$$

The theoretical values of R_0 and u_0 calculated according to the LJ potential are compared with the experimental values in Table 6.2. The discrepancy increases when the mass decreases, because of the increasing effect of zero-point energy (see below). For He, no stable solid phase is possible.

6.3.5 Effect of zero-point vibrations

The zero-point vibrational energy per atom of a noble-gas crystal with one atom per primitive cell is given by (see Chapter 7):

$$E_{zp} \;=\; \frac{1}{N}\sum_{\vec{q}s}\frac{1}{2}h\,\nu(\vec{q}s) \tag{6.38}$$

where N is the total number of atoms, the sum is over the $3N$ normal modes labelled by the N wavevectors $\vec{q}$ and the 3 branch indexes s. An exhaustive evaluation of E_{zp} would thus require a full knowledge of the vibrational dynamics of the crystal.

An approximate evaluation of the order of magnitude can anyway be made by the following considerations, concerning a single pair of atoms.

Let the interatomic distance R_0 be assumed as an estimate of the maximum positional uncertainty Δx (certainly an excess estimate). As a consequence of the uncertainty principle, the relative motion of the atomic pair is characterised by an average linear momentum of modulus $p \simeq h/R_0$ (uncertainty relation), to which there corresponds a kinetic energy $E_{zp} = p^2/2m = h^2/2mR_0$. Substituting numerical values, one finds

$E_{zp} \simeq 0.0026$ eV for He
$E_{zp} \simeq 0.0004$ eV for Ne

The main reason of the difference is the different mass of the two atoms. These values E_{zp} can be compared with the binding energies ϵ for a pair of atoms listed in the first column of table 6.2. One can see that E_{zp} is larger than ϵ for He, smaller than ϵ for Ne.

6.3.6 The case of Helium

All inert gases can be liquefied at ambient pressure (1 bar) by lowering the temperature; the boiling points of ^{4}He and Ne are 4.22 K and 27.1 K, respectively.

By further lowering the temperature, inert gases can be solidified at ambient pressure (the melting point of Ne is 24.48 K), with the exception of helium. The zero-point energy of He is larger than the static cohesive energy of the LJ potential at any temperature.

The phase diagram of ^{4}He is shown in Fig. 6.7. A solid phase of ^{4}He can be obtained only at pressures $P > 25$ bar.
The most striking peculiarity of the He phase diagram is the presence of two liquid phases, labeled He I and He II, respectively. ^{4}He has been liquefied in 1908 for the first time. In 1928 the λ transition between the two liquid phases was discovered (the name λ reflects the peculiar shape of the dependence of specific heat on temperature at the transition); in 1938 the super-fluidity of He II was discovered.
The superfluid phase has a number of striking peculiarities: anomalously low viscosity, negative thermal expansivity, anomalously high thermal conductivity, Fountain effect, creeping effects, several different "sound" waves, quantized vortices.
These behaviours are macroscopic manifestations of quantum effects.

6.4 Ionic crystals

In ionic crystals, the atoms are partially ionised. The ions are held together by Coulomb forces (isotropic). The fluctuating-dipole London forces are still present, but their contribution is negligible with respect to the Coulomb interaction.
The simplest model of an ionic crystal is represented by an assemby of $2N$ sperical ions, N with charge $= Ze$ and N with charge $-Ze$, respectively. This model is particularly well suited for I-VII compounds; its goodness progressively decreases when going to II-VI and III-V compounds, where covalent contributions are not negligible.
Electron charge density maps from X-ray elastic scattering confirm this picture of ionic bonds; for the I-VII compounds the charge density is confined within spherical regions around the nuclei; for II-VI and III-V compounds a non negligible charge density appears along the lines connecting pairs of atoms, indicating the partially covalent character of the bond.

6.4.1 A simple phenomenological approach

Let us consider the atoms of Na (column I of the periodic table, Fig. 6.8) and Cl (column VII). The energy balance of the ionic bond of an Na–Cl crystal can be calculated through the sequence of logical steps:

1. Ionisation of each Na atom, leading to the stable [Ne] configuration $1s^2\,2s^2\,2p^6$:
   $\mathrm{Na} + (5.14\ \mathrm{eV}) \rightarrow \mathrm{Na}^+ + e^-$
2. Addition of one electron to each Cl atom, leading to the stable [Ar] configuration: [Ne] $3s^2 3p^6$:
   $\mathrm{Cl} + e^- \rightarrow \mathrm{Cl}^- + (3.61\ \mathrm{eV})$
3. Coulomb attraction between all the ion pairs, which leads to the equilibrium distance within the crystal:
   $\mathrm{Na}^+ + \mathrm{Cl}^- \rightarrow \mathrm{NaCl} + (7.9\ \mathrm{eV\ per\ ion\ pair})$

The formation of the two single ions requires an external energy of 1.53 eV.
Globally, the formation of the entire NaCl crystal leads to a gain of 6.37 eV per Na–Cl pair (7.9 eV due to Coulomb attraction - 1.53 eV necessary to create the ion pair Na^+-C^- pair).

6.4.2 Ionicity and crystal structure

The Coulomb interaction between ions is isotropic, like the fluctuating-dipole London interaction. No close-packed crystal structures are however generated by ionic bonds. The main reason is the alternation of attractive and repulsive long-range Coulomb forces between any ion and the ions of the surrounding coordination shells (nearest-neighbours, next-nearest-neighbours, and so on). The condition that each ion is surrounded only by ions of the opposite charge imposes a steric constraint, so that the maximum coordination number encountered in ionic crystals is $z = 8$.

Example: Form a two-dimensional arrangement of oppositely charged spheres, trying to realise a close packing and contemporarily fulfil the requirement that each sphere is surrounded only

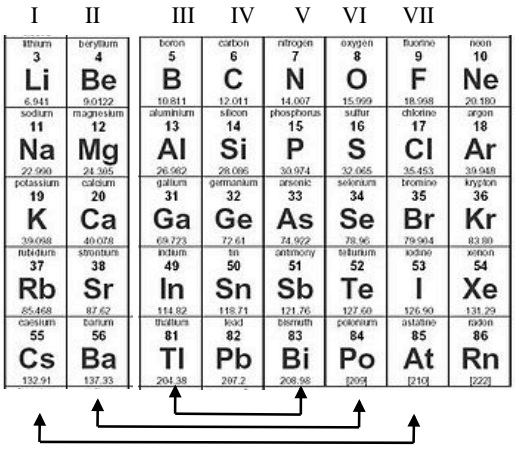

Figure 6.8: An extract of the Periodic Table.

by spheres of the opposite charge.

Besides, when going from I-VII to II-VI to III-V compounds (Fig. 6.8), the ionicity of the bond is progressively mixed with a covalent contribution, which increases the density of electronic charge in the regions between ions, thus creating directional bonds and further reducing the coordination number.
In more details:

a) In I–VII compounds (alkali halides) the atoms are singly ionized and the bond has a strong ionic character
   - a few crystals (CsCl, CsBr and CsI) share the CsCl structure, formed by two inter-penetrating simple cubic lattices of ions of opposite charge (Fig. 6.9, left); the coordination number is $z = 8$;
   - the majority of crystals share the NaCl structure, formed by two inter-penetrating fcc lattices of ions of opposite charge (Fig. 6.9, centre); the coordination number is $z = 6$;

b) In II–VI compounds, the atoms are doubly ionized
   - a number of crystals share the NaCl structure, with $z = 6$
   - other crystals (such as Be compounds, MgTe, CdTe) share the zincblende structure (Fig. 6.9, right) with $z = 4$; the low coordination indicates the presence of a mixed covalent and ionic bond

c) The III-V crystals, such as GaAs, exhibit the zincblende structure, $z = 4$; the bond shares ionic and covalent characters.

6.4.3 Potential energy of a single ion pair

The potential energy of an ion pair (e.g. Na^+ and Cl^-) is the sum of an attractive Coulomb term and a repulsive term due to the Pauli exclusion principle.
The repulsive term can be represented by different analytical expressions, such as

$$\phi_{\text{rep}}(r) \simeq \frac{A}{r^9}\,, \qquad \text{or} \qquad \phi_{\text{rep}}(r) \simeq \lambda e^{-r/\rho} \tag{6.39}$$

(the second expression, where $\lambda \gg 1$ eV, is known as Born-Mayer potential).
The attractive term is the Coulomb energy for an ion pair

$$\phi_{\text{coul}}(r) = -\frac{1}{4\pi\epsilon_0}\,\frac{(Ze)^2}{r}\,, \tag{6.40}$$

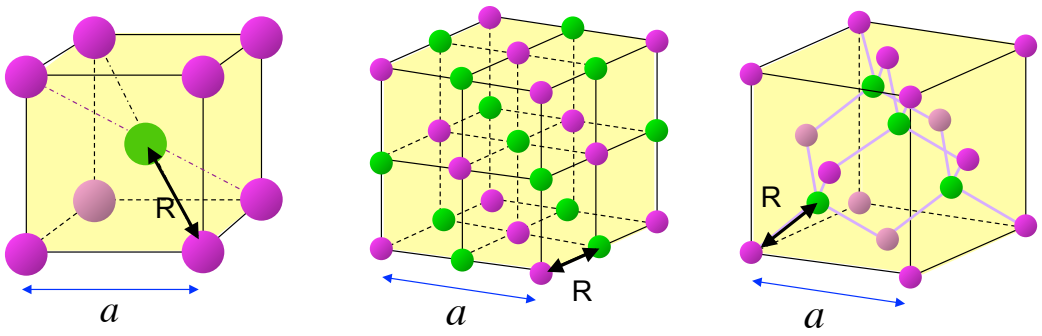

Figure 6.9: From left to right: the CsCl, NaCl and zincblende structures. CsCl: sc Bravais lattice with two ions of opposite sign per primitive cell in positions $(0,0,0)$ and $(1/2,1/2,1/2)$. NaCl: fcc Bravais lattice with two ions of opposite charge at positions $(0,0,0)$ and $(1/2,1/2,1/2)$. Zincblende: fcc Bravais lattice with two ions of opposite charge at positions $(0,0,0)$ and $(1/4,1/4,1/4)$.

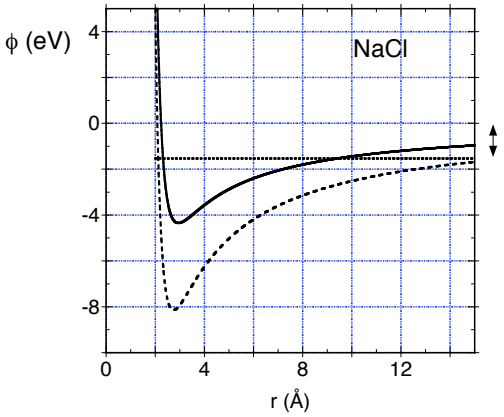

Figure 6.10: Coulomb potential energy of a single pair of Na^+ and Cl^- ions (continuous line) and for a pair embedded in a crystal (dashed line). The bottom of the dashed line corresponds to the energy necessary to dissociate the crystal into two ions (7.9 eV). The horizontal dotted line corresponds to the energy necessary to create two ions starting from two neutral atoms (1.53 eV). The cohesive energy is 7.9-1.53=6.37 eV.

where Ze is the ionic charge. For a singly ionised compound ($Z = 1$), the Coulomb energy (in eV) is related to the inter-ionic distance r (in Å) by the numerical relation

$$\phi_{\text{coul}}(r) = -\frac{14.386}{r} \tag{6.41}$$

In comparison with the fluctuating dipole London interaction, the Coulomb interaction is much stronger and long ranged.
The resulting potential energy for a single isolated pair of Na^+ and Cl^- ions, sum of the repulsive and Coulomb contributions, is represented in Fig. 6.10 as a continuous line. It is evident that only below $r \simeq 9.5$ Å does the attractive Coulomb energy overcome the energy necessary to create the two ion pairs (horizontal line at -1.5 eV).

6.4.4 Electrostatic cohesive energy – Madelung constant

Let us now calculate the total cohesive energy of an ionic crystal.
Each ion is surrounded by a very large number of coordination shells of alternating charge. For concreteness, let us consider the case of the NaCl structure (Fig. 6.9, centre). For a given central ion i of a given charge, the Coulomb contribution to the total energy is given by a sum over all its coordination shells:

$$\Phi_{\text{coul}}^{(i)} = -\frac{e^2}{4\pi\epsilon_0}\frac{1}{R}\left[6 - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \cdots\right] \tag{6.42}$$

where R is the nearest-neighbours distance. The first term within square parentheses is relative to the $z = 6$ first-shell ions of opposite charge at the distance R, the second term is relative to the $z = 12$ second shell ions of equal charge at the distance $R\sqrt{2}$, and so on.
The sequence of terms within the square parentheses depends on the crystal structure. Its sum is called the "Madelung constant" α, so that

$$\Phi_{\text{coul}}^{(i)} = -\,\alpha\,\frac{e^2}{4\pi\epsilon_0}\,\frac{1}{R}\,. \tag{6.43}$$

The evaluation of the Madelung constant is far from trivial, due to the long-range character of the Coulomb forces and to the consequent slow convergence of the series in (6.42). Efficient calculation schemes have been developed. For the NaCl structure, the Madelung constant has the value $\alpha = 1.7475$.

Let us now add a repulsive term, for example

$$\Phi_{\text{rep}}^{(i)} = z\,\lambda\,e^{-R/\rho}\,, \tag{6.44}$$

where z is the first-shell coordination number (number of nearest-neighbours, for NaCl $z = 6$).
To evaluate the total energy, let N be the number of ionic pairs, $2N$ the number of ions, so that

$$\Phi_{\text{tot}} = \frac{1}{2}\,2N\,\left[\Phi_{\text{rep}}^{(i)} + \Phi_{\text{coul}}^{(i)}\right] = N\,\left[z\,\lambda\,e^{-R/\rho} - \alpha\,\frac{e^2}{4\pi\epsilon_0}\,\frac{1}{R}\right] \tag{6.45}$$

The sum of the Coulomb attractive potential energy and of the repulsive potential energy for a pair of Na^+ and Cl^- ions embedded in a crystal, Φ_{tot}/N, is represented in Fig. 6.10 as a dashed line. The cohesive energy, say the energy necessary to decompose the crystal into neutral atoms at infinite distance, is the difference between the horizontal dotted and the dashed line.

6.5 Covalent crystals

The covalent bonds are characterised by the sharing of electrons by neighbouring atoms, leading to high electronic densities along the inter-nuclear distances. These electronic densities can be experimentally detected by x-ray diffraction experiments. Covalent bonds are strongly directional.

Different approaches have been developed to account for covalent bonds in molecules; they are generally based on the linear combination of one-electron atomic orbitals (LCAO). These approaches can be extended to covalent crystals, to gain a qualitative understanding of the origin of their structure.
An exhaustive quantitative understanding of the electronic properties of covalent crystals, including for example their electric conductivity, requires a full quantum-mechanical treatment based on the solution of the Schrödinger equation, supported by symmetry considerations.

6.5.1 Molecular orbital approach

In the molecular orbital (MO) approach, one-electron wavefunctions extending over the entire molecule are taken as basis functions.
The molecular orbital (MO) approach was considered in Section 5.6 with reference to the CH_4 molecule and to the irreducible representations of its symmetry group. There, we found that only two families of linear combinations of the atomic orbitals of the carbon and oxygen atoms were compatible with the character table of the T_d symmetry group of the molecule: one MO, basis of the A_1 irreducible representation, and three MO, basis of the T_2 irreducible representation; the two MOs correspond to two bonding energy levels, the lowest one (a_1) non-degenerate which can accommodate two electrons, the highest one (t_2) of degeneracy three which can accommodate six electrons.
The extension of the MO method to crystals structures is the basis of the "tight binding" approach for the calculation of the electron energy band structure.

6.5.2 Valence orbitals approach

An alternative approach reconstructs the high electron density along the bonding directions of a given atom (the valence orbitals) starting from the its electronic structure. In a sense, this approach can be considered as the extension to covalent bonds of the idea of pair interaction.
A key problem is to find all the possible atomic orbitals from which a set of valence orbitals can be formed, which accounts for the actual structure of the system. The directionality of covalent bonds is explained in terms of suitable hybridizations of atomic orbitals. This approach allows one to understand quite directly the origin of the local coordination of atoms in molecules and crystals, without however giving the possibility of estimating the structure of MOs.

As a simple example, let us consider here the case of carbon.

Carbon: Hibridisation of atomic orbitals

In the ground state, the electron configuration is: $1s^2, 2s^2, 2p_x^1, 2p_y^1, 2p_z^0$ (Fig. 6.11, top).
By spending an amount of energy of 3.6 eV, it is possible to promote one of the $2s$ electrons to the $2p_z$ empty orbital, giving rise to the excited configuration: $1s^2, 2s^1, 2p_x^1, 2p_y^1, 2p_z^1$ (Fig. 6.11, bottom).
The binding involves the $2s$ and $2p$ atomic orbitals. Neglecting the radial dependences, the angular dependences of the four orbitals, in standard spherical coordinates, are:

$$s = 1\,, \quad p_x = \sqrt{3}\sin\theta\cos\phi\,, \quad p_y = \sqrt{3}\sin\theta\sin\phi\,, \quad p_z = \sqrt{3}\cos\theta\,. \tag{6.46}$$

The energy spent for the excitation of a single atom is recovered when the atom is bonded via the hybridisation mechanism. Let us consider here the sp^3 and sp^2 hybridisations.

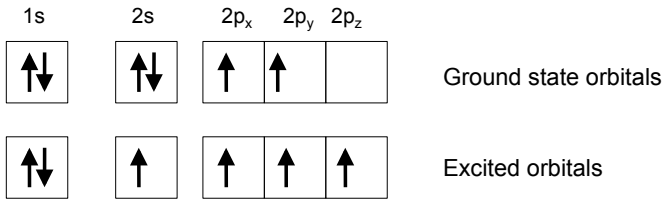

Figure 6.11: Ground state and excited orbitals of carbon atoms.

The sp^3 hibridisation

The $2s$ orbital and the three $2p$ orbitals of the excited state can be linearly combined to form four new orbitals

$$\begin{aligned}
\Psi_1 &= (1/2)\,(s + p_x + p_y + p_z) \\
\Psi_2 &= (1/2)\,(s + p_x - p_y - p_z) \\
\Psi_3 &= (1/2)\,(s - p_x + p_y - p_z) \\
\Psi_4 &= (1/2)\,(s - p_x - p_y + p_z)
\end{aligned} \tag{6.47}$$

The four new hybrid orbitals have tetrahedral symmetry (Fig. 6.12, left panel), and transform only among themselves under the symmetry transformations of the T_d group: they thus form a basis for a 4-dimensional (reducible) representation of the T_d group.
The sp^3 hybrid orbitals can give rise to strong σ bonds with neighbouring atoms, each bond consisting of two electrons, one for each of the two participating atom.
Examples of tetrahedral sp^3 hybridisation are the methane molecule CH_4 and the diamond crystalline structures of C, Si, Ge.

Group theory considerations for sp^3 hybridisation

The above procedure can be reversed: given the set of bond angles which determines the crystal structure, the group theory can help in finding the atomic orbitals whose hybridisation can originate that given structure at the local level.
Let us consider again the character table of the T_d group, which is reproduced in Table 6.3.
In order to get tetrahedral bonds, we need valence orbitals which are transformed into themselves by the symmetry operations of the T_d group. Such orbitals are the basis of a 4×4 reducible representation of the T_d group; the characters of such a representation can be evaluated by counting the orbitals that are unchanged by the different symmetry operations, and are listed in the last line of Table 6.3.
One can easily check that the 4×4 reducible representation can be decomposed into two irreducible representations, A_1 and T_2, whose bases are the s and p orbitals, respectively.

Table 6.3: Character table of the symmetry group T_d. The last line contains the characters if the reducible 4×4 representation based on the binding σ orbitals of the tetrahedral coordination.

T_d	e	$8C_3$	$3C_2$	6σ	$6S_4$	orbitals
A_1	1	1	1	1	1	s
A_2	1	1	1	-1	-1	
E	2	-1	2	0	0	$(d_{z^2}, d_{x^2-y^2})$
T_1	3	0	-1	-1	1	
T_2	3	0	-1	1	-1	$(p_x, p_y, p_z)(d_{xy}, d_{xz}, d_{yz})$
$\Gamma^{(4\times4)}$	4	1	0	2	0	

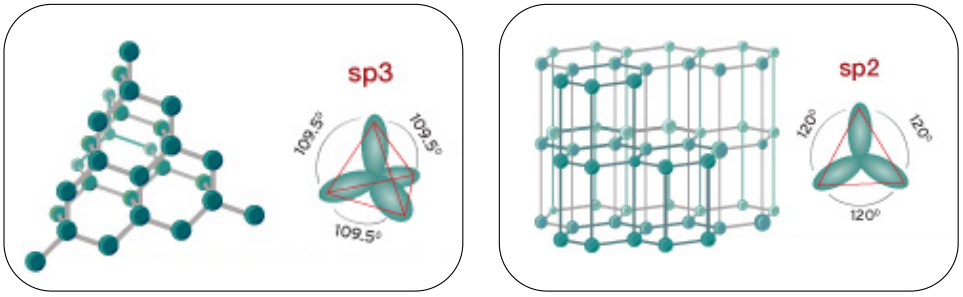

Figure 6.12: Diamond structure and sp^3 hybridisation (left); stacking of graphene planes in the graphite structure and sp^2 hybridisation (right).

The sp^2 hybridisation

In the sp^2 hybridisation, three binding orbitals are created by the linear combination of the $2s$ and two $2p$ atomic orbitals:

$$\begin{aligned} \Psi_1 &= (1/\sqrt{3})\,[s + \sqrt{2}\,p_x] \\ \Psi_2 &= (1/\sqrt{3})\,[s - (1/\sqrt{2})\,p_x + (\sqrt{3/2})\,p_y] \\ \Psi_3 &= (1/\sqrt{3})\,[s - (1/\sqrt{2})\,p_x - (\sqrt{3/2})\,p_y] \end{aligned} \qquad (6.48)$$

The three new hybrid orbitals have triangular symmetry (Fig. 6.12, right panel) and transform only among themselves under the symmetry operations of the D_{3h} group; they thus form a basis for a 3-dimensional (reducible) representation of the D_{3h} group.

The sp^2 hybrid orbitals can give rise to three planar σ molecular orbitals with neighbouring atoms, corresponding to three strong bonds.
The p_z atomic orbital gives origin to a delocalized π molecular orbital.

Example 1: The ethilene molecule C_2H_4 and the benzene molecule C_6H_6.

Example 2: Graphene is a two dimensional carbon crystal with honeycomb exagonal structure. The nearest-neighbours distance is $R = 1.42$ Å. The delocalised π orbital guarantees high mechanical strength and high electrical conductivity.

Example 3: Graphite is formed by a stacking of parallel graphene sheets, held together by weak Van der Waals forces. The distance between the graphene sheets is 3.35 Å.

Group theory considerations for sp^2 hybridisation

As for the sp^3 hybridisation, even for the sp^2 hybridisation the above procedure can be reversed.
Let us consider the character table of the D_{3h} group, which is reproduced in Table 6.4.
In order to get triangular bonds, we need σ valence orbitals which are transformed into themselves by the symmetry operations of the D_{3h} group. The three σ orbitals are the basis of a 3×3 reducible representation of the D_{3h} group; the characters of such a representation can be evaluated by counting the orbitals that are unchanged by the different symmetry operations, and are listed in the Γ^σ line of Table 6.4.
One can easily check that the 3×3 Γ^σ reducible representation can be decomposed into two irreducible representations, A_1' and E', whose bases are the s and (p_x, p_y) orbitals, respectively.
If d orbitals were participating to the bond, other hybridisations could be possible: $sd^2 = s + (d_{xy}, d_{x^2-y^2})$, $dp^2 = d_{z^2} + (p_x, p_y)$ and $d^3 = d_{z^2} + (d_{xy}, d_{x^2-y^2})$

Table 6.4: Character table of the symmetry group D_{3h}. Single prime $'$ and double prime $''$ refer to basis functions that don't or do change sign under σ_h inversion. The last line contains the characters if the 3×3 reducible representation based on the σ orbitals of the triangular coordination.

D_{3h}	e	σ_h	$2C_3$	$2S_3$	$3C_2'$	$3\sigma_v$	orbitals
A_1'	1	1	1	1	1	1	s, d_{z^2}
A_2'	1	1	1	1	-1	-1	
A_1''	1	-1	1	-1	1	-1	
A_2''	1	-1	1	-1	-1	1	p_z
E'	2	2	-1	-1	0	0	$(p_x, p_y), (d_{xy}, d_{x^2-y^2})$
E''	2	-2	-1	1	0	0	(d_{xz}, d_{yz})
Γ^σ	3	3	0	0	1	1	

6.5.3 Full quantum approach

The localisation of the density of valence electrons in covalent solids explains their property of being electrical insulators at $T = 0\,\mathrm{K}$ (as ionic crystals are). A number of covalent crystals, such as Si and Ge, exhibit weak conductivity when T increases, and are classified as semi-conductors. The conductivity of semiconductors increases with temperature.

A full understanding of the structural and conductivity properties of covalent (and metallic) crystals can be obtained from the solution of the Schrödinger equation (6.1), which is facilitated by the constraints due to the translational symmetry of crystals; the behaviour of the electron wavefunctions in a periodic potential, described in the reciprocal space, leads naturally to the band structure and to the conductivity properties. This approach is considered in Chapter 9.
We only note here that various approximate methods have been devised for solving the electron Schrödinger equation in crystals; two limiting approaches are worth to be mentioned:

1. The nearly free electrons approach, which considers the valence electrons as moving in a weak periodic field due to the interaction with the ion lattice; this approach is particularly suitable for metals.
2. The tight binding approach, which extends to the crystals the molecular orbitals formalism; this approach is particularly fruitful for covalent crystals.

6.6 Metals

Metals are characterised, among other properties, by a high electrical conductivity, of the order of $10^8 - 10^6$ S/m. For a comparison, the room-temperature conductivities of some substances are compared in Table 6.5

Table 6.5: Electrical conductivities of some metals and insulators.

System	conductivity	unit
copper	5.9×10^7	S/m
silicon	1.5×10^{-3}	S/m
glasses	$10^{-11} - 10^{-15}$	S/m
sulfur	10^{-16}	S/m

The high electrical conductivity of metals suggests a high mobility of the valence electrons. The electrical conductivity of metals decreases when the temperature increases.

The simplest model to explain the electrical conductivity consists in considering the valence electrons as de-localised, forming a gas of free or nearly free particles obeying the Fermi-Dirac statistics. This crude model can also explain the main features of the electronic contribution to the specific heat.

The model suggests an appealing explanation for the crystal structure of metals: the electron gas permeating the lattice structure guarantees the attractive force that brings together the ions. Contrary to the case of the localised bond in covalent crystals, the de-localised electron gas in metals gives rise to an isotropic bond.
One can thus expect that metals crystallise in the close-packed fcc or hcp structures, with coordination number 12 (Fig. 6.13, left and centre). This is actually the case, for example, of Cu, Ag, Au (fcc) or Be, Mg, Cd (hcp).
Other metals however, such as Fe, Cr and Mo, exhibit the less packed bcc structure, with coordination number 8 (Fig. 6.13, right), suggesting a non-negligible influence of the directional d atomic orbitals on the bond formation.
Moreover, polymorphism is frequently found in metal structures; for example, Li and Na are hcp at low temperatures, bcc at high temperatures. These facts evidence the inadequacies of the free electron gas model to fully account for all properties of metals.

A deeper understanding of the properties of electrical conductivity of metals and of the differences with respect to covalent crystals (insulators or semiconductors) requires to take into account the interaction of electrons with the ion lattice which gives rise to the band structure (Chapter 9).
A further improvement is to take into account the interaction between electrons, which is responsible for important phenomena such as ferro-magnetism. It is now quite well stablished that the non close-packed structure of some metals is correlated to magnetic properties. As a matter of fact, it has been recently found that Fe undergoes the bcc to hcp phase transition and loses its ferromagnetic properties when subjected to a pressure higher than about 13 GPa.

6.7 Hydrogen-bonded solids

A number of organic and inorganic molecules are linked together by the presence of hydrogen ions in between. Such hydrogen bonds share some properties of both covalent and ionic bonds, with

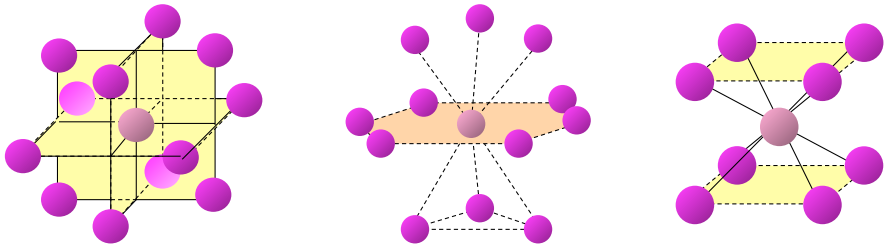

Figure 6.13: Atomic coordination in face centred cubic (left), hexagonal close packed (centre) and body centred cubic (right) structures).

several relevant peculiarities:

1. Hydrogen has a high ionisation potential (13.6 eV); a perfectly ionic bond, based on the complete ionisation of H, would require a large amount of energy.
2. Only one $1s$ electron is present in hydrogen. The stable configuration of the $1s$ shell is $1s^2$ (He); an hydrogen atom can be involved in only one covalent bond (as in the H_2 or H_2O molecules).
3. The hydrogen ion H^+ is a just proton, whose size is of the order of 10^{-15} m, several orders of magnitude smaller than the size of any other ion.

The water molecule

Oxygen and hydrogen atoms are held together in the water molecule H_2O by covalent bonds with non-negligible ionic character. The electronic configuration of oxygen is $1s^2, 2s^2, 2p^4$.

Two of the four $2p$ electrons of the oxygen atom form covalent bonds with the two hydrogen atoms: each bond is formed by a pair of electrons, one from oxygen, one from hydrogen. The two bonds form an angle of 104.45°. The hydrogen ions (protons) are at a distance of 0.958 Å from the oxygen nucleus.

The remaining two $2p$ electrons of the oxygen atoms, not participating to the covalent bonds with hydrogen, form a "lone pair".

Liquid and solid water

The hydrogen bond is responsible for the structures of liquid and solid water.
Each oxygen ion of a given molecule can be linked to two hydrogen ions of different molecules, via the two electrons of the lone pair. Otherwise stated, each water molecule can be connected to up to four other molecules:

a) two via the hydrogen atoms of the molecule, which are attracted by the electrons of the lone pairs of other molecules
b) two via the two electrons of the lone pair, which are attracted by the hydrogen ions of other molecules

The hydrogen bond is weaker that the covalent bond, but stronger than the Van der Waals bond. That's why the melting and ebullition points of water are much higher than the corresponding points of other molecular crystals.
Several different structures are possible for solid water. At ambient pressure, the structure of ice is hexagonal. The position of protons is not ordered in the crystal structures of ice, giving to different equivalent configurations and to a residual entropy at $T = 0\,\mathrm{K}$.

The hydrogen bond is present also in other inorganic and organic compounds; among them one can cite the DNA molecule, where the hydrogen bond is responsible of folding.

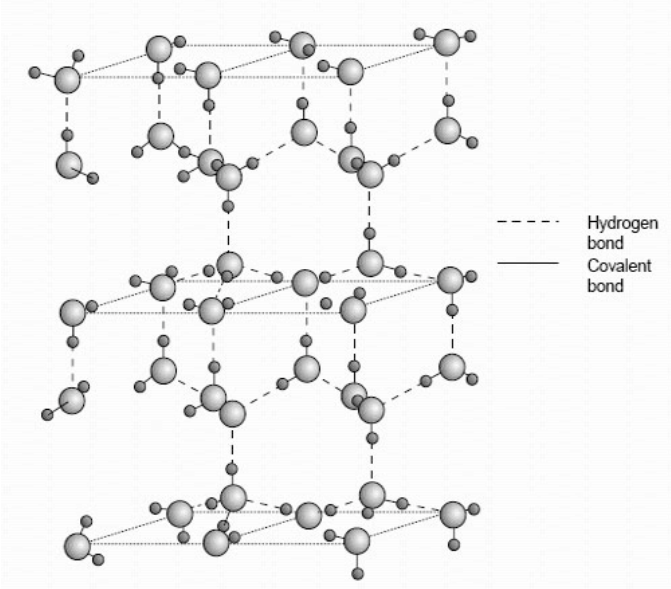

Figure 6.14: A sketch of the crystal structure of ice. Each oxigen ion (large circles) is covalently bound to the two hydrogen ions belonging to its molecule, and to two other farther hydrogen ions belonging to other molecules.

6.8 Non-crystalline solids

Non-crystalline atomic aggregations are frequently found in both natural and man-made materials, e.g. volcanic lava and windowpanes, respectively.
The interest for non-crystalline solids is increasing, in view of their many possible technological applications but also because understanding their structural, electronic and vibrational properties represents a challenge for basic research.

6.8.1 Introduction

Thermodynamics considerations

According to Thermodynamics, the equilibrium state of a pure substance is characterised by the atomic configurations that give rise to the absolute and local minima of the Gibbs function $G(T, p)$ with respect to the T and p variables. The absolute minimum corresponds to stable equilibrium; local minima correspond to metastable states.
Thermodynamic equilibrium states (stable and metastable) correspond to crystal structures (Section 6.2). Non-crystalline structures are thermodynamically out of equilibrium.

Structural considerations

Non-crystalline solids are characterised by the lack of long-range order. The experimental signature of long range order is the presence of Bragg peaks in diffraction patterns (Chapter 11): the degree of crystallinity of a sample is thus measured by the presence and intensity of Bragg peaks. For a completely non-crystalline solid, Bragg peaks are absent at all (Fig.6.15).
The analysis of diffraction patterns from crystalline solids allows the reconstruction of the three-dimensional lattice + basis structure (with some possible limitations due to the so called phase problem). For non crystalline solids, as for liquids, the analysis of diffraction patterns leads to one-dimensional statistical descriptions in terms of radial distributions (Section 2.6).

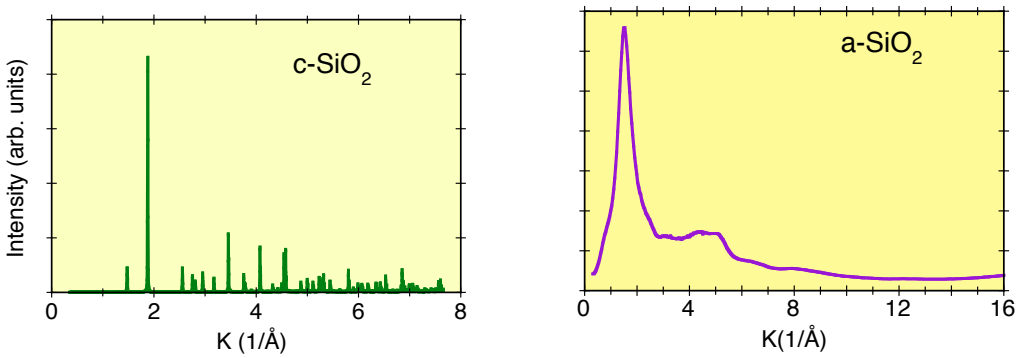

Figure 6.15: X-ray diffraction pattern of powdered crystalline SiO_2 (left) and of amorphous SiO_2 (right). The quantity on the horizontal axis is the magnitude of the scattering vector $K = 4\pi \sin\theta/\lambda$, where θ is the scattering angle and λ is the X-ray wavelength.

Even in the absence of long range order, short-range order is always present even in non-crystalline solids. It is connected to correlations between nearest-neighbour atoms, due to directional bonds in covalent structures or to close-packing in metallic structures. The short range order is revealed by the well defined first peak in the radial distribution function RDF (Section 2.6) reconstructed from diffraction patterns.
In many cases, some forms of medium range order are present too, revealed by the structures of the RDF beyond the first peak.

Example 1: In glassy silica SiO_2, silicon atoms are always coordinated to four oxygen atoms and each oxygen atom is always coordinate to two silicon atoms, as in crystalline SiO_2; the bond distance and bond angles are however not all perfectly equal.

Example 2: In metallic structures, each atom tends to coordinate the nearest neighbour atoms at the distance of minimum approach.

6.8.2 Production methods

To obtain a non-crystalline solid, the preparation procedure must be such that the thermodynamic stable equilibrium cannot be attained. Actually, a number of substances can be found in both crystalline and non crystalline form (e.g. SiO_2 or Ge), depending on the preparation procedure. Various are the factors that influence the choice of a substance between the crystalline and the non-crystalline form. Among them, the type of chemical bond and the velocity of the process.

Three broad categories of preparation methods can be distinguished; starting point can be a disordered phase (liquid or vapour) or an ordered crystalline phase.

1. *Fast cooling of a liquid phase.*
   The formation of the equilibrium crystalline phase at the melting point requires the formation of small crystalline nuclei and their progressive growth. When the liquid phase is cooled, its viscosity progressively increases. If the cooling is sufficiently fast, the atomic kinetics which leads to the process of nucleation and growth is prevented by the increasing viscosity, the liquid is undercooled and at last it is frozen into a disordered solid structure (glass structure). The kinetics of atomic rearrangement is easier for non directional bonds (metals) than for directional covalent bonds. Accordingly, different are the cooling speeds that one should use to avoid crystallisation.

2. *Vapour condensation on a cold surface.*
   Atoms of one element or a mixture of elements are deposited on a surface (substrate), which is maintained sufficiently cold to prevent their ordered crystalline arrangement. In general,

thin amorphous films are produced. If the temperature of the substrate is sufficiently high and the deposition rate sufficiently low, a crystalline phase is instead formed.
Different method have been devised to obtain the initial vapours from solid sources, from Joule heating to sputtering with ion beams.

3. *Disordering of a crystalline phase.*
Amorphisation of a crystalline solid can be achieved by different methods, e.g. impact of fast ions or neutrons (radiation damage) or mechanical friction or impact of shock waves.

Influence of the chemical bond

Although no simple general rule exists, some general considerations can be made about the relation between the nature of the chemical bond and the easiness to solidify in a non-crystalline state.
Isotropic bonds represent very little constraints to atomic kinetics, so that the melts of isotropically bonded solids (such as metals) easily crystallise event at very high cooling rates.
Directional covalent bonds represent an impediment for the atomic rearrangements necessary to obtain a crystalline state starting from a disordered state.

6.8.3 Glasses

An important category of non-crystalline solids is represented by glasses. Glasses are non-crystalline solid obtained by solidification from melt through a process that is called "glass transition".
If the cooling of the melt is fast enough (depending on the material), the liquid can be undercooled below the fusion temperature T_f (temperature of transition to the crystalline state). When the undercooled liquid is further cooled, the increased viscosity prevents the crystallisation, and when the viscosity attains a value of typically 10^{13} poise (1 poise $= 0.1\,\mathrm{N\,s/m^{-2}}$) the substance is frozen into a state that maintains the topological disorder of the liquid state but is solid for all the other respects.

Viscosity

The dynamic viscosity of a fluid expresses its resistance to shearing flows. Let us consider a layer of fluid trapped between two horizontal plates, moving horizontally in the x direction at different constant speeds. The horizontal shear force per unit area A necessary to maintain the relative motion can be expressed as

$$\frac{F}{A} = \eta \frac{dv_x}{dz}\,, \tag{6.49}$$

where dv_x/dz is the vertical gradient of the horizontal velocity and η is the viscosity coefficient, which is measured in $\mathrm{N\,s\,m^{-2}}$. It is customary to use the old cgs unit, $1\,\mathrm{poise} = 0.1\,\mathrm{N\,s\,m^{-2}}$.

The glass transition

The transition to the glassy state is called glass transition. The temperature dependence of the volume of a system which solidifies from melt is different for different cooling rates. In Fig. 6.16 the glass transition behaviour (continuous lines) is compared with the first-order liquid-solid phase transition (dashed line) The glass transition is characterised by a glass transition temperature T_g, which is not as sharply defined as the melting temperature T_m; its value can vary depending on the cooling rate; one better speaks of a glass transition range around T_g.
Plots similar to Fig. 6.16 can be drawn for the entalpy or the viscosity.
At the glass transition, the specific heat and the thermal expansion coefficient diverge.

The cooling rate required to form a glass depends on the substance and on the type of bonds.
To form a crystal, time has to be allowed to organise the regular lattice structure, breaking and reconstructing bonds. For covalent substances, obtaining a good crystal generally requires a slow cooling.

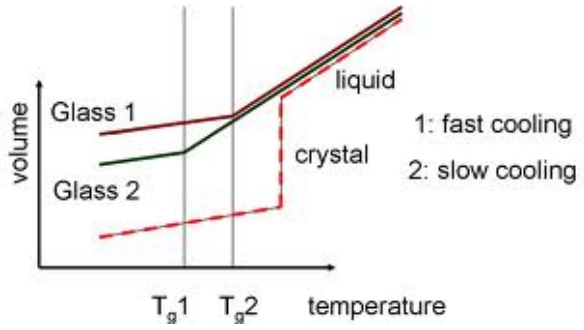

Figure 6.16: Volume variation for the glass transition (continuous lines) and for a crystalline solidification (dashed line).

One can expect that it is easier to form crystals from metallic melts (isotropic bonds) than from covalent liquids (directional rigid bonds).

Oxide glasses

Glasses are easily formed from many oxides. One distinguish

- glass former oxides, such as silica SiO_2 or boron oxide B_2O_3, which form glasses by themselves,
- glass modifier oxides, such as Na_2O, which can be added to the former oxides to modify the chemical and physical properties of the resulting glass.

Let us consider some examples of oxide glasses.

1. Silica SiO_2 is one of the most common-glass formers.
   Silica can be found in a large number of crystalline polymorphs (α and β quartz, α and β tridymite,...). All the crystalline structures are formed by Si–O_4 regular tetrahedra joined by the corners in different configurations. Each Si atom is coordinated to 4 O atoms and each O atom is coordinated to 2 Si atoms.
   In the glass form, the structure is still based on Si–O_4 tetrahedra connected by the corners. The tetrahedra are however not perfect and the tetrahedral groups are connected with random orientations. The structure is generally depicted as a continuous random network (Fig. **??**).
   Silica is the main ingredient of most commercial glasses, from windowpanes to containers for liquids. The glass transition of pure silica is quite high, above 1500 K. To reduce the glass transition and to ameliorate the chemical durability, typically about 25% of the composition is made by other oxides, such as Na_2O, CaO, etc. A typical composition of industrial glasses could be: 70 SiO_2, 20 Na_2O,10 CaO.

2. Boron oxide B_2O_3 is perhaps the most effective glass former. Actually, B_2O_3 can be crystallised with great difficulty.
   In the glass form, each boron atom is coordinated to 3 O atoms and each O atom is coordinated to 2 B atoms. Several hexagonal boroxol rings have been observed, formed by alternating corner of B and O.
   Boro-silicate glasses, based on silicon and boron oxides and characterised by low thermal expansivity, are used in chemistry laboratories and for cooking. A typical composition is: 80.6 SiO_2, 12.6 B_2O_3, 4.2 Na_2O, 2.2 Al_2O_3, 0.04 Fe_2O_3, 0.1 CaO, 0.05 MgO, 0.1 Cl.

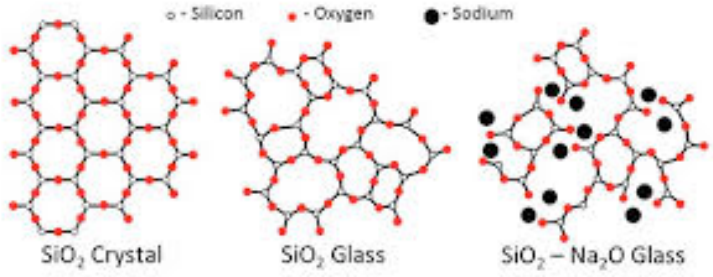

Figure 6.17: Crystalline structure of SiO_2 (left), continuous random network of SiO_2 glass (centre), modified random network of SiO_2-Na_2O glass (right).

Metallic glasses

The first metallic glasses have been obtained in 1960, by fast cooling (about 10^6 K/s) of an alloy $Au_{75}Si_{25}$ in the form of a thin ribbon. The fast cooling is necessary to prevent the formation of the crystal structure, eased by the isotropy of the metallic bond. The alloy composition creates some impediment to crystallisation and helps the glass formation.
During the years, new alloys have been found and the cooling rates have been reduced in some cases down to 1 K/s and bulk metallic glasses (BMG) are now available; generally the large number of alloy components favour some sort of "confusion effect" that prevents the formation of the regular crystalline structure.
Amorphous metals can be also created by vapour deposition (again very fast cooling) or by ion irradiation of crystalline samples.
Among the possible advantages of amorphous metal alloys is the absence of grain boundaries, which allows a better resistance to corrosion.

6.8.4 Amorphous semiconductors

The production of amorphous semiconductors (e.g. a-Si, a-Ge, a-Se) is generally made by vapour deposition. The degree of disorder depends on the substrate temperature. A high substrate temperature favours the movement of adsorbed atoms and the formation of crystalline structures. In amorphous Si, a large number of atoms is four-coordinated as in the crystal, but some atoms are only three-coordinated and the fourth bond is 'dangling' (Fig. 6.18). Dangling bonds can be saturated by the insertion of hydrogen.

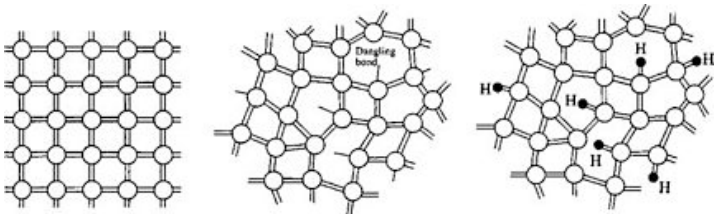

Figure 6.18: Crystalline structure of Si (left), continuous random network of amorphous Si with dangling bonds (centre), hydrogenated amorphous Si (right).

6.8.5 Configurational landscape

From the microscopic point of view, we can in principle describe the structure of a non-crystalline system by a configuration space with $3N$ dimensions, where N is the number of atoms of the

system. Each point of the configuration space corresponds to a choice of the cartesian coordinates of all the atoms.
In spite of its actual impracticability, such an abstract description is the starting point to try to elucidate some physical properties of non-crystalline materials, and in particular of glasses.
Let us suppose that the atoms interact via a pair potential, for example the Lennard-Jones potential. To each point of the configuration space one can in principle associate a total potential energy V. Some points of the configuration space are characterised by local minima of the total potential energy; they correspond to states of microscopic meta-stable equilibrium.
The absolute minimum corresponds to the crystalline structure (macroscopic thermodynamic equilibrium).

6.9 An introduction to Density Functional Theory (DFT)

Let us now come back to the Schrödinger equation for electrons (6.13) of the Born-Oppenheimer approximation. In this Section, we omit the term V_{nn} (ion-ion potential energy) as well as the indices e when not necessary:

$$H\,|\Psi\rangle \;=\; [T_e + V_{ne} + V_{ee}]\;|\Psi\rangle \;=\; E\,|\Psi\rangle\,. \tag{6.50}$$

For a given ionic configuration $\{R\}$, the electron wavefunctions $\Psi(\vec{r}_1, \vec{r}_2, \ldots \vec{r}_n)$ are defined in a $3n$-fold configuration space, where n is the number of electrons.
We focus attention in this Section on the ground-state energy of a system of many interacting electrons. The ground-state energy is a functional of the wavefunction

$$E \;=\; E\,[\Psi(\vec{r}_1, \vec{r}_2, \ldots \vec{r}_n)]\,. \tag{6.51}$$

For a large number n of electrons, the solution, even approximate, of the Schrödinger equations leads to unmanageable wavefunctions.
A dramatic improvement for the calculation of ground-state properties of many-electron systems is represented by the Density Functional Theory (DFT). In the following we will try to describe step by step some basic concepts about the DFT and its applications.

6.9.1 Ground state energy and density functional

Let us consider a system of interacting electrons in an external potential, e.g. V_{ne}, due to the interaction between electrons and ions
According to the Hohenberg and Kohn theorem (1964), the ground-state energy of the many-electron system can be expressed as the functional of the electronic charge density, which is defined in the three-dimensional space

$$E \;=\; E\,[\rho(\vec{r})]\,. \tag{6.52}$$

The problem is thus reduced from the evaluation of the wavefunction $\Psi(\vec{r}_1, \vec{r}_2, \ldots \vec{r}_n)$ to the evaluation of the density $\rho(\vec{r})$.
The energy functional can be decomposed as

$$E\,[\rho(\vec{r})] \;=\; \int V_{en}(\vec{r})\,\rho(\vec{r})\,d\vec{r} \;+\; F[\rho(\vec{r})]\,, \tag{6.53}$$

where the first term on the right depends on the external potential, the density functional $F[\rho(\vec{r})]$ is a universal function of the density, and includes the kinetic energy of the electrons and the potential energy of the electron-electron interaction.
According to the Hohenberg and Kohn theorem, there is a one to one correspondence between $\rho(\vec{r})$ and $V_{ne}(\vec{r})$, so that the knowledge of $\rho(\vec{r})$ implies the knowledge of $V_{ne}(\vec{r})$ and thus of the ground state wavefunction $|\Psi(\vec{r})\rangle$.
As a consequence, the density functional F can be expressed as

$$F[\rho(\vec{r})] \;=\; \langle\Psi[\rho]|\,[T_e + V_{ee}]\,|\Psi[\rho]\rangle\,. \tag{6.54}$$

6.9.2 Variational principle

According to a general theorem of Quantum Mechanics, if the number n of particles is constant, the energy of a system is minimum in correspondence of the ground-state wavefunctions.
The ground state energy can thus be sought by a variational procedure. A typical approach consists in varying the one-electron wavefunctions in a self-consistent field, with the constraint that the number of particles is fixed. This approach is used in the Hartree and Hartree-Fock methods for many-electron atoms.
In the present case, we can expect that the ground-state energy is minimised by varying the density $\rho(\vec{r})$ with the constraint

$$\int \rho(\vec{r})\, d\vec{r} \;=\; n\,. \tag{6.55}$$

The density can in turn be evaluated from the occupied one-electron orbitals. The next step is thus to understand how suitable one-electron orbitals can be chosen.

6.9.3 The exchange-correlation energy

The density functional

$$F[\rho] \;=\; T_e[\rho] + V_{ee}[\rho] \tag{6.56}$$

depends on both the kinetic energy and the electron-electron potential energy, both of which are functionals of the density.
Now, the kinetic energy of a system of interacting electrons and the electron-electron interaction energy (including the effects of the Pauli exclusion principle) are very difficult to be evaluated.
It is convenient to separate, in the expression of $F[\rho]$, some terms that are easily evaluated:

1. The classical (Hartree) interaction energy between electrons

$$E_{\mathrm{Har}} \;=\; \frac{1}{2}\,\frac{e^2}{4\pi\epsilon_0}\int \frac{\rho(\vec{r})\rho(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\, d\vec{r}\, d\vec{r}\,' \tag{6.57}$$

2. The kinetic energy of a system of non-interacting electrons

$$T_{\mathrm{n.i.}}[\rho] \;=\; -\frac{\hbar^2}{2m}\sum_i^{\mathrm{occ}} \int \phi_i^*(\vec{r})\nabla^2\phi_i(\vec{r})\, d\vec{r}\,, \tag{6.58}$$

   where the $\phi_i(\vec{r})$ are single-particle orbitals, the sum is over the occupied orbitals, and the density is

$$\rho(\vec{r}) \;=\; \sum_i^{\mathrm{occ}} \phi_i^*(\vec{r})\phi_i(\vec{r}) \tag{6.59}$$

The ground-state energy can thus be expresed as

$$E[\rho] \;=\; \int V_{en}(\vec{r})\,\rho(\vec{r})\, d\vec{r} \;+\; T_{\mathrm{n.i.}}[\rho] \;+\; E_{\mathrm{Har}} \;+\; E_{\mathrm{xc}}[\rho]\,. \tag{6.60}$$

The last term $E_{\mathrm{xc}}[\rho]$ in (6.60), the exchange-correlation energy, represents our ignorance of $F[\rho]$ for the interacting system; it contains:

a) the electron-electron interaction energy beyond the classical Hartree term,
b) part of the kinetic energy of the interacting electrons, not accounted for by $T_{\mathrm{n.i.}}[\rho]$.

6.9.4 The Kohn-Sham equations

The problem of evaluating the ground-state energy (6.60) has not yet been solved. To that purpose we have to

a) choose a suitable density ρ to which the variational principle could be applied in order to minimise the ground-state energy
b) estimate the exchange-correlation unknown term E_{xc}.

The first problem (choice of the density ρ) can be solved by considering two systems:

1. the original material system of interest, with interacting electrons moving in the external potential V_{ne},
2. a fictitious system, with the same number of *non interacting* electrons, moving in an effective one-body potential $V_{\text{eff}}(\vec{r}) = V'(\vec{r})$ (whose actual form will be determined below).

The Schrödinger equation for the fictitious system is

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V'(\vec{r})\right] \phi'_i(\vec{r}) \;=\; \epsilon'_i\, \phi'(\vec{r})\,, \tag{6.61}$$

where ϕ'_i are single-electron orbitals, from which the effective density can be calculated

$$\rho'(\vec{r}) \;=\; \sum_i |\phi'(\vec{r})|^2 \tag{6.62}$$

The kinetic energy for the non interacting electrons can then be expressed as a function of the effective density $\rho'(\vec{r})$ in terms of the eigenvalues ϵ'_i

$$T_{\text{n.i.}}[\rho'] \;=\; \sum_i \epsilon'_i - \int V'(\vec{r})\, \rho'(\vec{r})\, d\vec{r}\,. \tag{6.63}$$

The ground state energy (6.60) can now be expressed as a functional of the effective density $\rho'(\vec{r})$:

$$E[\rho'] \;=\; \int V_{en}(\vec{r})\, \rho'(\vec{r})\, d\vec{r} \;+\; T_{\text{n.i.}}[\rho'] \;+\; E_{\text{Har}}[\rho'] \;+\; E_{\text{xc}}[\rho'] \tag{6.64}$$

According to the variational method, the ground state energy is obtained by varying the one-electron orbitals and the corresponding effective density in order to obtain $\delta E[\rho'] = 0$.
The condition $\delta E[\rho'] = 0$ leads to the following expression for the effective potential

$$V_{\text{eff}}(\vec{r}) \;=\; V'(\vec{r}) \;=\; V_{ne}(\vec{r}) + V_{\text{H}}(\vec{r}) + \frac{\delta E_{xc}[\rho]}{\delta \rho(\vec{r})}\,, \tag{6.65}$$

where

$$V_H(\vec{r}) \;=\; \frac{e^2}{4\pi\epsilon_0} \int \frac{\rho(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d\vec{r}\,'\,, \tag{6.66}$$

and transforms the Schrödinger equations (6.64) for the fictitious system into a set of equations (Kohn-Sham equations) where the actual potentials appear,

$$\left[\frac{p^2}{2m} + V_{ne}(\vec{r}) + V_H(\vec{r}) + V_{xc}(\vec{r})\right] \phi_i(\vec{r}) \;=\; \epsilon_i\, \phi_i(\vec{r})\,, \tag{6.67}$$

where $\phi_i(\vec{r})$ are one-electron orbitals and the exchange-correlation term $V_{xc}(\vec{r})$ is defined as the functional derivative

$$V_{xc}(\vec{r}) \;=\; \frac{\delta E_{xc}[\rho]}{\delta \rho(\vec{r})}\,. \tag{6.68}$$

According to the Kohn-Sham equations, electrons can thus be viewed as independent particles moving in the effective potential

$$V_{\text{eff}}(\vec{r}) \;=\; V_{ne}(\vec{r}) + V_{\text{H}}(\vec{r}) + V_{xc}(\vec{r})\,. \tag{6.69}$$

In principle, if E_{xc} were known, a self-consistent solution of the Kohn-Sham equations would give the electron density and the ground state energy of the interacting electron system.

Approximations for E_{xc}

The exchange-correlation energu $E_{xc}[\rho(\vec{r})]$ is generally expressed as

$$E_{xc}[\rho(\vec{r})] \;=\; \int \rho(\vec{r})\,\epsilon_{xc}(\vec{r})\,d\vec{r}\,, \tag{6.70}$$

where $\epsilon_{xc}(\vec{r})$ is an exchange-correlation energy density.
To evaluate the exchange-correlation energy is far from trivial. Relatively simple is the calculation for a homogeneous electron gas, say an electron gas embedded in a continuous and uniform distribution of positive charge. To this simple system one makes reference for the two most popular approximations

1. In the local density approximation (LDA), $\epsilon_{xc}(\vec{r})$ is assumed to be a function of $\rho(\vec{r})$
2. In the generalised gradient approximation (GGA), $\epsilon_{xc}(\vec{r})$ is assumed to be a function of $\rho(\vec{r})$ and of $\vec{\nabla}\rho(\vec{r})$

In the LDA approximation, each small volume of the system, into which the density can be considered as uniform, is assumed to contribute the same ϵ_{xc} as a homogeneous gas at the same density. The final expression of V_{xc} is

$$V_{xc}(\rho) \;\propto\; \rho^{1/3}\,. \tag{6.71}$$

Note: Only the density ρ and the ground-state energy E have physical meaning. The eigenvalues ϵ_i are simply Lagrange multipliers in the variational procedure , and don't correspond to actual excitation energies.

6.10 Bibliography of Chapter 6

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions). Chapters 9 and 10.
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapters 3 and 19 (non-crystalline solids).
- B. K. Vainshstein, V. M. Fridkin and V. L. Indenbom: *Modern Crystallography*, Springer 19812. Vol. 2, Chapter 1 (formation of atomic structure of crystals).
- C. Cohen-Tannoudji, B. Diu, F. Laloe: *Quantum Mechanics*, Wiley. Complement C_{XI}: Van der Waals forces.
- J. Zarzycki: *Les verres et l'état vitreux*, Masson 1982 (intriduction to glasses, in French).
- M. L. Cohen and S. G. Louie: *Fundamentals of Condensed Matter Physics*, Cambridge University Press 2016. Chapter 16: Density Functional Theory.
- A. D. Becke: *Perspective: Fifty years of density-functional theory in chemical physics*, J. Chem. Phys. **140**, 18A301 (2014).

Chapter 7

Vibrational dynamics

Atomic vibrations are responsible for a wide variety of physical phenomena, such as storage and transport of energy and thermal expansion; the corresponding quantities are the lattice specific heat, the lattice thermal conductivity and the coefficient of thermal expansion. Atomic vibrations affect infrared, Raman and neutron spectroscopies. The interaction between atomic vibrations and electrons accounts for other phenomena, such as the electrical resistivity of metals and the superconductivity.

Within the adiabatic approximation (Section § 6.1), the Schrödinger equation for nuclei is

$$[T_n + V(R)]\ \Phi_n \ = \ E\,\Phi_n\,, \tag{7.1}$$

where $V(R) = E_e(R)$ is the potential energy surface (PES) defined in the configuration space of the nuclear coordinates $R = \{R_1, R_2, \ldots, R_N\}$.
The static equilibrium properties of the system correspond to the minimum of the PES. The vibrational properties depend on the local curvature of the PES near the equilibirum position.
The study of the local curvature of the PES can be made at different levels of approximation: the harmonic approximation is sufficient to account for the low-temperature behaviour of the vibrational specific heat, the inclusion of anharmonicity is necessary to account for thermal expansion and for thermal resistivity.

In § 7.1, the harmonic approximation and the anharmonic behaviour are compared for a system with one degree of freedom. For many-atomic systems, only the harmonic approximation is considered in § 7.2, leading to the normal modes. In § 7.3 the case of crystals is considered and the advantages of translational symmetry are exploited. The quantum picture in terms of phonons is introduced in § 7.4. The constraints on the wavevectors of normal mode, the meaning of the first Brillouin Zone and the dispersion curves are introduced in § 7.5.

7.1 Two-atomic system (one degree of freedom)

Before considering the dynamical behaviour of many-atomic systems, such as crystals, it is convenient to review some basics on the harmonic approximation and on anharmonicity for the simple case of a system with one degree of freedom (e.g. a two-atomic molecule).

7.1.1 Power expansion of the potential energy

Let $V(r)$ be the potential energy of a system of two atoms or of two ions (Fig. 7.1, left). If only small variations of distance with respect to the position r_0 of the potential energy minimum (often referret to as "rest position") are considered, it is convenient to expand the potential energy in a Taylor series:

$$V(r) = V(r_0) + \left(\frac{dV}{dr}\right)_0 (r - r_0) + \frac{1}{2!}\left(\frac{d^2V}{dr^2}\right)_0 (r - r_0)^2 + \frac{1}{3!}\left(\frac{d^3V}{dr^3}\right)_0 (r - r_0)^3 + \cdots \tag{7.2}$$

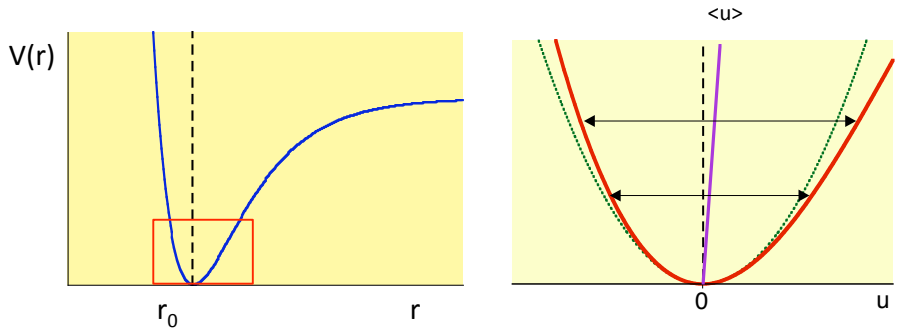

Figure 7.1: Left: realistic potential energy of interaction for a two-atomic system as a function of the interatomic distance r (the static binding energy $V(r_0)$ is not considered). Right: enlarged view as a function of the displacement $u = r - r_0$ (continuous line); the dotted line is the harmonic approximation; the dashed line slightly inclined to the right corresponds to the rest distance, the continuous nearly vertical line corresponds to the average distance (thermal expansion).

- The first term $V(r_0)$ is the static binding energy.
- The second term is zero, because the first derivative of the potential energy is zero at the minimum position.
- The harmonic approximation consists in considering the expansion only up to the second-order term.
- Higher order terms are called anharmonic terms.

For our present purposes, it is convenient to set $V(r_0) = 0$ and rewrite (7.2) in terms of the ionic displacements with respect to the rest position $u = r - r_0$ (Fig. 7.1, right), introducing the force constants k_i:

$$V(u) = \frac{1}{2}\, k_0\, u^2 + k_3\, u^3 + k_4\, u^4 + \cdots \tag{7.3}$$

Here k_0 is the harmonic (second order) force constant, $k_i (i > 2)$ are the anharmonic force constants. In general, the third-order force constant is negative, $k_3 < 0$, corresponding to a repulsive branch of the potential energy steeper than the attractive branch.

7.1.2 The harmonic approximation

Although no realistic potential energy is truly harmonic, the harmonic approximation $V(u) = k_0 u^2/2$ is very important, for a number of reasons.

- The Schrödinger equation of the harmonic oscillator can be exactly solved, leading to the quantisation of energy levels $E_n = (n + 1/2)\hbar\omega_0$, where $\omega_0 = \sqrt{k_0/\mu}$.
- The harmonic approximation allows a good reproduction of the main features of several important thermal properties of matter, such as specific heats.
- The harmonic approximation is a good starting point for treating anharmonic effects (quasi-harmonic approximation, perturbative approaches).

Classical harmonic oscillator

The dynamical state of a classical harmonic oscillator is defined by the values of position u and momentum $p = m\, dv/dt$, which can be both derived from the equation

$$u(t) = A\cos(\omega_0 t) + \phi\,, \quad \text{where} \quad \omega_0 = \sqrt{k_0/\mu}\,. \tag{7.4}$$

The quantity ω_0 is independent of the amplitude A and corresponds to the angular frequency of oscillations.

The total energy is the Hamiltonian

$$H = E = \frac{p^2}{2\mu} + \frac{1}{2}\mu\omega_0^2 u^2 = \frac{1}{2}\mu\omega_0^2 A^2 \tag{7.5}$$

where μ is the reduced mass of the two-atomic system.
From (7.4) and (7.5), one can easily calculate the average values

$$\langle u \rangle = 0\,, \qquad \langle u^2 \rangle = \frac{A^2}{2} = E/\mu\omega_0^2\,. \tag{7.6}$$

The average distance is equal to the rest distance:

$$\langle r \rangle = r_0 + \langle u \rangle = r_0\,. \tag{7.7}$$

Quantum harmonic oscillator

The dynamical state of a quantum harmonic oscillator (§ 7.6.1) is characterised by the eigenvector $|n\rangle$ of the Hamiltonian H_0, which is contemporarily eigenvector of the number operator $N = a^\dagger a$ The dynamical state is thus characterised by the number n of energy quanta $\hbar\omega_0$. The total energy of state $|n\rangle$ is

$$E_n = (1/2 + n)\,\hbar\omega_0\,. \tag{7.8}$$

If the harmonic oscillator is in a stationary state $|n\rangle$, there is no information on the position $u(t)$. One can anyway calculate the average values of u and u^2; one finds (see § 7.6.1) that, as for the classical approximation,

$$\langle u \rangle = 0\,, \qquad \langle u^2 \rangle = E_n/\mu\omega_0^2\,. \tag{7.9}$$

Note 1: For the quantum oscillator in a stationary state $|n\rangle$, it makes no sense to consider ω_0 as an oscillation frequency, since one doesn't know the law $u(t)$.

Note 2: It is worth remembering that the true dissociation energy D of a two-atomic molecule is determined by the joint effect of the static binding energy and of the zero-point energy:

$$D = -V(r_0) - \hbar\omega_0/2 = |V(r_0)| - \hbar\omega_0/2\,. \tag{7.10}$$

Classical statistics of the harmonic oscillator

Let us now consider a canonical ensemble of harmonic oscillators at a temperature $T = 1/k_B\beta$, and evaluate the average value $\langle u \rangle$ and the mean square displacement $\langle u^2 \rangle$, first in the classical approximation, then for the quantum case.

In the *classical approximation*, the average energy of a harmonic oscillator is connected to the temperature by the equipartition theorem, and is thus proportional to T:

$$\langle E \rangle \;=\; k_B T \;=\; 1/\beta\,. \tag{7.11}$$

The average displacement from the equilibrium position is (see § 7.6.3)

$$\langle u \rangle \;=\; \frac{\int u \exp[-\beta V(u)]\; du}{\int \exp[-\beta V(u)]\; du} \;=\; 0\,, \tag{7.12}$$

where $V(u) = k_0 u^2/2$. There is no thermal expansion.
The mean square displacement is

$$\sigma^2 \;=\; \langle u^2 \rangle \;=\; \frac{\int u^2 \exp[-\beta V(u)]\; du}{\int \exp[-\beta V(u)]\; du} \;=\; \frac{k_B T}{k_0} \;=\; \frac{\langle E \rangle}{\mu\omega_0^2}\,, \tag{7.13}$$

proportional to the temperature (and to the energy) and inversely proportional to the force constant k_0 (Fig.7.2, right panel, dashed line).

Quantum statistics of the harmonic oscillator

The average energy of the harmonic oscillator in the energy representation is (see § 7.6.3)

$$\langle H_0 \rangle \;=\; \mathrm{Tr}(H_0 w) \;=\; \frac{1}{Z_0}\,\mathrm{Tr}\left(H_0\, e^{-\beta H_0}\right) \;=\; \hbar\omega_0 \left[\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega_0}-1}\right]. \tag{7.14}$$

where w is the statistical density operator

$$w \;=\; \frac{1}{Z_0}\, e^{-\beta H_0} \tag{7.15}$$

and Z_0 is the partition function for the harmonic oscillator

$$Z_0 \;=\; \mathrm{Tr}\,(e^{-\beta H_0}) \;=\; \frac{e^{-\beta\hbar\omega_0/2}}{1-e^{-\beta\hbar\omega_0}}. \tag{7.16}$$

The probability density for the displacement u has a gaussian shape centred on $\langle u \rangle = 0$:

$$\rho(u) \;=\; (1/\sigma\sqrt{2\pi})\, e^{-u^2/2\sigma^2}\,. \tag{7.17}$$

The moments of the distribution can be calculated as (see § 7.6.3)

$$\langle u^k \rangle \;=\; \frac{1}{Z_0}\,\mathrm{Tr}\,(u^k e^{-\beta H_0}) \;=\; \frac{1}{Z_0}\sum_{n=0}^{\infty} \langle n|u^k|n\rangle\, e^{-\beta E_n}. \tag{7.18}$$

One can see that

$$\langle u \rangle \;=\; 0\,, \tag{7.19}$$

(there is no thermal expansion) and that the mean square displacement is

$$\sigma^2 \;=\; \langle u^2 \rangle \;=\; \frac{\hbar}{\mu\omega_0}\left[\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega_0}-1}\right] \;=\; \frac{\langle E \rangle}{\mu\omega_0^2}\,. \tag{7.20}$$

For $T \to \infty$, σ^2 tends to the classical limit $k_B T/k_0$ (Fig.7.2, right panel, continuous line). For $T \to 0$, instead, the quantum behaviour is at variance with the classical behaviour:

$$\sigma^2 \to \sigma_0^2 = \hbar/2\sqrt{k_0\mu} = \hbar/2\mu\omega_0 \tag{7.21}$$

7.1.3 Anharmonicity effects

Approximate classical approach

An introductory idea of the main effects of anharmonicity can be gained by considering a single oscillator with a weak third-order term in the potential energy

$$V(u) = \frac{1}{2}\, k_0\, u^2 + k_3\, u^3\,, \tag{7.22}$$

where $k_3 < 0$, and looking for an approximate classical solution of the equation of motion.

Note: A purely cubic potential (7.22) would lead to system instability for large enough oscillations; a suitable fourth-order term is necessary to guarantee stability. We consider here only the cubic term for simplicity, with the basic assumption that oscillations are sufficiently weak to guarantee the stability.

The classical acceleration is

$$\ddot{u} \;=\; \frac{F}{\mu} \;=\; -\frac{k_0}{\mu}\, u - \frac{3k_3}{\mu}\, u^2 \;=\; -\omega_0^2\, u - s\,\omega_0^2\, u^2 \tag{7.23}$$

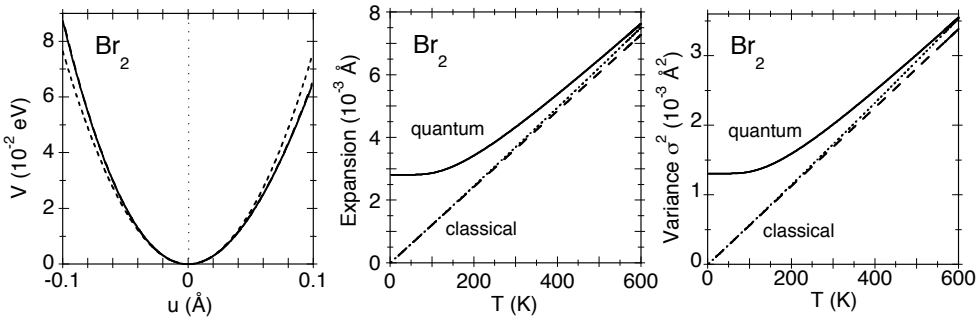

Figure 7.2: Bromine molecule, thermodynamic quantities calculated from the force constants of an anaharmonic potential, determined by spectroscopic investigations (k_0 = 15.28 eV/Å^2, k_3 = -10.96 eV/Å^3, k_4 = 6.6 eV/Å^4). Left: potential energy: only harmonic approximation (dashed line) and including the third order term (continuous line); the contribution of the fourth order term cannot be appreciated. Centre and right: thermal expansion $\langle u \rangle$ and variance $\langle u^2 \rangle$; dashed line: first-order classical approximation; dotted line: second-order classical approximation; continuous line: first-order quantum approximation.

where $s = 3k_3/k_0$. The relevance of the cubic term of the potential energy increases when the amplitude A of the vibrations increases. The assumed weakness of the cubic term corresponds to the condition $sA \ll 1$.
The simplest correction to the harmonic oscillator solution is of the form

$$u(t) = u_0 + A\left[\cos(\omega t) + \eta \cos(2\omega t)\right] , \tag{7.24}$$

where the oscillation centre u_0 and the frequency ω are a-priori unknown and the anharmonicity is supposed to introduce a weak oscillating perturbation ($\eta \ll 1$) with frequency 2ω. By inserting (7.24) into (7.23), taking into account that $\eta \ll 1$ and $sA \ll 1$ and considering only the leading terms terms, one gets two important results:

1. The centre of oscillation u_0 depends on the amplitude A according to

$$u_0 = -\frac{3k_3}{2k_0} A^2 > 0 \tag{7.25}$$

When the amplitude A increases, the average distance increases (k_3 is negative), say there is a bond expansion:

$$\langle r \rangle = r_0 + \langle u \rangle = r_0 + u_0 = r_0 - \frac{3k_3}{2k_0} A^2 \tag{7.26}$$

2. The frequency of oscillations decreases when the amplitude A increases, according to

$$\omega^2 = \omega_0^2 (1 - s^2 A^2) \tag{7.27}$$

By substituting A^2 from (7.25) in (7.27) one finds

$$\omega^2 = \omega_0^2 \left(1 + \frac{6k_3}{k_0}\, u_0\right) \tag{7.28}$$

To first approximation, thus, the decrease of frequency is proportional to the thermal expansion u_0. In general, the relation between thermal expansion and frequency is not linear. For some applications, one can consider the oscillator as if it where harmonic, the only effect of anharmonicity being the variation of frequency (quasi-harmonic approximation).

Approximate quantum approach

The quantum approach is based on considering the anharmonic terms of the potential energy as a perturbation with respect to the harmonic oscillator. Let us consider again only the third-order anharmonic term.
Making use of the formalism of creation and annihilation operators, one can show (see § 7.6.2) that

1. The energy levels are down-shifted with respect to the harmonic levels, the extent of the shift increasing with the quantum number n. This fact corresponds to a reduction of the frequency of the fundamental absorption or emission line of the harmonic oscillator, and to the appearance of less intense satellite lines.
2. The eigenfunctions of the level E_n are now a linear combination of the harmonic oscillator eigenfunctions $|n\rangle, |n+1\rangle, |n-1\rangle, |n+3\rangle, |n-3\rangle$.
   As a consequence, the position and momentum operators u and p can have non-zero matrix elements for eigenstates of the Hamiltonian, which are superpositions of harmonic eigenstates with different quantum numbers.
   In particular, one can show that $\langle u \rangle \neq 0$.

Classical statistics of the anharmonic oscillator

For a canonical ensemble of classical anharmonic oscillators, the average values of $\langle u \rangle$ and $\langle u^2 \rangle$ are calculated trhough (7.12) and (7.13), respectively, where now the potential energy $V(u)$ includes the anharmonic terms (7.3). One finds that the average displacement is

$$\langle u \rangle \;=\; -\,\frac{3k_3}{k_0^2}\,k_B T \;-\; \frac{3k_3}{k_0^4}\left(\frac{45k_3^2}{k_0} - 32k_4\right)(k_B T)^2 \;+\; \ldots \tag{7.29}$$

To first order, $\langle u \rangle$ is positive ($k_3 < 0$): there is a positive thermal expansion which linearly depends on temperature (Fig.7.2, central panel, dashed line). The second-order correction is proportional to T^2, and is generally quite small (Fig.7.2, central panel, dotted line).
The mean square displacement is

$$\langle u^2 \rangle \;=\; \frac{k_B T}{k_0}\left[1 + \frac{3}{k_0^2}\left(\frac{12k_3^2}{k_0} - 4k_4\right)(k_B T) \;+ \ldots\right] \tag{7.30}$$

To first order, one recovers the same expression as for the harmonic oscillator,

$$\langle u^2 \rangle \;=\; \langle A^2 \rangle / 2 \;=\; k_B T / k_0\,, \tag{7.31}$$

(see Fig.7.2, central panel, dashed line). The frequency is again $\omega_0 = (k_0/\mu)^{1/2}$.
The second order correction is proportional to T^2 (Fig.7.2, central panel, dotted line) and is generally positive. Neglecting the contribution of k_4, the second order correction gives a softening of the force constant with respect to the harmonic value k_0, corresponding to a progressive reduction of the frequency when T increases:

$$\omega_0^2 \;\rightarrow\; \omega^2 = \omega_0^2\left[1 - \left(\frac{36k_3^2}{k_0^3}\right)k_B T\right] \tag{7.32}$$

Two important consequences of the third-order term in the potential energy expansion are thus the following:

1. There is thermal expansion, positive if $k_3 < 0$. Within the classical approximation, to first order the thermal expansion is proportional to T.
2. The frequency of oscillations decreases when the temperature increases.

Quantum statistics of the anharmonic oscillator

The hamiltonian operator for the anharmonic oscillator can be expressed as the sum

$$H \;=\; H_0 + H_1\,, \tag{7.33}$$

where

$H_0 = p^2/2\mu + k_0 u^2/2$ is the harmonic Hamiltonian
$H_1 = k_3 u^3 + k_4 u^4 + \ldots$ is the contribution of anharmonicity.

In the statistical density operator, H_0 is substituted by H: the factor $\exp(-\beta H_0)$ is substituted by $\exp(-\beta H)$

$$w \;=\; \frac{1}{Z}\, e^{-\beta H}\,, \qquad Z \;=\; \mathrm{Tr}(e^{-\beta H}) \tag{7.34}$$

and eq. (7.18) for the moments $\langle u^k \rangle$ becomes

$$\langle u^k \rangle \;=\; \frac{1}{Z}\,\mathrm{Tr}\,(u^k e^{-\beta H})\,. \tag{7.35}$$

One can easily see that the un-normalized statistical density operator $\tilde{w} = e^{-\beta H}$ obeys the differential equation

$$\frac{\partial}{\partial \beta}\, e^{-\beta H} \;=\; -H\, e^{-\beta H}\,. \tag{7.36}$$

The initial condition is that for $\beta = 0$ (say for $T = \infty$) $\tilde{w} = e^{-\beta H} = 1$, corresponding to a uniform distribution over all the available energy states.

Note: It is worth noting that (7.36) is formally similar to the time-dependent Schrödinger equation, provided the substitution $\beta \to it/\hbar$ is made.

Eq. (7.36) can be exactly solved for the harmonic Hamiltonian. The statistical quantum approach to anharmonicity is based on considering the anharmonic part H_1 as a small perturbation with respect to the unperturbed harmonic Hamiltonian H_0, whose eigenvalues E_n and eigenvectors $|n\rangle$ can be exactly calculated.
Let us consider the derivative

$$\frac{\partial}{\partial \beta}\,\left[e^{-\beta H}\, e^{\beta H_0}\right] \;=\; -\, e^{\beta H_0}\, H_1\, e^{-\beta H} \tag{7.37}$$

Integrating from $\beta' = 0$ to $\beta' = \beta$ and taking into account the initial condition for $\beta = 0$, one finds that, within the perturbative approach, $\exp(-\beta H)$ is related to $\exp(-\beta H_0)$ by the integral equation

$$e^{-\beta H} \;=\; e^{-\beta H_0}\left[1 - \int_0^\beta e^{\beta' H_0}\, H_1\, e^{-\beta' H}\, d\beta'\right]. \tag{7.38}$$

The first order solution is obtained by imposing $H = H_0$ in the last factor within the integral,

$$e^{-\beta H} \;\simeq\; e^{-\beta H_0}\left[1 - \int_0^\beta e^{\beta' H_0}\, H_1\, e^{-\beta' H_0}\, d\beta'\right]. \tag{7.39}$$

The factor $\exp(-\beta H)$ obtained from the solution to (7.39) can be inserted into (7.38) to obtain a second-order approximation; and so on.
The first-order approximation gives

$$\mathrm{Tr}\left[u^k\, e^{-\beta H}\right] \;\simeq\; \mathrm{Tr}\left[u^k\, e^{-\beta H_0}\,\left(1 - \int_0^\beta e^{\beta' H_0}\, H_1\, e^{-\beta' H_0}\, d\beta'\right)\right] \tag{7.40}$$

and

$$\frac{1}{Z} \;\simeq\; \frac{1}{Z_0}\left[1 + \frac{1}{Z_0}\,\mathrm{Tr}\left(e^{-\beta H_0}\int_0^\beta e^{\beta' H_0}\, H_1\, e^{-\beta' H_0}\, d\beta'\right)\right] \tag{7.41}$$

By calculating the integrals using E_n and $|n\rangle$ from H_0, one obtains

$$\begin{aligned}\langle u^k\rangle \;\simeq\; & \frac{1}{Z_0}\sum_n e^{-\beta E_n}\,\langle n|u^k|n\rangle \\ & + \frac{1}{Z_0}\sum_{n,n'}\frac{e^{-\beta E_n}-e^{-\beta E_{n'}}}{E_n-E_{n'}}\,\langle n|u^k|n'\rangle\langle n'|H_1|n\rangle \\ & + \frac{\beta}{Z_0^2}\sum_n e^{-\beta E_n}\,\langle n|u^k|n\rangle\sum_{n'} e^{-\beta E_{n'}}\,\langle n'|H_1|n'\rangle\end{aligned} \tag{7.42}$$

where for $n = n'$ the 0/0 factor in the second term should be replaced by

$$\frac{e^{-\beta E_n}-e^{-\beta E_{n'}}}{E_n-E_{n'}} \;=\; -\beta\, e^{-\beta E_n}\,. \tag{7.43}$$

To first order, one finds that the thermal expansion expansion is

$$\langle u\rangle \;\simeq\; -\frac{3k_3}{k_0}\,\frac{\hbar}{\mu\omega_0}\left[\frac{1}{2}+\frac{1}{\exp(\hbar\omega_0/k_BT)-1}\right] \;=\; -\frac{3k_3}{k_0}\,\sigma_0^2\,. \tag{7.44}$$

where σ_0^2 is the variance of the harmonic oscillator; the difference with respect to the classical behaviour is evident in Fig. 7.2 (central panel, continuous line).

As for the variance $\langle u^2\rangle$, to first order one finds again the variance of the harmonic oscillator (Fig. 7.2, right panel, continuous line):

$$\sigma^2 \;=\; \langle u^2\rangle \;\simeq\; \sigma_0^2 \;=\; \frac{\hbar}{\mu\omega_0}\left[\frac{1}{2}+\frac{1}{e^{\beta\hbar\omega_0}-1}\right] \;=\; \frac{\langle E_0\rangle}{\mu\omega_0^2}\,. \tag{7.45}$$

Note: It is worth to note the formal similarity between (7.38) and the integral equation for the time evolution operator

$$U(t_f,t_i) \;=\; U^0(t_f,t_i) \;+\; \frac{1}{i\hbar}\int_{t_i}^{t_f} U^0(t_f,t)\,H_{\rm int}\,U(t,t_i)\,dt \tag{7.46}$$

which is at the base of the time-dependent perturbation theory; here U_0 is the evolution operator of a stationary state

$$U^0(t_f,t_i) = e^{-iH_0(t_f-t_i)/\hbar}\,, \qquad U^0(t_f,t) = e^{-iH_0(t_f-t)/\hbar} \tag{7.47}$$

and the initial condition is $U(t_i,t_i) = 1$.

7.1.4 Final remarks

In both the classical approximation and the quantum treatment, one finds that the inclusion of anharmonic terms

a) produces only slight variations to the energy and to the variance of the distribution; as a consequence, the harmonic approximation is very effective in predicting the specific heats and the mean square atomic displacements;
b) is necessary to account for thermal expansion.

A note on nomenclature

It is convenient, in view of the applications to many-atomic systems, to distinguish the three different inter-atomic distances introduced above.

1. The rest distance r_0 corresponds to the minimum of the static potential energy $V(r)$; it is a constant, independent of temperature.
   The expression "rest distance" is conventional; it refers to a classical distance for $T = 0$.

2. The instantaneous distance $r(t)$ is variable with time.
3. The average distance is $\langle r(t)\rangle = r_0 + \langle u(t)\rangle$.
   - In the harmonic approximation, $\langle u\rangle = 0$ and $\langle r(t)\rangle = r_0$, the average distance coincides with the rest distance.
   - As an effect of anharmonicity, the average distance is different from the rest distance and depends on temperature.

7.2 Many-atomic systems

Let us now consider a system composed of $\mathcal{N}$ atoms: a molecule, a nano-cluster, a crystal, a non-crystalline solid. In this Section 7.2 we neglect the possible presence of translational symmetry, which is taken into account in Section 7.3.
Let the atoms of the system be labeled by the integer number m and let $\vec{r}(m)$ be te position of atom m. The potential energy surface (PES) $V[\{\vec{r}(m)\}]$, introduced in Section 6.1, is defined in the $3\mathcal{N}$-dimensional configuration space ($m = 1, 2, \ldots, \mathcal{N}$).
Our goal is here to study the oscillations of the system around the equilibrium configuration (minimum of the PES) due to zero-point energy and to thermal excitations at finite temperature. These oscillations of the whole system in the configuration space are connected to the oscillations of single atoms around their equilibrium positions in the real space. The instantaneous position $\vec{r}(m)$ of atom m is connected to the equilibrium position $\{\vec{R}(m)\}$ by

$$\vec{r}(m) \;=\; \vec{R}(m) + \vec{u}(m)\,, \tag{7.48}$$

where $\vec{u}(m)$ is the instantaneous displacement.

7.2.1 Expansion of the potential energy

If the atomic displacements $\vec{u}(m)$ are small, the potential energy of $V[\{\vec{r}(m)\}]$ of the system can be expanded as a power series of the atomic displacements about their average (equilibirum) positions (Born-von Karman expansion):

$$V\;[\{\vec{r}(m)\}] \;=\; V_0 + V_1 + V_2 + V_3 + \cdots \tag{7.49}$$

The **zero-order term**

$$V_0 \;=\; V\left[\{\vec{R}(m)\}\right] \tag{7.50}$$

is the potential energy corresponding to the average configuration $\{\vec{R}(m)\}$ of the system.
The **first-order term** is

$$V_1 \;=\; \sum_{\alpha m} \Phi(\alpha m)\, u_\alpha(m)\,, \tag{7.51}$$

where $\alpha = x, y, z$ labels the cartesian coordinates and the quantities

$$\Phi(\alpha m) \;=\; \left[\frac{\partial V}{\partial u_\alpha(m)}\right]_{\{\vec{R}\}} \tag{7.52}$$

are the first-order coupling parameters. Here and below, the derivatives are calculated at the equilibrium configuration $\{\vec{R}(m)\}$.
The **second-order term** is

$$V_2 \;=\; \frac{1}{2} \sum_{\substack{\alpha m \\ \alpha' m'}} \Phi\begin{pmatrix} \alpha\, m \\ \alpha' m' \end{pmatrix}\; u_\alpha(m)\, u_{\alpha'}(m') \tag{7.53}$$

where the quantities

$$\Phi\begin{pmatrix} \alpha\, m \\ \alpha' m' \end{pmatrix} = \left(\frac{\partial^2 V}{\partial u_\alpha(m)\, \partial u_{\alpha'}(m')}\right)_{\{\vec{R}\}} \tag{7.54}$$

are the second-order coupling parameters.
As a consequence of the invariance of second partial derivatives with respect to the order of derivation, the second-order coupling parameters are invariant with respect to the exchange of the indices:

$$\Phi \begin{pmatrix} \alpha\, m \\ \alpha' m' \end{pmatrix} = \Phi \begin{pmatrix} \alpha' m' \\ \alpha\, m \end{pmatrix} \tag{7.55}$$

The **third-order term** is

$$V_3 \;=\; \frac{1}{3!} \sum_{\substack{\alpha m \\ \alpha' m' \\ \alpha'' m''}} \Phi \begin{pmatrix} \alpha\, m \\ \alpha'\, m' \\ \alpha'' m'' \end{pmatrix} u_\alpha(m)\, u_{\alpha'}(m')\, u_{\alpha''}(m'')\,, \tag{7.56}$$

where

$$\Phi \begin{pmatrix} \alpha\, m \\ \alpha'\, m' \\ \alpha'' m'' \end{pmatrix} \;=\; \left(\frac{\partial^3 V}{\partial u_\alpha(m)\, \partial u_{\alpha'}(m')\, \partial u_{\alpha''}(m'')} \right)_{\{\vec{R}\}} \tag{7.57}$$

are the third-order coupling parameters.

And so on with higher order terms.

7.2.2 Harmonic approximation

The harmonic approximation consists in truncating the Taylor expansion (7.49) at the 2nd-order term V_2 (quadratic in the atomic displacements).
In the harmonic approximation, the average (or equilibirum) atomic positions $\{\vec{R}(m)\}$, around which atoms vibrate, are equal to rest positions $\{\vec{R}^0(m)\}$, which correspond to the minimum of the potential energy surface. As a consequence, the first order term (7.51) is zero, $V_1 = 0$.
For an anharmonic system, the average positions $\vec{R}(m)$ are different from the rest positions $\{\vec{R}^0(m)\}$ and are *a priori* unknown.

In this Chapter 7, we restrict our attention to the harmonic approximation for the vibrational dynamics of solids, beginning with the general case in this Section 7.2 and then considering the crystal case in Section 7.3. Some anharmonicity effects are considered in Chapter 8.
As usual, we start from the classical approach and introduce the normal modes of vibrations. The quantisation of normal modes is made as a second step, and is here considered for the case of crystals Section 7.4.

Note: The harmonic approximation here considered refers to the Taylor expansion of the potential energy of the whole many-atomic system. It doesn't imply that the pair interactions between single atoms can be considered harmonic.

7.2.3 Dynamical equations and harmonic coupling parameters

Hamiltonian function

The kinetic energy of the system is

$$T \;=\; \frac{1}{2} \sum_{\alpha m} M_m\, \dot{u}_\alpha^2(m)\,, \tag{7.58}$$

where M_m is the mass of atom m, so that the harmonic Hamiltonian of the system is

$$\mathcal{H} \;=\; V_0 \;+\; \frac{1}{2} \sum_{\alpha m} M_m\, \dot{u}_\alpha^2(m) \;+\; \frac{1}{2} \sum_{\substack{\alpha m \\ \alpha' m'}} \Phi \begin{pmatrix} \alpha\, m \\ \alpha' m' \end{pmatrix} u_\alpha(m)\, u_{\alpha'}(m')\,. \tag{7.59}$$

Forces

The force acting on atom m in direction α is

$$F_\alpha(m) = -\frac{\partial V_2}{\partial u_\alpha(m)} = -\sum_{\alpha'\ m'} \Phi\begin{pmatrix}\alpha\, m\\ \alpha' m'\end{pmatrix} u_{\alpha'}(m')\,. \tag{7.60}$$

Equation (7.60) expresses the force acting on atom m as the sum of the contributions from the interactions with all the other atoms.

Note: The expression (7.53) contains both quadratic terms , such as

$$\frac{1}{2}\sum_{\substack{\alpha m\\ \alpha m}} \Phi\begin{pmatrix}\alpha m\\ \alpha m\end{pmatrix} u_\alpha^2(m)\,, \qquad \frac{1}{2}\sum_{\substack{\alpha' m'\\ \alpha' m'}} \Phi\begin{pmatrix}\alpha' m'\\ \alpha' m'\end{pmatrix} u_{\alpha'}^2(m')$$

and cross-product terms, such as

$$\frac{1}{2}\sum_{\substack{\alpha m\\ \alpha' m'}} \Phi\begin{pmatrix}\alpha\, m\\ \alpha' m'\end{pmatrix} u_\alpha(m)\, u_{\alpha'}(m')\,, \qquad \frac{1}{2}\sum_{\substack{\alpha' m'\\ \alpha m}} \Phi\begin{pmatrix}\alpha' m'\\ \alpha\, m\end{pmatrix} u_{\alpha'}(m')\, u_\alpha(m)\,.$$

Because of the invariance (7.55) with respect to the exchange of indices, the cross terms are two by two equal, so that the factors 1/2 disappear in (7.60).

Meaning of coupling parameters

To grasp the physical meaning of the harmonic coupling parameters, let us choose a central atom m at rest and assume that all other atoms are at rest, with the exception of only one atom (m'), which is displaced by $u_{\alpha'}(m')$ along the α' direction. From (7.60), the α component of the force on m is

$$F_\alpha(m) = -\Phi\begin{pmatrix}\alpha\, m\\ \alpha' m'\end{pmatrix} u_{\alpha'}(m')\,. \tag{7.61}$$

Each coupling parameter $\Phi\left(\begin{smallmatrix}\alpha\, m\\ \alpha' m'\end{smallmatrix}\right)$ is thus the opposite of the force exerted on atom m in direction α when the atom m' is displaced a unit distance in direction α', all other atoms remaining fixed. According to (7.60), the force acting on a given atom depends on the contributions of all the atomic displacements, including the displacement of the atom m itself.

Note: In principle, the coupling parameters connect each atom m with all other atoms m' of the crystal. Actually, in many cases only the coupling parameters connecting atoms at relatively short distances are relevant.

Problem: Consider a linear chain of atoms connected by springs of elastic constant k. Evaluate the coupling parameters $\Phi(\ell, \ell')$, for $\ell' = \ell, \ell-1, \ell+1, \ell-2, \ell+2$. Express the corresponding contribution (7.53) to the potential energy. Calculate the contribution $F(\ell)$ to the force on atom ℓ due to the displacement of: atom ℓ itself, atoms $\ell-1$, $\ell+1$, $\ell-2$, $\ell+2$.

Constraint of the coupling parameters

Let us consider a rigid movement of the whole system, where all atoms undergo the same displacement $\vec{u}$ of components $u_{\alpha'}$. The force on the generic atom m induced by the equal displacements of all the atoms has to be zero:

$$0 \;=\; F_\alpha(m) = -\sum_{\alpha' m'} \Phi\begin{pmatrix}\alpha\, m\\ \alpha' m'\end{pmatrix} u_{\alpha'}(m') = -\sum_{\alpha'} u_{\alpha'} \sum_{m'} \Phi\begin{pmatrix}\alpha\, m\\ \alpha' m'\end{pmatrix} \tag{7.62}$$

Since the displacement $u_{\alpha'}$ is arbitrary, (7.62) gives

$$\sum_{m'} \Phi\begin{pmatrix}\alpha\, m\\ \alpha' m'\end{pmatrix} = 0\,. \tag{7.63}$$

From this equality, we can obtain, for the coupling parameters where $m' = m$:

$$\Phi \begin{pmatrix} \alpha\, m \\ \alpha' m \end{pmatrix} = - \sum_{m' \neq m} \Phi \begin{pmatrix} \alpha\, m \\ \alpha' m' \end{pmatrix} . \tag{7.64}$$

Physically, (7.64) means that displacing the atom m by a unit distance in a given direction, all other atoms staying at rest, has the same effect as displacing all the other atoms by an opposite unit distance, the atom m staying at rest.

Problem: Verify (7.64) for the case of the linear chain of the previous problem.

Dynamical equations

By solving the Hamilton equations, one obtains that the classical dynamical equation for the α-th component of the displacement of atom m (of mass M_m) is

$$\boxed{M_m\, \ddot{u}_\alpha(m) = - \sum_{\alpha' m'} \Phi \begin{pmatrix} \alpha\, m \\ \alpha' m' \end{pmatrix} u_{\alpha'}(m')} \tag{7.65}$$

There are $3\mathcal{N}$ such dynamical equations, three for each atom m. Globally, they form a system of $3\mathcal{N}$ coupled equations corresponding to $3\mathcal{N}$ coupled harmonic oscillators.

Solutions of the dynamical equations

The system of equations (7.65) describes the behaviour of $3\mathcal{N}$ coupled harmonic oscillators. It is thus reasonable to seek solutions of the oscillatory type

$$\boxed{u_\alpha(m,t) \;=\; \frac{1}{\sqrt{M_m}}\, \mathbf{Re}\left[\, w_\alpha(m)\, e^{-i\omega t}\right]} \tag{7.66}$$

The amplitudes $w_\alpha(m)$ are complex quantities, to account for the possible phase relationships between the displacements of different atoms, and depend both on the dynamical equation and on the initial conditions.
In (7.66), the mass M_m appears explicitly, for symmetry reasons that will be made clear below.
By substituting the values $u_\alpha(m,t)$ of (7.66) into the dynamical equation (7.65) one gets

$$-\sqrt{M_m}\, \omega^2\, w_\alpha(m)\, e^{-i\omega t} = - \sum_{\alpha' m'} \Phi \begin{pmatrix} \alpha\ m \\ \alpha'\ m' \end{pmatrix} \frac{1}{\sqrt{M_{m'}}}\, w_{\alpha'}(m')\, e^{-i\omega t}. \tag{7.67}$$

7.2.4 Interaction matrix and normal modes

By introducing the mass-adjusted coupling parameters

$$D^0 \begin{pmatrix} \alpha\ m \\ \alpha'\ m' \end{pmatrix} \;=\; \frac{\Phi \left(\begin{smallmatrix} \alpha\ m \\ \alpha'\ m' \end{smallmatrix} \right)}{\sqrt{M_m\, M_{m'}}}\,, \tag{7.68}$$

equations (7.67) can be rewritten as

$$\omega^2\, w_\alpha(m) \;-\; \sum_{\alpha' m'} D^0 \begin{pmatrix} \alpha\ m \\ \alpha'\ m' \end{pmatrix} w_{\alpha'}(m') \;=\; 0 \tag{7.69}$$

Equation (7.69) is representative of a system of $3\mathcal{N}$ coupled equations.
By considering

- the terms $w_\alpha(m)$ as elements of a $3\mathcal{N}$-size column vector U,
- the terms $D^0 \left(\begin{smallmatrix} \alpha\ m \\ \alpha'\ m' \end{smallmatrix} \right)$ as elements of a $3\mathcal{N} \times 3\mathcal{N}$ matrix, the mass-adjusted interaction matrix (or real-space dynamical matrix) $\mathbf{D^0}$,

the system of dynamical equations (7.69) can be synthesised as a matrix equation:

$$\boxed{\omega^2\, U \;=\; \mathbf{D^0}\, U} \tag{7.70}$$

or more explicitly

$$\omega^2 \begin{pmatrix} \cdots \\ w_\alpha(m) \\ \cdots \\ \cdots \end{pmatrix} = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & D^0 \left(\begin{smallmatrix} \alpha & m \\ \alpha' & m' \end{smallmatrix}\right) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \cdots \\ \cdots \\ w_{\alpha'}(m') \\ \cdots \end{pmatrix} \tag{7.71}$$

Eigenfrequencies and eigenvectors

The solution of (7.71) is a typical eigenvalue problem. By imposing that the secular determinant is zero, $\det[\mathbf{D^0} - \omega^2 \mathbf{I}] = 0$, one obtains an algebraic equation, whose solution gives in principle

→ $3\mathcal{N}$ eigenfrequencies ω_j $(j = 1, 2, \ldots, 3\mathcal{N})$, corresponding to $3\mathcal{N}$ independent normal modes. The number of normal modes is three times the number of atoms in the system.

The invariance of the force constants with respect to the exchange of indices (7.55) ensures that the real-space dynamical matrix $\mathbf{D}^0$ is symmetric, and the eigenvalues are real. The physical meaning requires that the eigenvalues ω^2 are non-negative quantities.

For each eigenfrequency ω_j, say for each normal mode, (7.71) gives an eigenvector U_j with $3\mathcal{N}$ components describing the relative amplitudes and the phase relationships of the atomic motions corresponding to the normal mode.
In total there are thus

→ $3\mathcal{N}$ eigenvectors U_j, each one with $3\mathcal{N}$ components.

By this treatment, the coupled motions of $\mathcal{N}$ atoms are described in terms of $3\mathcal{N}$ independent harmonic oscillators.
The practical solution of (7.71) is difficult for a large number $\mathcal{N}$ of atoms, e.g. for large molecules or nanoclusters. It represents an impossible task for a crystal: a dramatic improvement is however obtained if one takes into account the translational symmetry of the crystal.

Degrees of freedom and vibrational modes

A three-dimensional system of $\mathcal{N}$ atoms has $3\mathcal{N}$ degrees of freedom. Of these:

a) 3 are related to translation of the whole system
b) 3 are related to rotations of the whole system (only 2 for linear molecules)
c) the remaining $3\mathcal{N} - 6$ are related to vibrations ($3\mathcal{N} - 5$ for linear molecules)

For free molecules, the normal mode analysis leads to zero frequency for the modes corresponding to translation of rotation of the whole molecule.

Example 1: Water molecule H_2O.
3 atoms, 9 degrees of freedom, non-linear molecule.
3 translational degrees of freedom and 3 rotational degrees of freedom.
3 vibrational modes: symmetric stretching, asymmetric stretching and scissor bending.

Example 1: Carbon dioxide molecule CO_2.
3 atoms, 9 degrees of freedom, linear molecule.
3 translational degrees of freedom and 2 rotational degrees of freedom.
4 vibrational modes: symmetric stretching, asymmetric stretching, two bending modes.

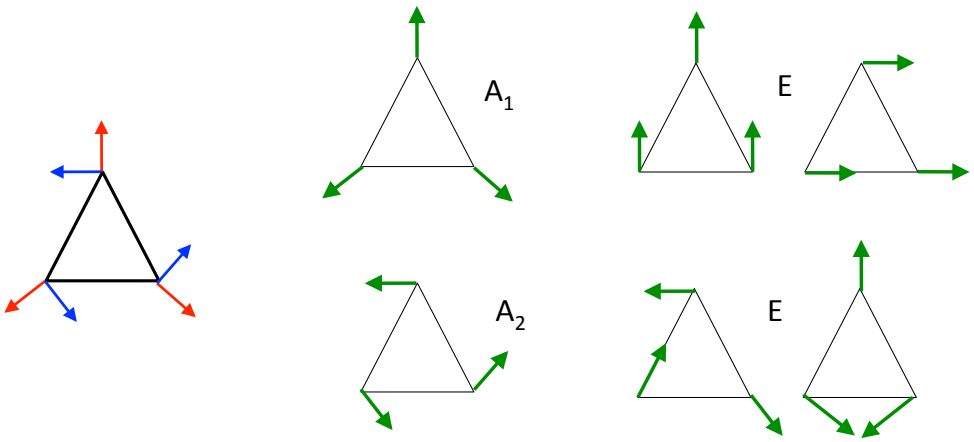

Figure 7.3: Planar triangular molecule. Left: directions of atomic displacements. Centre: non-degenerate normal modes, bases of A_1 and A_2 irreducible representations. Right: two-fold degenerate normal modes, bases of E irreducible representations.

7.2.5 Degenerate modes and group representation

It can happen that different normal modes are characterised by the same frequency; these normal modes are said to be degenerate.
Degenerate normal modes share the same frequency but are characterised by different eigenvectors.

As the degeneracy of quantum energy levels (Section 5.6), even the degeneracy of normal modes corresponds to the dimensionality of the irreducible representations of the symmetry group of the system.
As a simple example, let us consider a triangular molecule, whose symmetry transformations form the C_{3v} group (e.g. the basal plane of the ammonia molecule). The system is planar; the atomic displacements with respect to the equilibrium positions can be represented by six components, two for each atom. There are thus six normal modes, each one characterised by a different eigenvector, say a different pattern of atomic displacements.
The choice of the directions of the atomic displacements is to some extent arbitrary; a clever choice can however simplify the calculations. For each atom, it is convenient to choose the radial direction and the direction perpendicular to it (Fig. 7.3, left).
These six displacement coordinates represent a 6-dimensional basis $\Gamma^{6\times6}$ for the C_{3v} group. The characters of the symmetry transformations in the Γ representation can be easily evaluated and are listed in the last line of Table 7.1. It is easy to verify that the representation is reducible and can be decomposed as $\Gamma^{6\times6} = A_1 + A_2 + 2E$, say into four irreducible representations.
This means that there are only four different eigenfrequencies ω, two corresponding to non-degenerate normal modes, two corresponding to two two-fold degenerate normal modes.
A further analysis shows that

1. The basis of the A_1 representation is a breathing mode, where all atoms vibrate in phase in the radial direction (Fig. 7.3, top centre).
2. The basis of the A_2 representation is a mode where all atoms rotate in phase; for a free molecule this mode has zero frequency and is of no interest for vibrational dynamics (Fig. 7.3, bottom centre).
3. The bases of one of the E representations are two translational modes in normal directions, which are mixed by the symmetry transformation of the molecule; for a free molecule these modes have zero frequency (Fig. 7.3, top right).
4. The bases of the other E representation are two modes with relatively complicated displacement patterns, which are mixed by the symmetry transformations of the molecule (Fig. 7.3,

bottom right).

Table 7.1: Character table of the symmetry group C_{3v}. The last line contains the characters of the representation base on the six displacements described in the text.

	e	$2C_3$	$3\sigma_v$
A_1	1	1	1
A_2	1	1	-1
E	2	-1	0
$\Gamma^{6\times6}$	6	0	0

7.3 Vibrational dynamics of crystals

For crystals, in order to exploit the translational symmetry, it is convenient to factorize the total number $\mathcal{N}$ of atoms as $\mathcal{N} = Nn$, where N is the number of primitive cells, n is the number of atoms per primitive cell.
The positions of atoms are labelled by two indices, ℓ for the primitive cells and κ for the positions inside the cell:

$$\vec{R}(m) \;=\; \vec{R}(\ell\kappa) \;=\; \vec{R}(\ell) + \vec{R}(\kappa) \tag{7.72}$$

1. The positions of the primitive cells are given by the Bravais lattice vectors

$$\vec{R}(\ell) = \ell_1\vec{a}_1 + \ell_2\vec{a}_2 + \ell_3\vec{a}_3 \tag{7.73}$$

   where ℓ_1, ℓ_2, ℓ_3 are integer numbers. The index ℓ labels the primitive cells: $\ell = 1, \ldots, N$.
2. The positions of the atoms inside a primitive cell are given by

$$\vec{R}(\kappa) = x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 \tag{7.74}$$

   where x_1, x_2, x_3 are fractional coordinates. The index κ (the Greek letter kappa) is unnecessary for Bravais crystals, say for crystals with one atom per primitive cell (such as Cu, Ag, Au or the noble gas solids).

Note: When studying the effects of atomic vibrations ("lattice dynamics"), one considers only primitive cells, not conventional cells.

Crystal potential energy

The potential energy of the crystal $V[\{\vec{r}(\ell\kappa)\}]$ is defined in a $3nN$-dimensional configuration space. Atoms oscillate around the average positions $\{\vec{R}(\ell\kappa)\}$, and the instantaneous positions are

$$\vec{r}(\ell\kappa) \;=\; \vec{R}(\ell\kappa) + \vec{u}(\ell\kappa)\,. \tag{7.75}$$

We again rely on the potential energy expansion (7.49).
The **second-order term** (the only one relevant in the harmonic approximation) is

$$V_2 \;=\; \frac{1}{2}\sum_{\substack{\alpha\ell\kappa\\ \alpha'\ell'\kappa'}} \Phi\begin{pmatrix}\alpha\,\ell\,\kappa\\ \alpha'\ell'\kappa'\end{pmatrix}\, u_\alpha(\ell\kappa)\, u_{\alpha'}(\ell'\kappa') \tag{7.76}$$

where the second-order coupling parameters are

$$\Phi\begin{pmatrix}\alpha\,\ell\,\kappa\\ \alpha'\ell'\kappa'\end{pmatrix} = \left(\frac{\partial^2 V}{\partial u_\alpha(\ell\kappa)\,\partial u_{\alpha'}(\ell'\kappa')}\right)_{\{\vec{R}\}}\,. \tag{7.77}$$

Again, the coupling parameters are invariant with respect to the exchange of the indices:

$$\Phi\begin{pmatrix}\alpha\,\ell\,\kappa\\ \alpha'\ell'\kappa'\end{pmatrix} = \Phi\begin{pmatrix}\alpha'\ell'\kappa'\\ \alpha\,\ell\,\kappa\end{pmatrix} \tag{7.78}$$

7.3.1 Constraints on coupling parameters

Two general constraints on force constants Φ have been introduced above in (7.62) and (7.78). In crystals, further constraints are present, depending

a) on general invariance properties, such as invariance under translation or under rotation;
b) on the peculiar symmetry properties of a given crystal.

We introduce here the constraint due to translational invariance.
For an infinite crystal, the lattice translations shift the primitive cells: $\ell \to \ell + T$, $\ell' = \ell' + T$ and the atomic positions are modified accordingly

$$(\ell, \kappa) \longrightarrow (\ell + T, \kappa) \qquad (\ell', \kappa') \longrightarrow (\ell' + T, \kappa') \tag{7.79}$$

The coupling parameters are invariant with respect to the lattice translations:

$$\Phi \begin{pmatrix} \alpha, & \ell + T, & \kappa \\ \alpha', & \ell' + T, & \kappa' \end{pmatrix} = \Phi \begin{pmatrix} \alpha & \ell & \kappa \\ \alpha' & \ell' & \kappa' \end{pmatrix} \tag{7.80}$$

In particular, for $T = -\ell$ and for $T = -\ell'$, respectively, (7.80) becomes

$$\Phi \begin{pmatrix} \alpha & \ell & \kappa \\ \alpha' & \ell' & \kappa' \end{pmatrix} = \Phi \begin{pmatrix} \alpha, & 0, & \kappa \\ \alpha', & \ell' - \ell, & \kappa' \end{pmatrix} = \Phi \begin{pmatrix} \alpha, & \ell - \ell', & \kappa \\ \alpha', & 0, & \kappa' \end{pmatrix} \tag{7.81}$$

The force constants only depend on the differences $L = \ell - \ell'$, say on the relative positions of primitive cells, not on their absolute positions:

$$\Phi \begin{pmatrix} \alpha & \ell & \kappa \\ \alpha' & \ell' & \kappa' \end{pmatrix} = \Phi \begin{pmatrix} \alpha, & 0, & \kappa \\ \alpha', & -L, & \kappa' \end{pmatrix} = \Phi \begin{pmatrix} \alpha, & L, & \kappa \\ \alpha', & 0, & \kappa' \end{pmatrix} . \tag{7.82}$$

For finite crystals, the same conclusions hold, provided suitable boundary conditions (introduced in Section 2.5) are imposed.

7.3.2 Periodic boundary conditions

Let us consider a crystal of parallelepiped shape. The lengths of the three edges of the parallelepiped are $N_1 a_1, N_2 a_2, N_3 a_3$, respectively. There are $N = N_1 N_2 N_3$ primitive cells.
The periodic boundary conditions on the atomic displacements impose that

$$\vec{u}(\vec{r} + N_\alpha \vec{a}_\alpha) \;=\; \vec{u}(\vec{r}) \qquad (\alpha = 1, 2, 3) . \tag{7.83}$$

According to (7.83), the displacement of an atom sitting on one face of the parallelepiped is equal to the displacement of the corresponding atom on the opposite face.

Irreducible representations and boundary conditions

In Section 5.4, we studied the one-dimensional linear chain of N points spaced by a, with periodic boundary conditions. We found that there are N one-dimensional irreducible representations of the group of translations, whose simplest basis functions are (5.58)

$$e^{iqx} , \qquad q = \frac{2\pi}{a} \frac{m}{N} \quad (-N/2 < m \leq N/2) , \tag{7.84}$$

where the values q are confined within the first Brillouin Zone. A translation $T = sa$ is represented, in the basis e^{iqx} (say for a given q) by the complex number

$$\left(e^{isa}\right)^q \;=\; e^{iqsa} \;=\; e^{iqT} . \tag{7.85}$$

For a *three-dimensional* crystal with periodic boundary conditions, one can build up $N = N_1 N_2 N_3$ one-dimensional irreducible representations of the translation group, whose basis functions are

$$e^{i\vec{q}\cdot\vec{r}}\,, \qquad q_\alpha = b_\alpha \frac{m_\alpha}{N_\alpha} \quad (-N_\alpha/2 < m_\alpha \leq N_\alpha/2)\,, \tag{7.86}$$

where b_α is the magnitude of the $\vec{b}_\alpha$ primitive vector of reciprocal space. A translation $\vec{T}$ is represented, in the basis $e^{i\vec{q}\cdot\vec{r}}$ (say for a given $\vec{q}$) by the complex number

$$e^{i\vec{q}\cdot\vec{T}}\,. \tag{7.87}$$

The vectors $\vec{q}$ form a grid of equally spaced points within the first Brillouin Zone of the reciprocal space. They label the N irreducible representations of the translation group. When the crystal size increases, the number N of primitive cells in the real space increases, and the density of $\vec{q}$ points in reciprocal space increases accordingly.
The basis functions (7.86) are plane waves, of wavelength $\lambda = 2\pi/q$.

7.3.3 Spatial periodicity of eigenvectors

For an irreducible representation of the translation group, characterised by the wavevector $\vec{q}$, the effect of a translation $\vec{T}$ on each eigenvector is the multiplication by (7.87).
By identifying $\vec{T} = \vec{R}(\ell)$, one can thus express each eigenvector $w_\alpha(\ell\kappa)$ component, in the $\vec{q}$ representation, as a product

$$w_\alpha(\ell\kappa) \;=\; w_\alpha(\kappa)\; e^{i\vec{q}\cdot\vec{R}(\ell)}. \tag{7.88}$$

Note: For the monatomic one-dimensional linear chain, (7.88) reduces to $w(\ell) = \; w\, \exp[iqR(\ell)]$.
The atomic displacement (7.66), adapted to the crystal case

$$u_\alpha(\ell\kappa, t) \;=\; \frac{1}{\sqrt{M_\kappa}}\, \mathbf{Re}\left[\, w_\alpha(\ell\kappa)\, e^{-i\omega t}\right]\,, \tag{7.89}$$

is modified by (7.88) to

$$u_\alpha(\ell k, t) \;=\; \frac{1}{\sqrt{M_\kappa}}\, \mathbf{Re}\left[\, w_\alpha(\kappa)\, e^{i\vec{q}\cdot\vec{R}(\ell)}\, e^{-i\omega t}\right]\,. \tag{7.90}$$

7.3.4 The Fourier-transformed dynamical matrix

Let us now consider the dynamical equations (7.69), adapted to the crystal case

$$\omega^2\, w_\alpha(\ell\kappa) \;-\; \sum_{\alpha'\ell'\kappa'} D^0 \begin{pmatrix} \alpha\ \ell\ \kappa \\ \alpha'\ell'\kappa' \end{pmatrix} w_{\alpha'}(\ell'\kappa') \;=\; 0\,. \tag{7.91}$$

By substitute (7.88) for a given value of $\vec{q}$, one gets

$$\omega^2\, w_\alpha(\kappa)\; e^{i\vec{q}\cdot\vec{R}(\ell)} \;-\; \sum_{\alpha'\ell'\kappa'} D^0 \begin{pmatrix} \alpha\ \ell\ \kappa \\ \alpha'\ \ell'\ \kappa' \end{pmatrix} e^{i\vec{q}\cdot\vec{R}(\ell')}\, w_{\alpha'}(\kappa') \;=\; 0\,. \tag{7.92}$$

Moving $e^{i\vec{q}\cdot\vec{R}(\ell)}$ to the right of (7.92), substituting $L = \ell - \ell'$ as in (7.82) and substituting the sum over ℓ' by the sum over L, one obtains

$$\omega^2\, w_\alpha(\kappa) \;-\; \sum_{\alpha'\kappa'} \left\{ \sum_L D^0 \begin{pmatrix} \alpha\ L\ \kappa \\ \alpha'\ 0\ \kappa' \end{pmatrix} e^{-i\vec{q}\cdot\vec{R}(L)} \right\} w_{\alpha'}(\kappa') \;=\; 0\,. \tag{7.93}$$

For each one of the N wavevectors $\vec{q}$, one can define a quantity

$$D_{\vec{q}} \begin{pmatrix} \alpha\kappa \\ \alpha'\kappa' \end{pmatrix} \;=\; \sum_L D^0 \begin{pmatrix} \alpha\ L\ \kappa \\ \alpha'\ 0\ \kappa' \end{pmatrix} e^{-i\vec{q}\cdot\vec{R}(L)} \tag{7.94}$$

corresponding to the term included in curly brackets in (7.93).
For each one of the N wavevectors $\vec{q}$ there is thus a system of $3n \times 3n$ coupled equations

$$\omega^2 \, w_\alpha(\kappa) \;=\; \sum_{\kappa' \alpha'} D_{\vec{q}} \begin{pmatrix} \alpha\kappa \\ \alpha'\kappa' \end{pmatrix} w_{\alpha'}(\kappa'). \tag{7.95}$$

By considering

- the terms $w_\alpha(\kappa)$ as elements of a $3n$-sized column vector $U_{\vec{q}}$,
- the terms $D_{\vec{q}} \left({\alpha\kappa \atop \alpha'\kappa'} \right)$ as elements of a $3n \times 3n$ matrix, the *Fourier transformed dynamical matrix* $\mathbf{D}_{\vec{q}}$ (frequently simply referred to as dynamical matrix),

the system of equations (7.95) can be synthesised as a matrix equation:

$$\boxed{\omega^2 \, U_{\vec{q}} \;=\; \mathbf{D}_{\vec{q}} U_{\vec{q}}} \tag{7.96}$$

or more explicitly

$$\omega^2 \begin{pmatrix} \cdots \\ w_\alpha(\kappa) \\ \cdots \\ \cdots \end{pmatrix} = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & D_{\vec{q}} \left({\alpha\;\kappa \atop \alpha'\kappa'} \right) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \cdots \\ \cdots \\ w_{\alpha'}(\kappa') \\ \cdots \end{pmatrix} \tag{7.97}$$

Note: In principle, to evaluate the Fourier transformed dynamical matrix for a given $\vec{q}$ one should evaluate the sum (7.94) over all values of L. Actually, due to the short range of most interactions, often only a few terms of the sum are significant.

Eigenfrequencies and eigenvectors of the FT dynamical matrix

The solution of (7.97) is again an eigenvalue problem. By imposing that the secular determinant be zero, one obtains an algebraic equation, whose solution gives (for each $\vec{q}$)

→ $3n$ eigenvalues or eigenfrequencies $\omega(\vec{q}, s)$ which are labeled, in addition of $\vec{q}$, by a branch index $s = 1, 2, \ldots, 3n$; the frequencies correspond to $3n$ independent normal modes of wavevector $\vec{q}$.

The Fourier-transformed dynamical matrix $\mathbf{D}_{\vec{q}}$ has complex elements. One can show that it is Hermitian

$$D_{\vec{q}} \begin{pmatrix} \alpha'\kappa' \\ \alpha\kappa \end{pmatrix} \;=\; D^*_{\vec{q}} \begin{pmatrix} \alpha\kappa \\ \alpha'\kappa' \end{pmatrix}, \tag{7.98}$$

so that its eigenvalues are real. The stability of the lattice further requires that the eigenvalues ω^2 are non-negative quantities too.

For each one of the $3n$ eigenfrequency $\omega(\vec{q}s)$, say for each normal mode, eq. (7.97) gives an eigenvector $U_{\vec{q}s}$ with $3n$ components that describe the relative amplitudes and the phase relationships of the atomic motions corresponding to the normal mode.
In total, for a given wavevector $\vec{q}$, there are thus

→ $3n$ eigenvectors $U_{\vec{q}s}$, each one with $3n$ components $w_\alpha(\kappa|\vec{q}s)$.

Note: Degenerate modes. The solution of the Fourier transformed dynamical matrix can lead to degenerate modes, characterised by the same frequency and different atomic displacement patterns, say different eigenvectors.

Evaluation of the dynamical matrix

The coupling coefficients that form the dynamical matrix can be obtained from force-constant phenomenological models or from calculations from first principles.

The simplest models are based on short-range forces and are characterised by bond stretching force constants between nearest-neighbours and next-nearest-neighbours as well as by bond bending force constants. From the force constants of the models, calibrated on the macroscopic elastic constants or on other measurable properties such as the compressibility, one calculates the harmonic coupling coefficients of the Born-von Karman expansion.

More refined models take into account long-range forces, which are necessary to account for the dynamics in metals or in ionic crystals.

Calculation from first principles are nowadays based on perturbation approaches to the Density Functional Theory, which allow to directly evaluate the coupling coefficients from the shape of the PES in the vicinity of the equilibrium configuration.

7.3.5 Atomic displacements and normal coordinates

The displacement of each atom $(\ell\kappa)$ is the sum of the contributions of all the $3nN$ independent normal modes. The contribution of a normal mode $(\vec{q}s)$ to the displacement of atom $(\ell\kappa)$ is

$$u_\alpha(\ell\kappa, t \,|\, \vec{q}s) = \frac{1}{\sqrt{M_\kappa}} \mathbf{Re}\left[\, w_\alpha(\kappa|\vec{q}s)\, e^{i\vec{q}\cdot\vec{R}(\ell)}\, e^{-i\omega(\vec{q}s)t} \,\right]. \tag{7.99}$$

This contribution is determined from the dynamical matrix only to within a complex factor (amplitude and phase), which depends on the initial conditions. In particular, the amplitude of the eigenvector components $w_\alpha(\kappa|\vec{q}s)$ depends on the energy stored in the normal mode.
It is convenient to factorize the quantities $w_\alpha(\kappa|\vec{q}s)$, separating the information on the magnitude of the relative motion of atoms inside the primitive cell (geometrical information) from the contributions due to the total energy (which depends on temperature).
To this effect, (7.99) can be rewritten as

$$u_\alpha(\ell\kappa, t \,|\, \vec{q}s) = \frac{1}{\sqrt{M_\kappa}} \mathbf{Re}\left[\mathrm{e}_\alpha(\kappa|\vec{q}s)\, \frac{Q(\vec{q}s, t)}{\sqrt{N}} \quad e^{i\vec{q}\cdot\vec{R}(\ell)} \right] \tag{7.100}$$

where:

a) The factor $\mathrm{e}_\alpha(\kappa|\vec{q}, s)$ is a component of the *normalized eigenvector* $\mathrm{e}(\vec{q}, s)$, which describes the displacement patterns of the n atoms inside each primitive cell.

b) The quantity
$$Q(\vec{q}s, t) = Q_0(\vec{q}s) \exp[-i\omega(\vec{q}s)t] \tag{7.101}$$
is the normal coordinate of mode $(\vec{q}s)$; the amplitude Q_0 is a complex quantity, which depends on the initial conditions and on the energy stored in the normal mode $(\vec{q}s)$.

c) The factor $1/\sqrt{N}$ can be justified as follows. The number of primitive cells N is proportional to the number of normal modes $3nN$. The energy stored in each normal mode, proportional to the square of the normal coordinate (see below) is independent of N. The energy of the normal mode is distributed over all the atoms, whose number is proportional to N.

The total displacement of the atom $(\ell\kappa)$ can be calculated by summing up the contributions of all normal modes within the first B.Z.:

$$u_\alpha(\ell\kappa, t) = \frac{1}{\sqrt{N\, M_\kappa}} \sum_{\vec{q}s} \left[\mathrm{e}_\alpha(\kappa|\vec{q}s)\, e^{i\vec{q}\cdot\vec{R}(\ell)}\, Q(\vec{q}s, t) \right] \tag{7.102}$$

Any pair of values $\vec{q}$ and $-\vec{q}$ in the first B.Z. correspond to two plane waves traveling in opposite direction and sharing the same frequency.
The displacement $u_\alpha(\ell\kappa, t)$ is a real quantity, so that $Q(-\vec{q}s, t) = Q^*(\vec{q}s, t)$.

7.3.6 More on normal mode wavevectors

The periodic boundary conditions on the atomic displacements impose that

$$\vec{u}(\vec{r}+N_\alpha\vec{a}_\alpha) \;=\; \vec{u}(\vec{r}) \qquad (\alpha=1,2,3)\,. \tag{7.103}$$

We want now to show that (7.103) is automatically fulfilled by the choice of wavevectors $\vec{q}$ of (7.86), say $q_\alpha = b_\alpha\, m_\alpha/N_\alpha$. To this purpose let us express the contributions to the two displacements of (7.103) due to a given normal mode of wavevector $\vec{q}$ according to (7.99):

$$u_\alpha(\vec{R},t) \;=\; \frac{1}{\sqrt{M}}\,\mathbf{Re}\left\{w_\alpha(\kappa|\vec{q})\,\exp\left[i\vec{q}\cdot\vec{R}\right]\,\exp(-i\omega t)\;\right\} \tag{7.104}$$

$$u_\alpha(\vec{R}+N_\alpha\vec{a}_\alpha,t) \;=\; \frac{1}{\sqrt{M}}\,\mathbf{Re}\left\{w_\alpha(\kappa|\vec{q})\,\exp\left[i\vec{q}\cdot(\vec{R}+N_\alpha\vec{a}_\alpha)\right]\,\exp(-i\omega t)\;\right\} \tag{7.105}$$

Since $\vec{q}$ is a vector of the reciprocal space, such that

$$\vec{q} \;=\; (m_1/N_1)\vec{b}_1 + (m_2/N_2)\vec{b}_2 + (m_3/N_3)\vec{b}_3\,, \tag{7.106}$$

and recalling the relation between primitive vectors of real and reciprocal space:

$$\vec{a}_\alpha\cdot\vec{b}_\beta \;=\; 2\pi\,\delta_{\alpha\beta} \tag{7.107}$$

one can easily show that (7.105) is equal to (7.104).

The first Brillouin Zone

One can demonstrate that the frequencies and eigenvectors corresponding to any wavevector $\vec{q}$ are unmodified if a reciprocal lattice vector $\vec{G}$ is added to $\vec{q}$. Actually, the Fourier transformed dynamical matrices corresponding to $\vec{q}+\vec{G}$ and to $\vec{q}$ are identical

$$D_{\vec{q}+\vec{G}}\begin{pmatrix}\alpha\;\kappa\\ \alpha'\kappa'\end{pmatrix} \;=\; \sum_L D^0\begin{pmatrix}\alpha\;L\;\kappa\\ \alpha'\;0\;\kappa'\end{pmatrix}\exp\left\{i\left[\vec{q}+\vec{G}\right]\cdot\vec{R}(L)\right\} \;=\; D_{\vec{q}}\begin{pmatrix}\alpha\;\kappa\\ \alpha'\kappa'\end{pmatrix} \tag{7.108}$$

because $\exp(i\vec{G}\cdot\vec{R}) = 1$.

7.4 Energy of normal modes. Phonons

From the classical point of view, each normal mode $(\vec{q}s)$ is a travelling wave of wavevector $\vec{q}$, whose frequency is obtained by diagonalising the Fourier transformed dynamical matrix $\mathbf{D}_{\vec{q}}$.
We want now to express the total energy of the system as a sum of the contributions of the independent normal modes. To this purpose, we will use the lagrangian and hamiltonian formalisms. We will then switch to the quantum picture, by introducing creation and destruction operators for each normal mode, as linear combinations of the normal coordinates and their conjugate momenta. Following the same procedure as for the one-dimensional harmonic oscillator, the Hamiltonian of each normal mode will be expressed in terms of creation and destruction operators.

7.4.1 Energy and normal coordinates

The total kinetic energy depends on the squares of the first derivatives of atomic displacements. If the expressions (7.102) are substituted for the atomic displacements, one can demonstrate (see § 7.6.4 for details) that the total kinetic energy can be expressed as a function of the derivatives of the normal coordinates:

$$T \;=\; \frac{1}{2}\sum_{\alpha\ell\kappa} M_\kappa\dot{u}^2_\alpha(\ell\kappa) \;=\; \frac{1}{2}\sum_{\vec{q}s}\dot{Q}(\vec{q}s)\dot{Q}^*(\vec{q}s) \tag{7.109}$$

The potential vibrational energy, expressed in terms of atomic displacements, is

$$V_2 \;=\; \frac{1}{2} \sum_{\substack{\alpha \ell \kappa \\ \alpha' \ell' \kappa'}} \Phi_2 \begin{pmatrix} \alpha\, \ell\, \kappa \\ \alpha' \ell' \kappa' \end{pmatrix} \; u_\alpha(\ell\kappa)\, u_{\alpha'}(\ell'\kappa') \tag{7.110}$$

If the expressions (7.102) are substituted for the atomic displacements in (7.110), one finds (see again § 7.6.4 for details) that the harmonic part of the potential vibrational energy can be expressed as

$$V_2 \;=\; \frac{1}{2} \sum_{\vec{q}s} \omega^2(\vec{q}s)\, Q(\vec{q}s)\, Q^*(\vec{q}s)\,. \tag{7.111}$$

The cross terms appearing in the expression (7.110) as a function of atomic diplacements disappear in (7.111) as a function of normal coordinates. Normal modes contribute independently to the total energy of the system.

Note: The dimensions of the normal coordinate are $[Q] = [M^{1/2}L]$.

Lagrange function

The Lagrange function of the system of normal modes is

$$\mathcal{L} \;=\; T + V_2 \;=\; \frac{1}{2} \sum_{\vec{q}s} \left[\dot{Q}(\vec{q}s) \dot{Q}^*(\vec{q}s) - \omega^2(\vec{q}s) Q(\vec{q}s) Q^*(\vec{q}s) \right] , \tag{7.112}$$

It is interesting to compare each term of the Lagrange function (7.112) with the Lagrange function of a harmonic oscillator: $\mathcal{L} = mv^2/2 - m\omega^2 x^2/2$.
Each classical normal mode $(\vec{q}s)$ can be considered as an independent harmonic oscillator of frequency $\omega(\vec{q}s)$, whose amplitude and time dependence are described by the normal coordinate $Q(\vec{q}s, t) = Q_0(\vec{q}s)\, e^{i\omega(\vec{q}s)t}$.
The normal coordinates correspond to generalised coordinates of the Lagrangian formalism.

Hamilton function

Introducing the conjugate momenta $P(\vec{q}s) = \partial\mathcal{L}/\partial\dot{Q}(\vec{q}s) = \dot{Q}^*(\vec{q}s)$, one obtains the Hamiltonian

$$\mathcal{H} \;=\; \frac{1}{2} \sum_{\vec{q}s} \left[P(\vec{q}s) P^*(\vec{q}s) + \omega^2(\vec{q}s) Q(\vec{q}s) Q^*(\vec{q}s) \right] . \tag{7.113}$$

The classical energy stored in the normal mode $(\vec{q}s)$ is

$$E(\vec{q}s) \;=\; \frac{1}{2}\, \omega^2(\vec{q}s)\, Q_0^2(\vec{q}s)\,. \tag{7.114}$$

7.4.2 Quantization of normal modes

The quantum approach to vibrational dynamics consists in replacing the classical normal coordinates and momenta by the corresponding operators. By the same procedure followed for the one-dimensional harmonic oscillator, the destruction and creation operators and the number operator N are introduced
By introducing the operators (destruction and creation)

$$\begin{aligned} a(\vec{q}s) \;&=\; \frac{1}{\sqrt{2\hbar\omega(\vec{q}s)}} \left[\omega(\vec{q}s)\, Q(\vec{q}s) + iP^*(\vec{q}s) \right] \\ &=\; \frac{1}{\sqrt{2\hbar\omega(\vec{q}s)}} \left[\omega(\vec{q}s)\, Q(\vec{q}s) + iP(-\vec{q}s) \right] \end{aligned} \tag{7.115}$$

$$\begin{aligned} a^\dagger(\vec{q}s) \;&=\; \frac{1}{\sqrt{2\hbar\omega(\vec{q}s)}} \left[\omega(\vec{q}s)\, Q^*(\vec{q}s) - iP(\vec{q}s) \right] \\ &=\; \frac{1}{\sqrt{2\hbar\omega(\vec{q}s)}} \left[\omega(\vec{q}s)\, Q(-\vec{q}s) - iP(\vec{q}s) \right] \end{aligned} \tag{7.116}$$

one can express the normal coordinates and momenta as

$$Q(\vec{q}s) = \sqrt{\frac{\hbar}{2\omega(\vec{q}s)}}\,\left[a(\vec{q}s)+a^\dagger(-\vec{q}s)\right] \tag{7.117}$$

(7.118)

$$P(\vec{q}s) = -i\sqrt{\frac{\hbar\omega(\vec{q}s)}{2}}\,\left[a(-\vec{q}s)-a^\dagger(\vec{q}s)\right] \tag{7.119}$$

(7.120)

Taking into account the commutation relations

$$\left[a(\vec{q}s),a^\dagger(\vec{q}'s')\right]=\delta_{\vec{q}\vec{q}'}\delta_{ss'}\,,\quad \left[a(\vec{q}s),a(\vec{q}'s')\right]=\left[a^\dagger(\vec{q}s),a^\dagger(\vec{q}'s')\right]=0\,, \tag{7.121}$$

the Hamiltonian (7.113) can be written as

$$\mathcal{H} = \sum_{\vec{q}s}\left[a^\dagger(\vec{q}s)\,a(\vec{q}s)+\frac{1}{2}\right]\hbar\omega(\vec{q}s) = \left[n(\vec{q}s)+\frac{1}{2}\right]\hbar\omega(\vec{q}s)\,. \tag{7.122}$$

The energy eigenvalues of each normal mode $(\vec{q}s)$ are given by

$$E(\vec{q}s) = \left[n(\vec{q}s)+\frac{1}{2}\right]\hbar\omega(\vec{q}s)\,. \tag{7.123}$$

In the energy representation, the quantum state of the normal mode $(\vec{q}s)$ is characterised by the quantum number $n(\vec{q}s)$. The quantum state of the entire system is characterised by the set of quantum numbers of all normal modes:

$$|n(\vec{q}_1s_{11}),n(\vec{q}_1s_{12}),\ldots n(\vec{q}_is_{ij}),\ldots\rangle = |\{n(\vec{q}s)\}\rangle \tag{7.124}$$

7.4.3 Statistical thermodynamics of normal modes

A crystal can exchange heat with its surroundings. Let us consider here a non-metallic crystal, where thermal energy is only stored by lattice vibrations (in metals, electrons can store energy too). Within the harmonic approximation, normal modes are distinguishable independent oscillators. A level of total energy of the crystal can be expressed as

$$\mathcal{E}_j = V_0+\sum_{\vec{q}s}\left[n(\vec{q}s)+\frac{1}{2}\right]\hbar\omega(\vec{q}s) \tag{7.125}$$

The partition function of the whole crystal is the sum over all total energy levels

$$Z = \sum_{j=1}^{\infty}\exp\left[-\beta\mathcal{E}_j\right]\,, \tag{7.126}$$

where $\beta=1/k_BT$. By substituting (7.125) into (7.126), after some simple calculations, one finds

$$Z = \exp\left[-\beta V_0\right]\prod_{\vec{q}s}\frac{\exp\left[-\beta\hbar\omega(\vec{q}s)/2\right]}{1-\exp\left[-\beta\hbar\omega(\vec{q}s)\right]}\,. \tag{7.127}$$

Note: Compare with the expression (7.16) of the partition function of the one-dimensional harmonic oscillator.

The internal energy of the crystal is

$$\begin{aligned} U &= \langle\mathcal{E}\rangle = k_BT^2\left(\frac{\partial}{\partial T}\ln Z\right) \\ &= V_0+\sum_{\vec{q}s}\frac{1}{2}\hbar\omega(\vec{q}s)+\sum_{\vec{q}s}\frac{\hbar\omega(\vec{q}s)}{\exp[\hbar\omega(\vec{q}s)/k_BT]-1}\,. \end{aligned} \tag{7.128}$$

Three contributions are singled out in (7.128):

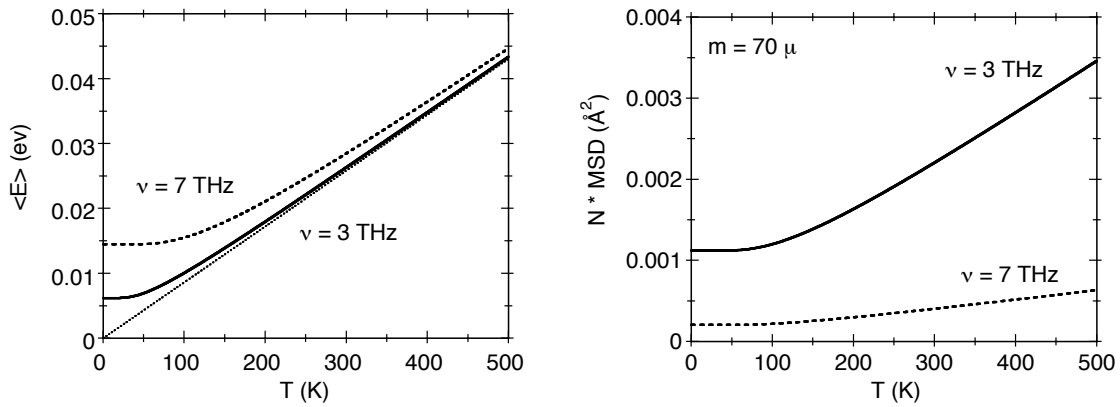

Figure 7.4: Left: average energy as a function of temperature for two normal modes of frequencies $\nu = 3$ and $\nu = 7$ THz, respectively ($\nu = \omega/2\pi$); the straight dotted line is the classical approximation, $\langle E\rangle = k_B T$. Right: Mean square displacement MSD of an ion of mass 70 μ along a given direction induced by N modes of frequencies $\nu = 3$ and $\nu = 7$ THz, respectively.

1. the static cohesive energy V_0
2. the vibrational zero-point energy
3. the vibrational thermal energy (dependent on temperature)

The average vibrational energy stored in normal mode $(\vec{q}s)$ at temperature T is thus

$$\langle E(\vec{q}s|T)\rangle \;=\; \left\{\frac{1}{2} + \frac{1}{\exp[\hbar\omega(\vec{q}s)/k_B T] - 1}\right\} \hbar\omega(\vec{q}s) \tag{7.129}$$

Equivalent alternative expressions are:

$$\langle E(\vec{q}s|T)\rangle \;=\; \frac{\hbar\omega(\vec{q}s)}{2} \coth\frac{\hbar\omega(\vec{q}s)}{2k_B T} \;=\; \frac{\hbar\omega(\vec{q}s)}{2}\,\frac{1+z}{1-z}\,, \tag{7.130}$$

where $z = \exp[-\hbar\omega(\vec{q}s)/k_B T]$.
The dependence on temperature of the average energy of two normal modes of different frequencies is shown in Fig. 7.4 (left). At high temperatures, the average energy tends to the classical value $\langle E\rangle \to k_B T$, independent of frequency. At low temperatures, the larger is the frequency, the larger is the average energy.

7.4.4 Phonons

The quanta of energy of the normal modes are called *phonons*, by analogy with photons, the quanta of the normal modes of the electromagnetic field. The number $n(\vec{q}s)$ appearing in (7.123) is the number of phonons of mode $(\vec{q}s)$.
The quantum state of a normal mode is characterised by the number $n(\vec{q}s)$ of phonons, say by its energy content. The vibrational quantum state of a crystal is characterised by the numbers of phonons for each normal mode. A phonon is characterized by a frequency ω, a wavevector $\vec{q}$ and an eigenvector $w(\vec{q}s)$.
The number of phonons is not constant. The weak exchange of energy between normal modes and with the surrounding environment, which guarantees the thermodynamic equilibrium, corresponds to the creation and annihilation of phonons. Phonons are created or annihilated when the temperature of a crystal increases or decreases, respectively.
The average energy of a normal mode can be expressed as

$$\langle E(\vec{q}s)\rangle \;=\; \left[\langle n(\vec{q}s)\rangle + \frac{1}{2}\right] \hbar\omega(\vec{q}s)\,. \tag{7.131}$$

By comparing with (7.129), the average value of $n(\vec{q}s)$ is given by

$$\langle n(\vec{q}s)\rangle \;=\; \frac{1}{\exp[\hbar\omega(\vec{q}s)/k_BT]-1}\,. \tag{7.132}$$

This expression corresponds to the Bose-Einstein distribution of a set of indistinguishable objects whose total number is variable. Actually phonons of the same frequency are indistinguishable and can be created and destroyed.
A number of phenomena, such as thermal conductivity, can be interpreted by considering phonons as quasi-particles. One often speaks of "phonon gas".

7.4.5 Temperature dependence of normal coordinates

As a consequence of (7.114) and (7.129), the amplitude of the normal coordinate of mode $(\vec{q}s)$ depends on temperature according to

$$\langle Q_0^2(\vec{q}s|T)\rangle \;=\; \left\{\frac{1}{2}+\frac{1}{\exp[\hbar\omega(\vec{q}s)/k_BT]-1}\right\}\frac{2\hbar}{\omega(\vec{q}s)}\,. \tag{7.133}$$

Equivalent expression:

$$\langle Q_0^2(\vec{q}s|T)\rangle \;=\; \frac{\hbar}{\omega(\vec{q},s)}\coth\frac{\hbar\omega(\vec{q},s)}{2k_BT}\,. \tag{7.134}$$

7.4.6 Mean square atomic displacements

Let us now consider the mean square displacement (MSD) of atom κ along a given direction α induced by a normal mode $(\vec{q}s)$:

$$MSD_\alpha \;=\; \langle u_\alpha^2(\kappa|\vec{q}s)\rangle\,. \tag{7.135}$$

By taking into account the expression (7.100) of the atomic instantaneous displacement, one gets

$$\begin{aligned}\langle u_\alpha^2(\kappa|\vec{q}s)\rangle \;&=\; \frac{1}{NM_\kappa}\,[\mathrm{e}_\alpha(\kappa|\vec{q}s)]^2\,\langle Q_0^2(\vec{q}s|T)\rangle\\ &=\; \frac{2}{NM_\kappa}\,[\mathrm{e}_\alpha(\kappa|\vec{q}s)]^2\,\frac{\langle E(\vec{q}s|T)\rangle}{\omega^2(\vec{q}s)}\,.\end{aligned} \tag{7.136}$$

The MSD of an atom κ along a direction α depends on the component of the normalized eigenvector e and on the temperature.
The dependence on temperature of the MSD induced by a sets of N normal modes of equal frequencies is illustrated in Fig. 7.4 (right) for the two cases $\nu = 3$ and $\nu = 7$ THz, respectively. As one can see, the high frequency modes induce a smaller average displacement and a weaker dependence on temperature than the low frequency modes.
The total mean square displacement of an atom is the sum of the contributions of all normal modes. It gives rise to the thermal spread of the distribution of atomic positions, measured by the Debye-Waller factor in scattering experiments.

7.4.7 Mean square relative displacements

At last, let us consider the mean square relative displacement (MSRD) of atom κ' relative to atom κ along a given direction α induced by a normal mode $(\vec{q}s)$. One can easily see that

$$\langle[u_\alpha(\kappa'|\vec{q}s)-u_\alpha(\kappa|\vec{q}s)]^2\rangle \;=\; \frac{1}{N}\,\langle Q_0^2(\vec{q}s|T)\rangle\,\left|\frac{\mathrm{e}_\alpha(\kappa'|\vec{q}s)e^{i\vec{q}\cdot\vec{R}}}{\sqrt{m(\kappa')}}-\frac{\mathrm{e}_\alpha(\kappa|\vec{q}s)}{\sqrt{m(\kappa)}}\right|^2 \tag{7.137}$$

The total MSRD of an atomic pair is the sum of the contribution of all normal modes. The MSRD gives rise to the damping of the EXAFS signal.

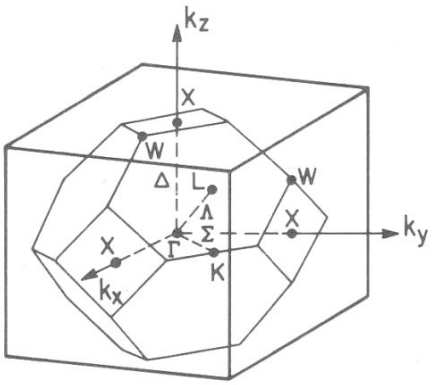

Figure 7.5: First Brillouin Zone of face centered cubic (fcc) Bravais lattices. In the figure, the wavevectors are indicated by $\vec{k}$ instead of $\vec{q}$.

7.5 First Brillouin Zone and dispersion relations

In Section 7.2 we introduced the periodic boundary conditions (7.83) and the N different wavevectors $\vec{q}$ labelling the N irreducible representations of the translation group for a finite crystal with N primitive cells. The tips of the wavevectors $\vec{q}$ represent a grid of equally spaced points within the first Brillouin Zone. To each value of $\vec{q}$ it corresponds a $3n \times 3n$ Fourier transformed dynamical matrix.

7.5.1 First Brillouin Zone

The central point of the First BZ ($\vec{q} = 0$) is labelled by the Greek letter Γ (upper-case gamma). The Γ point doesn't correspond to any normal mode (an infinite wavelength would correspond to a rigid translation of the whole crystal).

Travelling and stationary waves

Each point $\vec{q}$ of the 1st BZ different from the Γ point corresponds to a travelling plane wave in the real space, which classically affects the atom at $\vec{R}(\ell\kappa)$ proportionally to

$$\exp[i\vec{q}\cdot\vec{R}(\ell\kappa) - i\omega t]\,. \tag{7.138}$$

For each wavevector $\vec{q}$ there is a symmetric wavevector $-\vec{q}$, corresponding to a wave travelling in the opposite direction with the same amplitude. The two waves sum up to give a stationary wave. In thermodynamic equilibrium, the crystal dynamics is thus constituted by a huge number of stationary waves. If thermodynamic equilibrium is broken, waves travelling in opposite directions can have different amplitudes, giving rise to transport phenomena (e.g. heat transport, Section 8.3).

Problem: Consider a monatomic linear chain in the real space, with lattice parameter a. Determine the first Brillouin Zone (BZ) in the reciprocal space. Verify that the motion of atoms is the same for different wavevector values that differ by a reciprocal lattice vector. Find the relation between the lattice spacing a and the wavelength corresponding to the BZ border.

First BZ of cubic crystals

The shape of the first BZ depends on the Bravais lattice structure. Let us consider here only the cubic lattices. It is easy to verify that

1. The reciprocal lattice of a simple cubic Bravais lattice is again simple cubic.

2. The reciprocal lattice of a bcc (body centred cubic) Bravais lattice is an fcc (face centred cubic) lattice. The BZ of a bcc Bravais lattice has thus the shape of the Wigner-Seitz cell of the fcc lattice.
3. The reciprocal lattice of an fcc (face centred cubic) Bravais lattice is a bcc (body centred cubic) lattice. The BZ of an fcc Bravais lattice has thus the shape of the Wigner-Seitz cell of the bcc lattice (Fig. 7.5).

Face centered cubic Bravais lattices in the real space are shared by many important elements and compounds. For example:

a) The structure of many metals, such as Cu, Ag, Au, and of noble gas solids (Ne, Ar, Kr...) is fcc, with one atom per primitive cell.
b) The diamond structure (C, Si, Ge) and the related zincblende structure (ZnS, GaAs, CdTe...) is based on an fcc Bravais lattice with two atoms per primitive cell.

We focus here our attention on fcc Bravais lattices. The BZ (Fig. 7.5) is a truncated octahedron, with six square faces and eight hexagonal faces.
The highest symmetry points on the surface of the BZ, and the corresponding directions, are labelled by conventional symbols, for example:

- [100] or Δ direction, from Γ to X point
- [110] or Σ direction, from Γ to K point
- [111] or Λ direction, from Γ to L point

Problem: Consider an fcc Bravais lattice in the real space and the corresponding bcc lattice in the reciprocal space. Verify that the relation between the lattice constants is $b^3 = (2\pi)^3/a^3$. Determine the shape of the first Brillouin Zone (BZ) in the reciprocal space. For each of the three directions [100], [110] and [111], find the relations between the inter-planar spacing d and the wavelength corresponding to the BZ border.

7.5.2 Dispersion relations

Each normal mode $(\vec{q}s)$ is characterised by a frequency $\omega(\vec{q}s)$ and by a normalised eigenvector $\mathrm{e}(\vec{q}s)$ of components $\mathrm{e}_\alpha(\kappa|\vec{q}s)$. A basic information on the lattice dynamical properties of a crystal in the harmonic approximation is contained in the dispersion relations, which express the eigenfrequencies ω as a function of the wavevectors $\vec{q}$:

$$\vec{q} \quad \rightarrow \quad \omega(\vec{q}s)\,. \tag{7.139}$$

The dispersion relation for the normal modes of the electromagnetic field (say for photons) is simply $\omega = cq$. For vibrational normal modes (say for phonons) the dispersion relations $\omega(\vec{q}s)$ depend on the values of $\vec{q}$ and s and on the force constants of the dynamical matrix, which in turn depend on the point symmetry properties of the crystal.

In principle, the full knowledge of the dispersion relations would require the evaluation of ω for each one of the $3nN$ points $\vec{q}$ of the first B.Z. Actually, for symmetry reasons the BZ can be decomposed into a number of equivalent parts. Any such irreducible part of the BZ contains all the information on the lattice dynamics. Experiments or calculations are performed by choosing a convenient representative number of points $\vec{q}$ into an irreducible part of the BZ and solving the Fourier transformed dynamical matrix for each chosen $\vec{q}$.

Example: For a system with full cubic symmetry (group O_h) there are 48 point symmetry transformations, each one of which transforms an eigenvector $\vec{q}$ into an equivalent eigenvector $\vec{q}' = R\,\vec{q}$. The BZ can thus be divided into 48 equivalent irreducible parts.

The resulting dispersion relations are displayed and compared in conventional two-dimensional plots representing ω versus $\vec{q}$ along selected high-symmetry directions. The lines connecting the points are the dispersion curves. Examples are given in Figs. 7.6 and 7.7.

Information on the eigenvectors of the dynamical matrix is much less frequently sought than information on the eigenfrequencies and is not contained in the plots of dispersion relations.

Example: The knowledge of eigenvectors can be useful to evaluate the relative vibrational amplitude of atoms of different mass.

Dispersion relations for monatomic lattices

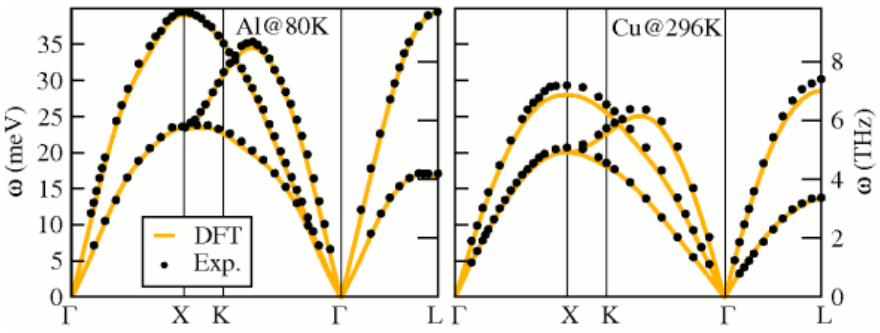

Figure 7.6: Phonon dispersion curves for Al (left panel) and Cu (right panel). The dots are experimental points from inelastic neutron scattering, the continuous lines are ab-initio calculations by the density functional (DFT) approach [From Grabowski et al., Phys. Rev. 2007]. Recall that 1 THz corresponds to 4.13 meV; the two plots have the same vertical scale.

Crystals with one atom per primitive cell represent the simplest cases. They are the three-dimensional analogue of the monatomic linear chain.
The Fourier transformed dynamical matrix $\mathbf{D}_{\vec{q}}$ is 3×3. For each $\vec{q}$ one has then

1. three eigenfrequencies $\omega_1, \omega_2, \omega_3$
2. three normalized eigenvectors, each one with three components, so that they correspond to standard euclidean vectors $\vec{w}_1, \vec{w}_2, \vec{w}_3$

One can demonstrate that the symmetry of the dynamical matrix $\mathbf{D}_{\vec{q}}$ implies that the eigenvectors are mutually orthogonal

$$\sum_{\alpha} w_\alpha(\vec{q}s)\ w_\alpha(\vec{q}s') = \delta_{ss'} \tag{7.140}$$

For wavevectors $\vec{q}$ parallel to high-symmetry directions

1. one of the eigenvectors is parallel to the wavevector $\vec{q}$; it corresponds to *longitudinal* vibrations of atoms in the direction of $\vec{q}$;
2. the remaining two eigenvectors are perpendicular to $\vec{q}$; they correspond to *transverse* vibrations of atoms with respect to the direction of $\vec{q}$.

Correspondingly, the dispersion curves can be classified in three branches:

1. two transverse acoustic (TA) branches;
   for high-symmetry directions, the TA modes are often degenerate, say $\omega(\vec{q}, 1) = \omega(\vec{q}, 2)$; this is the case, for example, for $\vec{q}$ along the [100] direction in fcc crystals;
2. one longitudinal acoustic (LA) branch.

The dispersion curves of the fcc metals Al and Cu are shown in Fig. 7.6. As usual, the plots show the dispersion curves measured and calculated along high symmetry directions within the first Brillouin Zone.
The frequencies ν are of the order of the tera-Hertz (10^{12} Hz). The frequency of ν=1 THz corresponds to an energy $h\nu = \hbar\omega$=4.13×10^{-3} eV. The phonon energies are of the order of the meV (compare with the thermal energy $k_B T$).

Near to the Γ point, say for long wavelengths of normal modes, all branches show a linear behaviour $\omega(\vec{q}s) = c_s q$, where c_s is the sound velocity of the given branch along the given direction.
It is evident in Fig. 7.6 that LA modes have highest frequencies and highest velocities than TA modes. This is not surprising, since the longitudinal stretching forces are stronger than the transverse shear forces.

When q increases (say when the wavelength decreases), the linear approximation becomes progressively worse.
The group velocity is $v = d\omega/dq$; it decreases progressively when going from the BZ centre to the BZ borders, where it frequently is equal to zero.

Table 7.2: Longitudinal and transverse velocity of sound for Al and Cu compared with the value for air.

	Al	Cu	air
c_{long} (m/s)	6420	4760	340
c_{tra} (m/s)	3040	2325	

Dispersion relations for many-atomic lattices

Crystal structures with two or more atoms per primitive cell are the three-dimensional equivalent of the many-atomic linear chain. It should be noted, however, that in the case of crystals the two (or more) atoms per primitive cell can correspond to the same atomic species and can have the same environment (a simple example is given by Si and Ge, whose structure is made by an f.c.c. Bravais lattice with two atoms per primitive cell).
For a crystal with n atoms per primitive cell one has

1. 3 acoustic branches, with the same general properties as in monatomic lattices;
2. $3n-3$ optic branches, containing modes whose frequency never goes to zero, and is generally higher than the frequency of acoustic modes.

One can demonstrate the orthogonality relation of eigenvectors

$$\sum_{\alpha\kappa} w_\alpha(\kappa|\vec{q}s)\; w_\alpha(\kappa|\vec{q}s') = \delta_{ss'} \tag{7.141}$$

The case of germanium is shown in Fig. 7.7.

7.6 Complements and demonstrations

7.6.1 The one-dimensional quantum harmonic oscillator

The Hamiltonian operator for the quantum harmonic oscillator is

$$H_0 \;=\; \frac{P^2}{2m} + \frac{1}{2} m\omega_0^2 X^2\,, \tag{7.142}$$

where X and P are the position and momentum operators, whose commutator is $[X, P] = i\hbar$. It is convenient to introduce the a-dimensional reduced observables

$$\hat{X} = \sqrt{\frac{m\omega_0}{\hbar}}\, X\,, \qquad \hat{P} = \frac{1}{\sqrt{m\hbar\omega_0}}\, P\,, \tag{7.143}$$

whose commutator is $[\hat{X}, \hat{P}] = i$; the reduced Hamiltonian is

$$\hat{H}_0 = \frac{H_0}{\hbar\omega_0} \;=\; \frac{1}{2}\,(\hat{X}^2 + \hat{P}^2)\,. \tag{7.144}$$

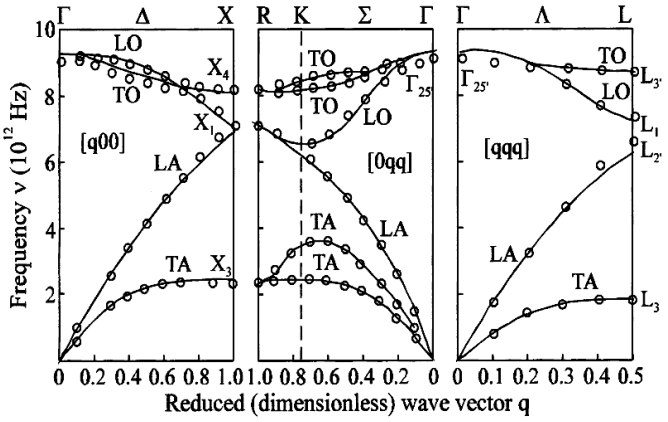

Figure 7.7: Phonon dispersion curves for Ge. Calculations by the valence shell model of Weber, 1977.

It is further convenient to substitute $\hat{X}$ and $\hat{P}$ by the (non hermitian) operators defined by the linear combinations

$$a^\dagger = \frac{1}{\sqrt{2}}(\hat{X} - i\hat{P}), \qquad a = \frac{1}{\sqrt{2}}(\hat{X} + i\hat{P}). \tag{7.145}$$

so that

$$\hat{X} = \frac{1}{\sqrt{2}}(a^\dagger + a), \qquad \hat{P} = \frac{i}{\sqrt{2}}(a^\dagger - a) \tag{7.146}$$

From $[\hat{X}, \hat{P}] = i$ one can easily verify that the commutator is $[a, a^\dagger] = 1$ and that

$$a^\dagger a \;=\; \frac{1}{2}(\hat{X}^2 + i\hat{X}\hat{P} - i\hat{P}\hat{X} + \hat{Y}^2) \;=\; \frac{1}{2}(\hat{X}^2 + \hat{P}^2 - 1). \tag{7.147}$$

The reduced Hamiltonian can then be expressed as

$$\hat{H}_0 = a^\dagger a + \frac{1}{2} \;=\; aa^\dagger - \frac{1}{2}. \tag{7.148}$$

By defining the (hermitian) number operator

$$N = a^\dagger a, \tag{7.149}$$

the reduced Hamiltonian becomes

$$\hat{H}_0 = N + 1/2 \;: \tag{7.150}$$

the eigenvectors of $\hat{H}_0$ are eigenvectors of N and viceversa.
One can then solve the eigenvalues equation for N

$$N\,|\phi_\nu\rangle \;=\; \nu\,|\phi_\nu\rangle \tag{7.151}$$

One finds (see Quantum Mechanics textbooks) that the spectrum of N is represented by non negative integer numbers

$$\nu = 0, 1, 2, \ldots n, \ldots \qquad |\phi_n\rangle = |n\rangle, \tag{7.152}$$

the energy levels are non degenerate and the eigenvalues of the Hamiltonian H_0 are

$$E_n^0 \;=\; (1/2 + n)\,\hbar\omega_0 \qquad (n = 0 \ldots \infty) \tag{7.153}$$

The energy difference between any two levels is

$$E_n - E_{n-1} = \hbar\omega_0. \tag{7.154}$$

The action of the $a^\dagger$ and a operators on the eigenstates is as follows:

$$a^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle, \qquad a\,|n\rangle = \sqrt{n}\,|n-1\rangle. \tag{7.155}$$

They are called creation and destruction operators, respectively, since they create or destroy a quantum of energy.
From (7.155) and from

$$\hat{X} = \frac{1}{\sqrt{2}}\,(a^\dagger + a) \tag{7.156}$$

one can easily see that the average displacement $\langle X\rangle$ is zero in the eigenstates $|n\rangle$:

$$\langle X\rangle = \sqrt{\frac{\hbar}{2m\omega_0}}\,\langle n|\,a^\dagger + a\,|n\rangle = \sqrt{\frac{\hbar}{2m\omega_0}}\,\left[\langle n|\,a^\dagger|n\rangle + \langle n|\,a|n\rangle\right] = 0\,. \tag{7.157}$$

One can further see that

$$\langle X^2\rangle = \frac{\hbar}{2m\omega_0}\,\langle n|(a^\dagger + a)^2|n\rangle = \frac{\hbar}{2m\omega_0}\,\langle n|a^\dagger a + aa^{dag}|n\rangle \tag{7.158}$$

$$= \frac{\hbar}{m\omega_0}\left(\frac{1}{2} + n\right) = \frac{E_n}{m\omega_0^2}\,. \tag{7.159}$$

7.6.2 The one-dimensional quantum anharmonic oscillator

Let us consider here for simplicity only the cubic term in the potential energy.
The harmonic hamiltonian H_0 is modified by adding a weak cubic term $W = k^3X^3$.

$$H = H_0 + W = \frac{1}{2}\,m\omega_0^2\,X^2 + \frac{1}{2m}P^2 + k^3X^3\,. \tag{7.160}$$

Introducing again the a-dimensional reduced observables $\hat{X}$ and $\hat{P}$ and expressing the third order term as $W = \sigma\hbar\omega_0\hat{X}^3$, where $\sigma = k^3\hbar^{1/2}/m^{3/2}\omega_0^{5/2}$ (a-dimensional), the Hamiltonian becomes

$$H = H_0 + W = \frac{1}{2}\,\hbar\omega_0\,(\hat{X}^2 + \hat{P}^2) + \sigma\hbar\omega_0\hat{X}^3\,. \tag{7.161}$$

If the third order term is sufficiently weak, one can consider it as a perturbation to the harmonic Hamiltonian.

Energy levels

According to the stationary perturbation theory, the energy eigenvalues E_n of the anharmonic oscillator are obtained from the energy eigenvalues E_n^0 of the harmonic oscillator as

$$E_n = E_n^0 + \langle n|W|n\rangle + \sum_{n\neq n'} \frac{|\langle n'|W|n\rangle|^2}{E_n^0 - E_{n'}^0} + \ldots \tag{7.162}$$

Taking into account that $\hat{X} = (a^\dagger + a)/\sqrt{2}$, one finds

$$\hat{X}^3 = \frac{1}{2^{3/2}}\,\left[(a^\dagger)^3 + a^\dagger aa^\dagger + aa^\dagger a^\dagger + aaa^\dagger + a^\dagger a^\dagger a + a^\dagger aa + aa^\dagger a + a^3\right]\,, \tag{7.163}$$

and exploiting the commutation relations of $a^\dagger$, a and $N = a^\dagger a$

$$\hat{X}^3 = \frac{1}{2^{3/2}}\,\left[(a^\dagger)^3 + 3Na^\dagger + 3(N+1)a + a^3\right]\,, \tag{7.164}$$

where each term corresponds to the creation or destruction of three quanta or of one quantum of energy.

The first-order term in the perturbation expansion, where the final state is equal to the initial state, is thus zero.
The only non-zero matrix elements of W are those connecting states that differ for one or three quanta:

$$\frac{|\langle n+3|W|n\rangle|^2}{E_n^0 - E_{n+3}^0} = \frac{(\sigma\hbar\omega_0)^2}{8}\frac{(n+1)(n+2)(n+3)}{-3\hbar\omega_0} = -\sigma^2\hbar\omega_0\frac{n^3+6n^2+11n+6}{24} \quad (7.165)$$

$$\frac{|\langle n-3|W|n\rangle|^2}{E_n^0 - E_{n-3}^0} = \frac{(\sigma\hbar\omega_0)^2}{8}\frac{n(n-1)(n-2)}{+3\hbar\omega_0} = +\sigma^2\hbar\omega_0\frac{n^3-3n^2+2n}{24} \quad (7.166)$$

$$\frac{|\langle n+1|W|n\rangle|^2}{E_n^0 - E_{n+1}^0} = \frac{(\sigma\hbar\omega_0)^2}{8}\frac{9(n+1)^3}{-\hbar\omega_0} = -\sigma^2\hbar\omega_0\frac{9n^3+27n^2+27n+9}{8} \quad (7.167)$$

$$\frac{|\langle n-1|W|n\rangle|^2}{E_n^0 - E_{n-1}^0} = \frac{(\sigma\hbar\omega_0)^2}{8}\frac{9n^3}{+\hbar\omega_0} = +\sigma^2\hbar\omega_0\frac{9n^3}{8} \quad (7.168)$$

By adding to the harmonic energy the sum of the four perturbation terms (mind the signs) one finds

$$E_n = \left(n+\frac{1}{2}\right)\hbar\omega_0 - \sigma^2\,\hbar\omega_0\,\frac{30n^2+30n+11}{8} \quad (7.169)$$

or equivalently

$$E_n = \left(n+\frac{1}{2}\right)\hbar\omega_0 - \frac{15}{4}\sigma^2\left(n+\frac{1}{2}\right)^2\hbar\omega_0 - \frac{7}{16}\sigma^2\hbar\omega_0 + \ldots \quad (7.170)$$

The energy levels are lowered with respect to the unperturbed harmonic case.
The difference between the energies of adjacent levels is not constant (as for the harmonic oscillator), but depends on the value n:

$$E_n - E_{n-1} = \hbar\omega_0\left[1 - \frac{15}{2}\,\sigma^2\,n\right]. \quad (7.171)$$

Eigenstates

At last, one can show that the eigenstates of the anharmonic oscillator are different from the eigenstates $|n\rangle$ of the harmonic oscillator, being contaminated by the states $|n+1\rangle, |n-1\rangle, |n+3\rangle, |n-3\rangle$.

7.6.3 Statistics of the one-dimensional harmonic oscillator

Let us consider an harmonic oscillator in equilibrium with a reservoir at temperature $T = 1/k_B\beta$.

Classical statistics

In the classical approximation, the instantaneous state of a one-dimensional system is defined by its position x and momentum p. For a system in thermal equilibrium with a reservoir at temperature $T = 1/k_B\beta$, the probability density of the point (x,p) in phase space is given by

$$\rho(x,p) = \frac{e^{-\beta E(x,p)}}{Z} = \frac{e^{-\beta E(x,p)}}{\int dx \int dp\, e^{-\beta E(x,p)}}, \quad (7.172)$$

and the average value of a quantity $A(x,p)$ is given by

$$\langle A(x,p)\rangle = \frac{1}{Z}\int dx\,dp\,A(x,p)\,\rho(x,p). \quad (7.173)$$

If the potential energy is independent of p, the total energy is $E(x,p) = p^2/2m + V(x)$, so that the partition function Z can be calculated as

$$Z \;=\; \int dp\, e^{-\beta p^2/2m} \int dx\, e^{-\beta V(x)} \;=\; \sqrt{2m\pi/\beta} \int dx\, e^{-\beta V(x)}. \tag{7.174}$$

The probability density for a given value of position x is then

$$\rho(x) \;=\; \frac{\int dp\, e^{-\beta p^2/2m}\, e^{-\beta V(x)}}{Z} \;= \frac{e^{-\beta V(x)}}{\int dx\, e^{-\beta V(x)}}, \tag{7.175}$$

and the average value of a quantity $A(x,p)$ is

$$\langle A(x)\rangle \;=\; \frac{\int dx\, A(x)\, e^{-\beta V(x)}}{\int dx\, e^{-\beta V(x)}}. \tag{7.176}$$

For the harmonic oscillator, $V(x) = k_0 x^2/2$.

Quantum statistics

If the oscillator is in equilibrium with a reservoir at temperature T, its state is a statistical mixture of eigenstates $|n\rangle$, weighted by $\exp(-\beta E_n)$, where $\beta = 1/k_B T$.
The thermodynamical properties can be calculated from the statistical density operator

$$\hat{w} \;=\; (1/Z_0)\; e^{-\beta \hat{H}_0}, \qquad Z_0 \;=\; \mathrm{Tr}\,(e^{-\beta \hat{H}_0}). \tag{7.177}$$

In the energy representation, where $\hat{H}_0$ is diagonal, the harmonic partition function Z_0 is

$$Z_0 \;=\; \sum_{n=0}^{\infty} \langle n|e^{-\beta \hat{H}_0}|n\rangle \;=\; \sum_{n=0}^{\infty} e^{-(n+1/2)\beta\hbar\omega} \;=\; \frac{e^{-\beta\hbar\omega/2}}{1-e^{-\beta\hbar\omega}}. \tag{7.178}$$

The average energy is given by

$$\langle \hat{H}_0\rangle \;=\; \mathrm{Tr}(\hat{H}_0\hat{w}) \;=\; \frac{1}{Z_0}\,\mathrm{Tr}(\hat{H}_0\, e^{-\beta \hat{H}_0}) \;=\; \frac{1}{Z_0} \sum_{n=0}^{\infty} (n{+}1/2)\hbar\omega e^{-\beta(n+1/2)\hbar\omega} \;=\; \hbar\omega \left[\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega}-1}\right]. \tag{7.179}$$

a) For $T \to 0$ ($\beta \to \infty$), it easy to see that $\langle H\rangle \to \hbar\omega/2$ (zero point energy). The zero point energy is proportional to the angular frequency ω.

b) For $T \to \infty$ ($\beta \to 0$), one can show that $\langle H\rangle \to k_B T$ (classical approximation). The demonstration is based on the expansion $1/(e^x - 1) \simeq [x + x^2/2 + \ldots]^{-1} \simeq (1/x)[1 - x/2 + \ldots]$, where $x = \beta\hbar\omega$. The behaviour for $T \to \infty$ is independent of ω.

The probability density for the displacement x has a gaussian shape centred on $\langle x\rangle = 0$:

$$\rho(x) \;=\; (1/\sigma\sqrt{2\pi})\, e^{-x^2/2\sigma^2}\,. \tag{7.180}$$

The moments of the distribution can be calculated as

$$\langle x^k\rangle \;=\; \frac{1}{Z_0}\,\mathrm{Tr}\,(x^k e^{-\beta H_0}) \;=\; \frac{1}{Z_0} \sum_{n=0}^{\infty} \langle n|x^k|n\rangle\, e^{-\beta E_n}. \tag{7.181}$$

7.6.4 Normal coordinates and energy

We want to express the kinetic and potential energies in terms of the normal coordinates.

Useful relation

As a preliminary step, it is convenient to introduce the relation

$$\sum_{\ell} e^{i\vec{q}\cdot\vec{R}(\ell)} = \begin{cases} N & \text{for } \vec{q}=0 \text{ or } \vec{q}=\vec{G}\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{7.182}$$

Relation (7.182) is a consequence of the lattice periodicity; the sum is unchanged if an arbitrary $\vec{R}(\ell')$ is added to $\vec{R}(\ell)$ in the sum:

$$\sum_{\ell} e^{i\vec{q}\cdot\vec{R}(\ell)} = \sum_{\ell} e^{i\vec{q}\cdot[\vec{R}(\ell)+\vec{R}(\ell')]} = \left(\sum_{\ell} e^{i\vec{q}\cdot\vec{R}(\ell)}\right) e^{i\vec{q}\cdot\vec{R}(\ell')}\,. \tag{7.183}$$

This is possible only if conditions (7.182) are fulfilled.

As a consequence of (7.182),

$$\sum_{\ell} e^{i(\vec{q}+\vec{q}')\cdot\vec{R}(\ell)} = \begin{cases} N & \text{for } \vec{q}+\vec{q}'=0 \text{ or } \vec{q}+\vec{q}'=\vec{G}\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{7.184}$$

Kinetic energy

Substituting the atomic displacements (7.102) in the expression of the kinetic energy, since only the normal coordinates depend on time, the kinetic energy becomes

$$T = \frac{1}{2}\frac{1}{N}\sum_{\alpha\kappa}\sum_{\vec{q}s}\sum_{\vec{q}'s} \mathrm{e}_\alpha(\kappa|\vec{q}s)\,\mathrm{e}_\alpha(\kappa|\vec{q}'s')\,\dot{Q}(\vec{q}s)\,\dot{Q}(\vec{q}'s') \sum_{\ell} e^{i(\vec{q}+\vec{q}')\cdot\vec{R}(\ell)}\,. \tag{7.185}$$

Taking into account that according to (7.184) the last sum is N if $\vec{q}=-\vec{q}'$, otherwise it is zero, and considering the orthonormality relation of eigenvectors, one gets

$$T = \frac{1}{2}\sum_{\vec{q}s}\dot{Q}(\vec{q}s)\,\dot{Q}(-\vec{q}s) = \frac{1}{2}\sum_{\vec{q}s}\dot{Q}(\vec{q}s)\,\dot{Q}^*(\vec{q}s)\,. \tag{7.186}$$

Potential energy

Substituting the atomic displacements (7.102) in the expression for the potential energy, the potential energy becomes

$$V_2 = \frac{1}{2N}\sum_{\substack{\vec{q}s\\ \vec{q}'s'}}\sum_{\substack{\alpha\ell\kappa\\ \alpha'\ell'\kappa'}} \frac{\Phi\left(\begin{smallmatrix}\alpha & m\\ \alpha' & m'\end{smallmatrix}\right)}{\sqrt{m_\kappa\, m_{\kappa'}}}\, e^{i\vec{q}\cdot\vec{R}(\ell)}\, e^{i\vec{q}'\cdot\vec{R}(\ell')}\, \mathrm{e}_\alpha(\kappa|\vec{q}s)\,\mathrm{e}_{\alpha'}(\kappa'|\vec{q},'s')\; Q(\vec{q}s)\; Q(\vec{q}'s') \tag{7.187}$$

Multipliying by

$$1 = e^{i\vec{q}'\cdot[\vec{R}(\ell)-\vec{R}(\ell)]}$$

and recalling the definition of Fourier transformed dynamical matrix, (7.187) becomes

$$V_2 = \frac{1}{2N}\sum_{\substack{\vec{q}s\\ \vec{q}'s'}}\sum_{\substack{\alpha\kappa\\ \alpha'\kappa'}} D_{\vec{q}'}\left(\begin{smallmatrix}\alpha\kappa\\ \alpha'\kappa'\end{smallmatrix}\right)\left[\sum_{\ell} e^{i(\vec{q}+\vec{q}')\cdot\vec{R}(\ell)}\right] \mathrm{e}_\alpha(\kappa|\vec{q}s)\,\mathrm{e}_{\alpha'}(\kappa'|\vec{q}'s')\; Q(\vec{q}s)\; Q(\vec{q}'s') \tag{7.188}$$

Taking again into account that according to (7.184) the sum in square parentheses is N if $\vec{q}=-\vec{q}'$, otherwise it is zero and making use of the eigenvalue equation of $\mathrm{e}^*_\alpha(\kappa|-\vec{q}s)=\mathrm{e}_\alpha(\kappa|\vec{q}s)$, one obtains

$$V_2 = \frac{1}{2}\sum_{\substack{\vec{q}s\\ s'}}\sum_{\alpha\kappa} \omega^2(\vec{q}s)\; \mathrm{e}_\alpha(\kappa|\vec{q}s)\,\mathrm{e}^*_\alpha(\kappa|\vec{q}s')\; Q(\vec{q}s)\; Q(-\vec{q}s') \tag{7.189}$$

At last, considering the orthonormality relation of eigenvectors, one finds taht the potential vibrational energy can be expressed as

$$V_2 \;=\; \frac{1}{2}\sum_{\vec{q}s}\omega^2(\vec{q}s)\,Q(\vec{q}s)\,Q(-\vec{q}s) \;=\; \frac{1}{2}\sum_{\vec{q}s}\omega^2(\vec{q}s)\,Q(\vec{q}s)\,Q^*(\vec{q}s)\,. \qquad (7.190)$$

A classical normal mode $(\vec{q}s)$ can be considered as an independent harmonic oscillator of frequency $\omega(\vec{q}s)$, whose amplitude and time dependence are described by the normal coordinate $Q(\vec{q}s,t) = Q_0(\vec{q}s)\,e^{i\omega(\vec{q}s)t}$. The dimensions of the normal coordinate are $[Q] = [M^{1/2}L]$.

7.7 Bibliography of Chapter 7

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions).
  Chapters 22 (classical theory of the harmonic oscillator), 23 (quantum theory of the harmonic oscillator) and 24 (measuring phonon dispersion relations).
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapters 4 (phonons: crystal vibrations) and 5 (phonons: thermal properties).
- L. Landau and E. Lifshitz: *Theoretical Physics* (various editions in different languages).
  Vol. 1: *Mechanics*, Chapter 5 (Oscillations).
- C. Cohen-Tannoudji, B. Diu, F. Laloe: *Quantum Mechanics*, Wiley. Chapter 5 (Harmonic oscillator). Complement J_v (Phonons). Complement A_{xi} (Anharmonic oscillator).
- P. Brüesch: *Phonons: theory and experiment*, Springer 1986
  (Introduction to vibrational dynamics of crystals).
- M. Dove: *Introduction to lattice dynamics*, Cambridge University Press 1993.
- B.T.M. Willis and A.W. Pryor: *Thermal vibrations in crystallography* Cambridge University Press 1975.
- G. Leibfried and W. Ludwig: *Theory of anharmonic effects in crystals*, Solid State Physics, vol 12
- R.P. Feynman: *Statistical mechanics*, Benjamin 1972
  Statistics of the anharmonic oscillator

Chapter 8

Vibrational thermal properties

In this chapter we discuss some equilibrium and non-equilibrium thermal properties of crystals that are connected to the ion vibrational dynamics.
Two properties can be accounted for within the harmonic approximation with good accuracy:

1. the contribution of ion vibrations to the specific heat of crystals (§ 8.1)
2. the effect of ion vibrations on the Bragg diffraction peaks, quantified by the Debye-Waller factors, which will be treated in Chapter 11.

Two properties can be accounted for only by considering anharmonicity:

3. the thermal expansion (§ 8.2)
4. the vibrational contribution to the heat conduction (§ 8.3)

8.1 Vibrational heat capacity

Specific heats (and heat capacities) of solids depend on the ion vibrations and, for conductors, also on electronic properties.
In insulators, the specific heat depends only on the dynamical properties of ions; more specifically, it depends on the normal mode frequencies, say on the eigenvalues of the dynamical matrix. The eigenvectors of the dynamical matrix have no influence on specific heats.

8.1.1 The contribution of a single normal mode

According to (7.129), the average energy stored in a normal mode of frequency ω at the temperature T is given by

$$\langle \epsilon \rangle \;=\; \hbar\omega \left[\frac{1}{2} + \frac{1}{e^{\hbar\omega/k_B T} - 1} \right] . \tag{8.1}$$

The constant-volume contribution of the single mode to the heat capacity is :

$$c_v \;=\; \left(\frac{\partial \langle \epsilon \rangle}{\partial T} \right)_v \;=\; k_B \left(\frac{\hbar\omega}{k_B T} \right)^2 \frac{e^{\hbar\omega/k_B T}}{\left(e^{\hbar\omega/kT} - 1 \right)^2} . \tag{8.2}$$

One can show (see§ 8.4.1) that

a) for $T \to \infty$ the single mode contribution $c_v \to k_B$ (classical limit);
b) at low temperatures

$$c_v \;\simeq\; k_B \left(\frac{\hbar\omega}{k_B T} \right)^2 e^{-\hbar\omega/k_B T} \tag{8.3}$$

and for $T \to 0$ the single mode contribution $c_v \to 0$.

8.1.2 The total heat capacity

The total vibrational energy of a crystal in equilibrium at temperature $T = 1/\beta k_B$ is the sum of the contributions of all normal modes, each one with frequency $\omega(\vec{q}s)$:

$$\langle E \rangle = \sum_{\vec{q}s} \frac{1}{2}\hbar\omega(\vec{q}s) + \sum_{\vec{q}s} \frac{\hbar\omega(\vec{q}s)}{e^{\hbar\omega(\vec{q}s)/k_B T} - 1} . \tag{8.4}$$

For $T \to \infty$ equation (8.4) reduces to the classical expression $E = 3Nnk_BT$, where N is the number of primitive cells and n is the number of atoms per primitive cell.
The constant-volume specific heat *per unit volume is*:

$$c_v = \frac{1}{V}\left(\frac{\partial\langle E\rangle}{\partial T}\right)_v = \frac{\partial}{\partial T}\sum_{\vec{q}s} \frac{\hbar\omega(\vec{q}s)}{e^{\hbar\omega(\vec{q}s)/k_B T} - 1} . \tag{8.5}$$

The specific heats of some crystals are compared in Fig. 8.1, left.
For $T \to \infty$ equation (8.5) reduces to the classical law of Dulong and Petit $C_v = 3Nnk_B$.
The low-temperature behaviour requires more attention than for a single mode.

Integration over the first Brillouin Zone

To calculate (8.5) it is convenient to substitute the sum over the discrete $\vec{q}$ values by an integral over the 1st Brillouin Zone (BZ):

$$\sum_{\vec{q}s} \cdots \longrightarrow \sum_s \int_{BZ} \cdots \rho(\vec{q})\, d\vec{q}, \tag{8.6}$$

where $\rho(\vec{q})$ is the number of $\vec{q}$ points per unit volume in the 1st BZ:

$$\rho(\vec{q}) = \frac{N}{V_b} = \frac{N}{(2\pi)^3/V_a} = \frac{NV_a}{(2\pi)^3} = \frac{V}{(2\pi)^3} \tag{8.7}$$

Here V_a and V_b are the volumes of the primitive cells in direct and reciprocal space, respectively, and V is the volume of the crystal. The relation (8.6) becomes

$$\sum_{\vec{q}s} \cdots \longrightarrow \frac{V}{(2\pi)^3} \sum_s \int_{BZ} \cdots d\vec{q}. \tag{8.8}$$

and the constant-volume specific heat per unit volume (8.5) can be evaluated by

$$c_v = \frac{1}{(2\pi)^3}\frac{\partial}{\partial T}\left[\sum_s \int d\vec{q}\, \frac{\hbar\omega(\vec{q}s)}{e^{\hbar\omega(\vec{q}s)/k_B T} - 1}\right] . \tag{8.9}$$

Low-temperature specific heat

A common behaviour is shared by all systems at low temperatures, where

a) only the low-frequency acoustic modes can store thermal energy
b) acoustic modes have a linear dispersion relation $\omega = c_s q$
c) the integration of (8.8) can anyway be extended to all the reciprocal space, since at low T only the low-frequency modes are excited and contribute to the integral.

As a consequence, one finds (see Ashcroft-Mermin for details) that for $T \to 0$ the specific heat per unit volume, for any system, tends to a cubic dependence on temperature:

$$c_v \longrightarrow \frac{2\pi^2}{5} k_B \left(\frac{k_B T}{\hbar c}\right)^3 \tag{8.10}$$

In spite of being the sum of a large number of Einstein-like single-mode contributions, the total specific heat has a low-temperature (8.10) behaviour different from the Einstein behaviour(8.3).

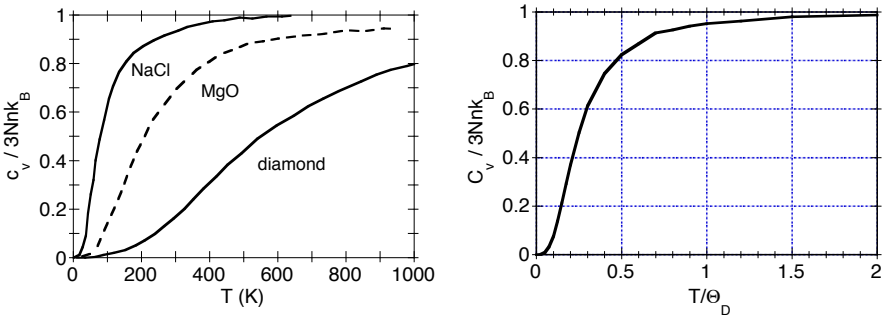

Figure 8.1: Left: comparison of specific heats of different crystals. Right: Debye specific heat as a function of the reduced temperature T/Θ_D.

One-dimensonal model

The physical origin of the low-temperature behaviour of specific heats (8.10) can be better grasped by considering the simplest case of the monatomic one-dimensional crystal.
At any temperature T, only the modes with frequency $\omega < \omega' = k_B T/\hbar$ are significantly excited, and their energy content is to a good approximation $\simeq k_B T$.
If T is sufficiently low, the dispersion relation $\omega(k)$ is linear, and the frequency values ω are equally spaced as the q values are; let the spacing between adjacent frequency values be $\delta\omega$. The number of thermally excited modes is thus $\simeq \omega'/\delta\omega = k_B T/\hbar\delta\omega \propto T$.
The total energy is thus proportional to T^2, and the specific heat of the one-dimensional system is proportional to T.
Note that at sufficiently high temperature all normal modes become excited (classical approximation); the number of excited modes is thus independent of temperature and the energy stored is now proportional to T, not to T^2.

The reasoning for the low-temperature behaviour can be extended to two and three-dimensional crystals, where the number of excited modes at low temperature is proportional to T^2 and T^3, respectively.

The vibrational density of states

The total energy and the heat capacity only depend on the frequency distribution, not on the eigenvectors of the dynamical matrix nor on the $\vec{q}$ directions. It is thus convenient to introduce a vibrational density of states (VDOS) $g(\omega)$ that is a function only of the frequency ω of normal modes, and is defined by

$$g(\omega)\, d\omega \;=\; \frac{V}{(2\pi)^3} \sum_s d\vec{q}\, \delta\left[\omega - \omega(\vec{q}s)\right] . \tag{8.11}$$

The relation between dispersion curves and VDOS is illustrated in Fig. 8.2 for the case of Germanium, where both acoustic and optic branches are present. In Fig. 8.3 the VDOS of Cu (only acoustic modes) are compared with the VDOS of CdTe (3 acoustic branches and 3 optic branches).
The total energy (8.4) can be expressed as

$$E \;=\; \int_0^\infty \left[\frac{1}{2} + \frac{1}{e^{\hbar\omega/k_B T} - 1}\right] \hbar\omega\, g(\omega)\, d\omega\,, \tag{8.12}$$

and the heat capacity is the integral of the Einstein-like one-mode contributions (8.2) weighted by the VDOS $g(\omega)$:

$$C_v \;=\; \left(\frac{\partial E}{\partial T}\right)_v \;=\; k_B \int_0^\infty \left(\frac{\hbar\omega}{k_B T}\right)^2 \frac{e^{\hbar\omega/k_B T}}{(e^{\hbar\omega/k_B T} - 1)^2}\; g(\omega)\, d\omega\,. \tag{8.13}$$

The calculation of the heat capacity through (8.13) requires that the VDOS $g(\omega)$ is known.

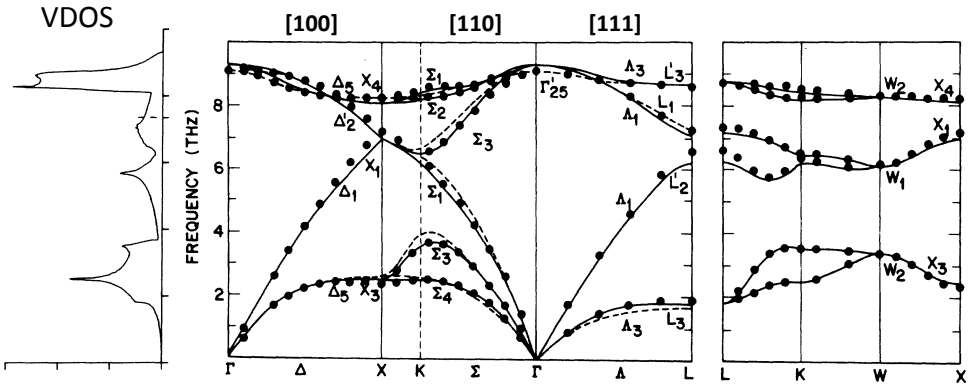

Figure 8.2: Relation between dispersion curves and VDOS $g(\nu)$ for Germanium.

The Einstein model

The Einstein approximate model consists in considering $3N$ normal modes with the same frequency ω_E, say $g(\omega) = \delta(\omega - \omega_E)$. The total vibrational energy of the crystal depends on temperature according to

$$\langle E \rangle \;=\; 3N\,\hbar\omega_E \left[\frac{1}{2} + \frac{1}{e^{\hbar\omega_E/k_BT} - 1}\right] . \tag{8.14}$$

The heat capacity in the Einstein model is

$$C_v \;=\; 3Nk_B \left(\frac{\hbar\omega_E}{k_BT}\right)^2 \frac{e^{\hbar\omega_E/k_BT}}{\left(e^{\hbar\omega_E/kT} - 1\right)^2} . \tag{8.15}$$

The Einstein model doesn't reproduce the T^3 low-temperature behaviour. For the interpretation of specific heats, the Einstein model has only historical interest.

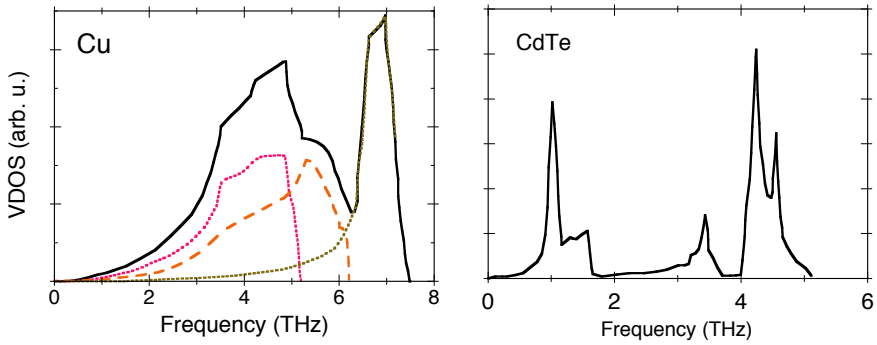

Figure 8.3: Left: VDOS for Cu; the contributions of TA modes give the VDOS below 6.3 THz, the contribution of LA modes gives the peak around 7 THz. Right: VDOS for CdTe; the contributions of TA modes (below 2 THz), LA modes (below 4THz) and optic modes (above 4 THz) can be easily distinguished.

8.1.3 The Debye model

The Debye approximate model is based on the following assumptions.

1. All phonon branches (acoustic + optic) are substituted by three acoustic branches with the same linear dispersion relation $\omega = c_s q$. (In refined treatments, three different sound

velocities are sometimes considered). Optic branches are transformed into extensions of the acoustic branches by a proper BZ zone extension.

2. The first BZ is substituted by a Debye sphere, whose volume is
$$V_D = \frac{4}{3}\pi q_D^3 = \frac{(2\pi)^3}{V_a} \tag{8.16}$$
where q_D is the radius of the sphere and V_a is the volume per atom in the real-space, corresponding to the volume of the primitive cell for monatomic lattices (such as Cu).
3. To the radius q_D of the Debye sphere there corresponds a Debye frequency $\omega_D = c_s q_D$, such that, taking into account of (8.16),
$$\omega_D^3 = \frac{6\pi^2 c_s^3}{V_a}\,. \tag{8.17}$$
4. Since the dispersion relation is linear, it is easy to see that the VDOS is parabolic. To this purpose, let us consider a sphere of radius $q = \omega/c_s$ in the reciprocal space, whose volume is $4\pi q^3/3 = 4\pi\omega^3/3c_s^3$. The number of wavevectors within the sphere is the product of the density times the volume
$$N = \frac{V}{(2\pi)^3}\,\frac{4}{3}\pi\frac{\omega^3}{c_s^3}\,, \tag{8.18}$$
so that the density of states VDOS is
$$g(\omega) \;=\; 3\,\frac{dN}{d\omega} \;=\; 3\frac{V}{(2\pi)^3}\,4\pi\,\frac{\omega^2}{c_s^3} \;=\; \frac{3V}{2\pi^2}\,\frac{\omega^2}{c_s^3} \tag{8.19}$$
where the factor 3 takes into account the number of acoustic branches.
Substituting $c_s = \omega_D/q_D$, taking into account (8.18) and remembering that $V/V_a = Nn$, one finds finally
$$g(\omega) \;=\; \frac{9Nn}{\omega_D^3}\,\omega^2 \tag{8.20}$$
One verify easily that $\int g(\omega)\,d\omega = 3Nn$, equal to the total number of degrees of freedom.

In spite of its quite crude approximations, the Debye model is still a useful reference, and the Debye temperature is a frequently used parameter.

Debye heat capacity

By substituting the Debye density of states (8.20) into the general expression for the heat capacity (8.13) one obtains
$$C_v \;=\; 9Nn\,k_B \int_0^{\omega_D} \left(\frac{\hbar\omega}{k_B T}\right)^2 \frac{e^{\hbar\omega/k_B T}}{(e^{\hbar\omega/k_B T}-1)^2}\,\frac{\omega^2}{\omega_D^3}\,d\omega\,. \tag{8.21}$$
It is convenient to define the new variable of integration, and modify accordingly the integration limit
$$x \;=\; \frac{\hbar\omega}{k_B T}\,, \qquad x_D \;=\; \frac{\hbar\omega_D}{k_B T} \;=\; \frac{\Theta_D}{T} \tag{8.22}$$
where the Debye temperature has been introduced
$$\Theta_D = \hbar\omega_D/k_B\,. \tag{8.23}$$
By substituting (8.22) into (8.21) one gets finally
$$C_v \;=\; 9Nn\,k_B \left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} \frac{x^4\,e^x}{(e^x-1)^2}\,dx\,. \tag{8.24}$$
The behaviour of the Debye specific heat is shown in Fig. 8.1, right.
The Debye model only depends on one parameter, the Debye frequency ω_D or the Debye temperature Θ_D. In spite of the differences between the Debye VDOS and the realistic VDOS, the reproduction of the specific heats is quite good also for non-Bravais crystals, say crystals exhibiting optical branches. The low-temperature T^3 behaviour is reproduced by the Debye model.

Meaning of the Debye temperature

The Debye model is used in a number of physical problems, of which the specific heat is only one example, perhaps the most important. The Debye temperature is a frequently used parameter, with different, although correlated, phenomenological interpretations:

a) separation between the low-T region where quantum effects are relevant and the high-T region where the classical approximation is suitable;
b) separation between the low-T region where the excited modes can be treated by the continuum approximation and the high-T region where the discrete description is necessary;
c) separation between the low-T region where the thermal contribution of long-wavelength modes prevails and the high-T region where all modes contribute on equal foot;
d) measure of the average strength of interatomic forces: higher Θ_D correspond to higher average frequencies and stronger average interatomic forces

Note that Debye temperatures obtained from different techniques (specific heat measurements, Debye-Waller factors of diffraction, etc) can be significantly different for the same system.

Table 8.1: Examples of Debye temperatures and their correlation with other physical properties.

	Θ_D (K)	Cohesive energy (eV/atom)	melting point (K)
Diamond	1860	5.81	3820
Silicon	625	4.63	1683
Copper	315	3.49	1358
Gold	170	3.81	1337
Lead	88	2.03	601

8.1.4 Anharmonicity effects

The main characteristics of the temperature dependence of specific heats can be accounted for within the harmonic approximation.
Only at high temperatures can the experimental behaviour deviate from the harmonic theory. Anharmonicity effects produce a slight linear increase with respect to the harmonic theory.

8.2 Thermal expansion

Anharmonicity is neccessary to account for thermal expansion. As we will see below, a perfectly harmonic crystal cannot undergo thermal expansion.
For one-dimensional systems, such as two-atomic molecules, thermal expansion is positive (§ 7.1). For crystals, where the potential energy is defined within a many-dimensional configuration space, thermal expansion is not necessarily positive.

8.2.1 Thermodynamics of thermal expansion

The coefficient of volume thermal expansion is defined as

$$\beta = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p . \tag{8.25}$$

The coefficient of linear thermal expansion is defined as

$$\alpha = \frac{1}{\ell}\left(\frac{\partial \ell}{\partial T}\right)_p . \tag{8.26}$$

For cubic crystals, by considering V as the volume of the conventional cell and $\ell = a$ the cell parameter, it is easy to see that $\beta = 3\alpha$.

The Grüneisen function

By exploiting some standard thermodynamic identities (see § 8.4.2) one can show that

$$\beta = \chi_T \left(\frac{\partial p}{\partial T}\right)_V = \chi_T \frac{C_v}{V} \left[\frac{\partial p}{\partial (U/V)}\right]_V = \chi_T \frac{c_v}{v} \left[\frac{\partial p}{\partial (U/V)}\right]_V , \tag{8.27}$$

where C_v is the constant-volume heat capacity, c_v is the constant-volume molar specific heat, U/V is the energy per unit volume, v is the molar volume and χ_T is the isothermal compressibility

$$\chi_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T . \tag{8.28}$$

Both specific heat and isothermal compressibility cannot be negative, in order to guarantee the stability of thermodynamic equilibrium.

The last factor in (8.27) is called Grüneisen function

$$\gamma(T) = \left[\frac{\partial p}{\partial (U/V)}\right]_V . \tag{8.29}$$

Equation (8.27) can thus be rewritten as

$$\beta = \chi_T \frac{c_v}{v}\gamma . \tag{8.30}$$

Physically, (8.27) shows that the variation of volume induced by the variation of temperature can be decomposed into three logical steps:

1. An increase of temperature corresponds to an increase of internal energy, measured by the specific heat c_v (which is never negative).
2. The increase of internal energy produces a variation of internal pressure that can be positive or negative and is measured by the Grüneisen function.
3. Since the crystal is in equilibrium with its surroundings at constant pressure, the increase of internal pressure when $\gamma > 0$ must be compensated by a reduction obtained through a volume increase (χ_T is never negative); viceversa, a reduction of pressure when $\gamma < 0$ must be compensated by an increase obtained through a reduction of volume.

An example of the relation between the four quantities of (8.27) is given in Fig. 8.4.

8.2.2 Vibrational dynamics and thermal expansion

Let us now seek a connection between the thermodynamical description of thermal expansion and the vibrational microscopic properties of crystals.

To that purpose, we will build on the harmonic approximation. We will show that the harmonic approximation cannot account for thermal expansion. On the other hand, the anharmonic perturbation theory is prohibitively intricate, even to first order.

It is however possible, for weakly anharmonic systems, to account for thermal expansion by a slight modification of the harmonic theory, which consists in letting the frequencies of normal modes to depend on the volume (quasi-harmonic approximation).

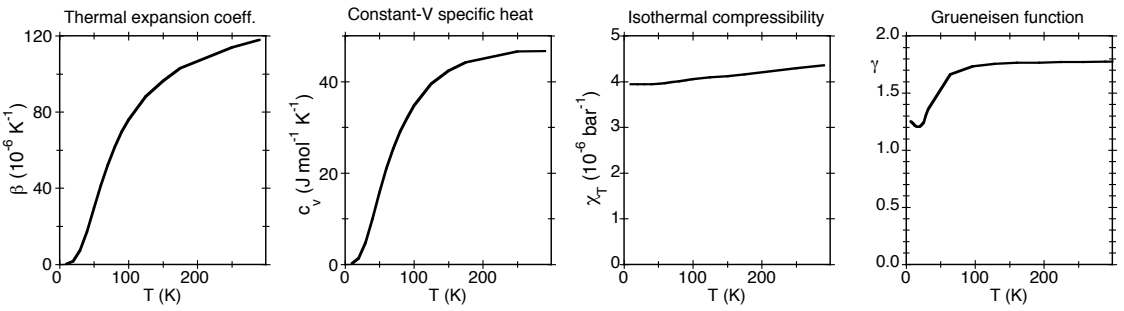

Figure 8.4: Thermal expansion of NaCl. From left to right: volume thermal expansion coefficient β, constant-volume molar specific heat c_v, isothermal compressibility χ_T and Grüneisen function $\gamma(T)$.

Dependence of pressure on temperature

According to the first equality of (8.27), the coefficient of thermal expansion β is proportional to the derivative of pressure with respect to temperature. To study the connection between thermal expansion and vibrational properties, it is convenient to find the dependence of pressure on the normal mode frequencies.
To this purpose, let us start from the Helmholtz function $F = U - TS$, whose differential is $dF = -S\,dT - p\,dV$. The pressure is

$$p = -\left(\frac{\partial F}{\partial V}\right)_T . \tag{8.31}$$

The Helmholtz function is connected to the partition function Z by $F = -k_B T \ln Z$.
According to (7.127), the partition function for the vibrational degrees of freedom of a harmonic crystal is

$$Z_{\text{vib}} = \prod_{\vec{q}s} \frac{e^{-\beta\hbar\omega(\vec{q}s)/2}}{1 - e^{-\beta\hbar\omega(\vec{q}s)}} . \tag{8.32}$$

The vibrational contribution to the Helmholtz function (Fig. 8.5) is

$$F_{\text{vib}} = -k_B T \ln Z_{\text{vib}} = \sum_{\vec{q}s} \left\{ \frac{1}{2}\hbar\omega(\vec{q}s) + k_B T \ln\left[1 - e^{-\beta\hbar\omega(\vec{q}s)}\right] \right\} . \tag{8.33}$$

By adding the static contribution to internal energy, the Helmholtz function becomes $F = U_0 + F_{\text{vib}}$ and pressure is

$$p = -\left(\frac{\partial F}{\partial V}\right)_T = -\left(\frac{\partial U_0}{\partial V}\right)_T - \frac{1}{2}\sum_{\vec{q}s}\left(\frac{\partial\hbar\omega(\vec{q}s)}{\partial V}\right)_T - \sum_{\vec{q}s}\frac{1}{e^{\beta\hbar\omega(\vec{q}s)} - 1}\left(\frac{\partial\hbar\omega(\vec{q}s)}{\partial V}\right)_T . \tag{8.34}$$

The last step is to evaluate the derivative of pressure with respect to temperature

$$\left(\frac{\partial p}{\partial T}\right)_V . \tag{8.35}$$

and insert it in (8.27). The first two sums on the right of (8.34), related to the static energy and to the zero-point energy, are independent of temperature. So, we have to evaluate the derivative of the last sum.
Before that, let us make two simple but important preliminary remarks.

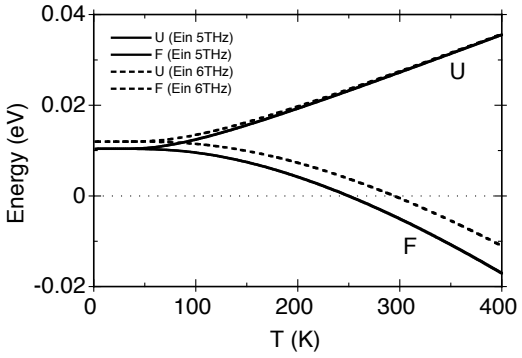

Figure 8.5: Internal energy U and Helmholtz function F for harmonic oscillators of different frequencies: 5 THz (continuous line), 6 THz (dashed line).

First remark: If the frequency is independent of volume

The temperature dependence of pressure is contained in the last sum of (8.27), where each temperature-dependent term is weighted by the volume derivative of the corresponding frequency. If the frequencies are independent of volume, then $(\partial p/\partial T)_V = 0$ and there is no thermal expansion, $\beta = 0$.

For the a two-atomic molecule, say a one-dimensional harmonic oscillator (Section 7.1), the frequency is independent of the amplitude of oscillations and there is no thermal expansion. For anharmonic oscillators, instead, the frequency is a function of the amplitude, according to (7.27) on page 167, and there is positive thermal expansion.

How can these remarks be extended to many-atomic systems such as crystals, where the harmonic approximation is related to the Born-von Karman expansion of the whole crystal potential energy?

Second remark: A harmonic crystal has zero thermal expansion

Let us consider the expansion of the crystal potential energy for a harmonic crystal:

$$V \;=\; V_0 + V_2 \;=\; V_0 \;+\; \frac{1}{2} \sum_{\substack{\alpha \ell \kappa \\ \alpha' \ell' \kappa'}} \Phi \begin{pmatrix} \alpha\, \ell\, \kappa \\ \alpha' \ell' \kappa' \end{pmatrix} u_\alpha(\ell\kappa)\, u_{\alpha'}(\ell'\kappa')\,, \tag{8.36}$$

where the atomic displacements are evaluated with respect to the average positions, which for a harmonic crystal coincide with the rest positions.

Let us now suppose that the crystal is strained by applying a force to each atom $(\ell\kappa)$ with components $F_\alpha(\ell\kappa)$, so that the volume of the strained crystal is modified and the potential energy becomes

$$V^* \;=\; V_0 \;-\; \sum_{\alpha\ell\kappa} F_\alpha(\ell\kappa) u_\alpha(\ell\kappa) \;+\; \frac{1}{2} \sum_{\substack{\alpha \ell \kappa \\ \alpha' \ell' \kappa'}} \Phi \begin{pmatrix} \alpha\, \ell\, \kappa \\ \alpha' \ell' \kappa' \end{pmatrix} u_\alpha(\ell\kappa)\, u_{\alpha'}(\ell'\kappa')\,, \tag{8.37}$$

The atoms of the strained crystal vibrate around new average positions. Let $\vec{\delta}(\ell\kappa)$ be the displacement of the atom $\ell\kappa$ from the old to the new average position, and $\vec{u}^*(\ell\kappa)$ be the atomic displacement with respect to the new average positions. The displacement of each atom with respect to the old average positions is then expressed as

$$u_\alpha(\ell\kappa) \;=\; \delta_\alpha(\ell\kappa) + u^*_\alpha(\ell\kappa)\,. \tag{8.38}$$

By substituting the expressions (8.38) for all atoms in (8.37), one finds that the second-order term V_2^* with respect to the new average positions has the same force constants.

$$V_2^* \;=\; \frac{1}{2} \sum_{\substack{\alpha \ell \kappa \\ \alpha' \ell' \kappa'}} \Phi \begin{pmatrix} \alpha\, \ell\, \kappa \\ \alpha' \ell' \kappa' \end{pmatrix} u^*_\alpha(\ell\kappa)\, u^*_{\alpha'}(\ell'\kappa')\,, \tag{8.39}$$

As a consequence, the eigen-frequencies $\omega(\vec{q}s)$ are not modified by the change of average positions, say by the change of volume, the last sum of (8.27) is zero, the derivative of pressure with respect to temperature is zero and $\beta = 0$.

Note: To grasp the basic idea, let us consider a one-dimensional harmonic oscillator represented by a mass attached to a spring. The force constant and the corresponding frequency are not modified if the system oscillates in a horizontal direction or in a vertical direction, although in the vertical direction the spring is strained by gravity.

Mode Grüneisen parameters

Let us now come back to (8.27) and insert the expression of pressure (8.34) into (8.27):

$$\begin{aligned} \beta \;=\; \chi_T \left(\frac{\partial p}{\partial T}\right)_V \quad &= \quad -\chi_T \sum_{\vec{q}s} \left[\frac{\partial}{\partial T}\left(\frac{1}{e^{\beta\hbar\omega(\vec{q}s)}-1}\right)\frac{\partial\,\hbar\omega(\vec{q}s)}{\partial V}\right] \\ &= \quad -\chi_T \sum_{\vec{q}s} \left[C_v(\vec{q}s)\,\frac{1}{\hbar\omega(\vec{q}s)}\,\frac{\partial\,\hbar\omega(\vec{q}s)}{\partial V}\right] \\ &= \quad \chi_T\,\frac{1}{v}\sum_{\vec{q}s} c_v(\vec{q}s)\,\gamma(\vec{q}s)\,. \end{aligned} \tag{8.40}$$

In the last line of (8.41), the quantities

$$\gamma(\vec{q}s) \;=\; -\frac{\partial \ln \omega(\vec{q}s)}{\partial \ln V} \tag{8.41}$$

are the mode Grüneisen parameters.
By comparing (8.41) with (8.30), one sees that the Grünesien function is connected to the mode Grüneisen parameters by

$$\gamma(T) \;=\; \frac{\sum_{\vec{q}s} c_v(\vec{q}s)\,\gamma(\vec{q}s)}{\sum_{\vec{q}s} c_v(\vec{q}s)}\,. \tag{8.42}$$

- Normal modes with positive Grüneisen parameters, for which an increase of volume is accompanied by a reduction of frequency and viceversa, give a positive contribution to thermal expansion.
- Normal modes with negative Grüneisen parameters, for which an increase of volume is accompanied by an increase of frequency and viceversa, give a negative contribution to thermal expansion.

The temperature dependence of the Grüneisen function is thus determined by the progressive population of modes with different Grüneisen parameters.

Quasi-harmonic approximation

The above treatment of thermal expansion in terms normal vibrational modes represents the so-called *quasi-harmonic approximation.*
The quasi-harmonic approximation consists in maintaining the description of atomic vibrations in terms of normal modes, and taking into account anharmonicity effects by a variation of the frequency of each normal mode as a function of the volume through the mode Grüneisen parameters.

8.2.3 Bond expansion and lattice expansion

In § 7.1, we have considered the effects of anharmonicity on a two-atomic molecule (one-dimensional oscillator). In this § 8.2 we have considered the volume expansion (or equivalently the linear lattice expansion) of a crystal.
It is of interest, for crystals, to compare the expansion of the lattice and the expansion of single inter-atomic bonds.

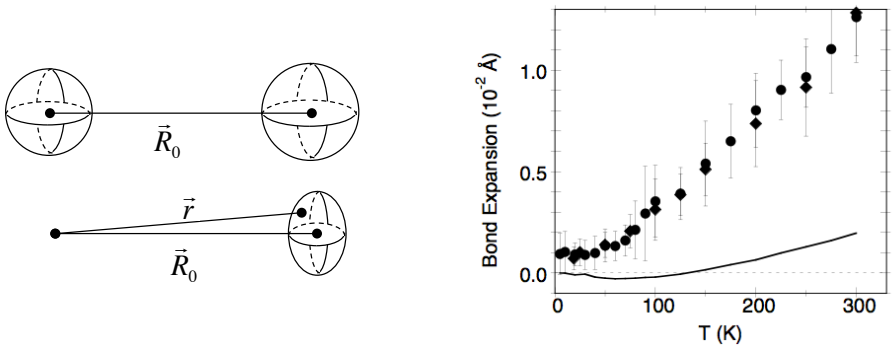

Figure 8.6: Left: Relation between equilibrium bond distance $\vec{R}_0$ and instantaneous bond distance $\vec{r}$ between neighbouring atoms; top: ellipsoids of atmic displacements; bottom: ellipsoid of relative displacements. Right: Thermal expansion of the average nearest-neighbour distance $\langle r \rangle$ in CdTe measured by EXAFS spectroscopy (symbols); the continuous line is the thermal expansion of the crystallographic distance R_c between the average atomic positions.

1. The lattice thermal expansion (expressed by the volume or linear expansion coefficients) depends on the average behaviour of all the atoms. The lattice thermal expansion can be measured by elastic scattering experiments (e.g. X-ray diffraction) or by macroscopic techniques, such as dilatometry.
   The lattice thermal expansion is connected to the anharmonicity of the total crystal potential energy expanded in power series of the atomic displacements $u_\alpha(\ell\kappa)$.
2. The thermal expansion of the bond between any pair of atoms can be measured by a suitable local probe, such as EXAFS (see Section 11.8).
   The bond expansion is connected to the anharmonicity of a one-dimensional pair potential expanded in power series of the variation of interatomic distance $u = r - R_0$. Note that we should consider here not the single pair potential, but an effective pair potential, taking into account the effect of interaction with all the atoms of the crystal.

Bond distance

Let us consider a pair of atoms A and B embedded in a crystal and let $R_c = |\langle \vec{r}_B \rangle - \langle \vec{r}_A \rangle|$ be the distance between the average atomic positions (crystallographic distance), which for simple crystals is proportional to the lattice parameters measured by elastic scattering. Its thermal expansion is proportional to the lattice thermal expansion and is quantified by the linear expansion coefficient.
The instantaneous bond distance (Fig. 8.6, left) is connected to the distance R_c by

$$\vec{r} \;=\; \vec{R}_c + \vec{u}_B - \vec{u}_A \;=\; \vec{R}_c + \Delta\vec{u}\,, \tag{8.43}$$

where $\vec{u}_B$ and $\vec{u}_A$ are the instantaneous displacements of the two atoms and $\Delta\vec{u} = \vec{u}_B - \vec{u}_A$ is the instantaneous relative displacement.
One can easily verify that the instantaneous scalar bond distance r can be expressed as

$$r \;=\; \left[\left(\vec{R}_c + \Delta\vec{u}\right)^2\right]^{1/2} \;\simeq\; R_c + \Delta u_\parallel + \frac{\Delta u_\perp^2}{2R_c} - \frac{\Delta u_\parallel \Delta u_\perp^2}{2R_c^2} + \frac{4\Delta u_\perp^4}{8R^3} + \cdots\,, \tag{8.44}$$

where $\Delta u_\parallel$ and $\Delta u_\perp$ are the projections of the relative displacement $\Delta\vec{u}$ along the bond direction and in the perpendicular plane.
Let us now consider the canonical average of (8.44), limited to the leading terms in atomic relative displacements:

$$\langle r \rangle \;\simeq\; R_c + \langle \Delta u_\parallel \rangle + \frac{\langle \Delta u_\perp^2 \rangle}{2R_c} \;=\; R_c + \frac{\langle \Delta u_\perp^2 \rangle}{2R_c}\,. \tag{8.45}$$

The average bond distance $\langle r \rangle = \langle |\vec{r}_B - \vec{r}_A| \rangle$ is larger that the crystallographic distance R_c owing to the effect of vibrations perpendicular to the bond direction, quantified by $\langle \Delta u_\perp^2 \rangle$. Since the amplitude of perpendicular vibration increases with temperature, the thermal expansion of the bond distance $\langle r \rangle$ is larger than the expansion of the crystallographic distance R_c.

A quantitative example of the difference between $\langle r \rangle$ and R_c and their temperature dependencies is shown in Fig. 8.6, right.

Pair potential and crystal potential

Let us suppose that the effective pair potential is perfectly harmonic, say quadratic with respect to the variations of the inter-atomic distance:

$$V = v_0 + k(r - R_c)^2/2 \,. \tag{8.46}$$

If each distance variation $r - R_c$ in (8.46) is expressed in terms of the atomic displacements $\vec{u}$, according to (8.44), it is easy to understand that higher-order terms in the power expansion of the crystal potential are introduced. The crystal potential is thus anharmonic, even if the effective pair potential is harmonic.

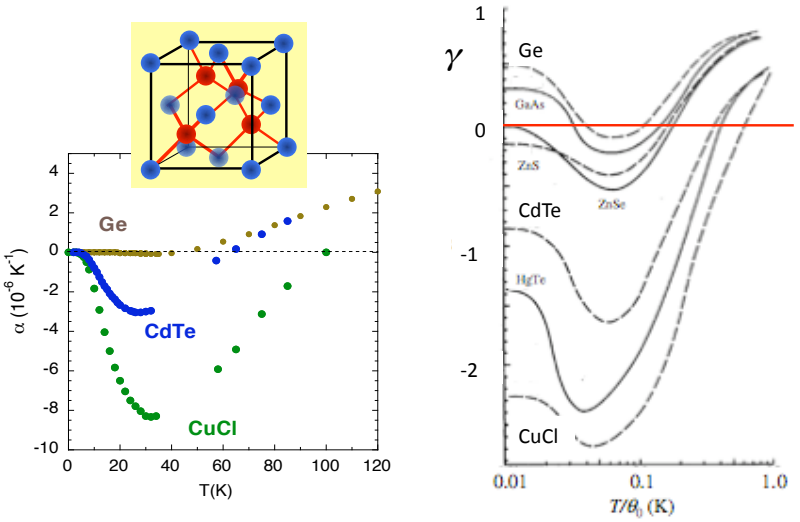

Figure 8.7: Linear expansion coefficient α (left) and Grünesien functions for some crystals with the diamond or zincblende structures.

8.2.4 Negative thermal expansion

Positive and negative contributions to thermal expansion

The lattice thermal expansion can be considered as the net effect of two opposite contributions, one positive and the other negative.
When, as it happens for the large majority of cases, the positive contribution prevails over the negative one, the crystal undergoes positive thermal expansion $\beta > 0$.
When the negative contribution prevails, the crystal undergoes negative expansion.
The two opposite contributions can be described from two different, and non equivalent, points of view, one based on a real space approach, the other on a reciprocal space approach.

In the first approach, the thermal expansion of the nearest-neigbour average distance $\langle r \rangle$ (bond thermal expansion) is expected to be positive as effect of the anharmonicity of the bond-strecthing potential energy.
The lattice thermal expansion is the sum of the following two opposite contributions:

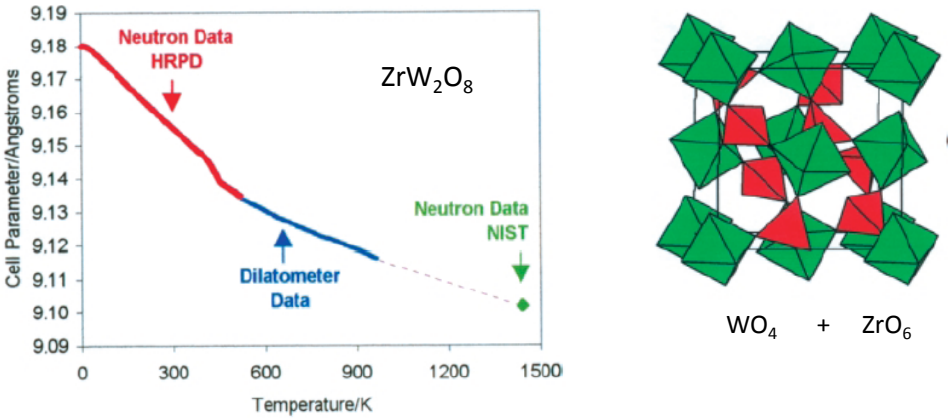

Figure 8.8: Left: thermal expansion of the cell parameter of ZrW_2O_8, from Mary, Evans, Vogt and Sleight, Science, 272 (1996). Right: sketch of the framework stcructure.

1. The positive contribution due to the bond stretching, corresponding to the expansion of the bond distance $\langle r \rangle$.
2. A negative contribution due to the relative vibrations perpendicular to the bond, measured by $\langle \Delta u_\perp^2 \rangle$, often referred to as "tension effects".

Within the alternative approach, based on the vibrational dynamics within the quasi-harmonic approximation, the positive contribution to the lattice thermal expansion is due to normal modes with positive Grüneisen parameters, and the negative contribution to thermal expansion is due to normal modes with negative Gruüneisen parameters. A crystal undergoes negative expansion when the total contribution of modes with negative Grüneisen parameters prevails over the total contribution of modes with positive Grüneisen parameters.

Examples of negative thermal expansion

Several crystals with the relatively open diamond or zincblende structures (such as Si, Ge, GaAs, CdTe) exhibit negative expansion in a limited low-temperature interval (Fig. 8.7). The negative contribution to expansion is attributed to some TA normal modes at the BZ boundary with negative Grüneisen parameters.
Some crystals exhibit negative expansion in very large temperature intervals. A typical example is ZrW_2O_8, whose lattice parameter contracts from 0 to 1400 K (Fig. 8.8). These crystals are characterized by framework structures, composed by tetrahedral and/or octahedral structural units joined by the corners. The rotation of structural units around the joining corners can give rise to strong tension effects with negative contribution to thermal expansion.

8.3 Thermal conductivity

Thermal conductivity in insulating crystals is only due to lattice vibrations. In metals, the contribution of electrons has to be added. We consider here only the lattice contribution to thermal conductivity.

8.3.1 Definition of thermal conductivity

The local flux of thermal energy along a one-dimensional system of cross-section A,

$$J_{\rm th} = \frac{1}{A}\frac{đQ}{dt}\,, \tag{8.47}$$

depends on the temperature gradient according to

$$J_{\rm th} \;=\; -K_{\rm th}\,\frac{dT}{dx}\,, \tag{8.48}$$

where $K_{\rm th}$ is the coefficient of thermal conductivity, measured in $\mathrm{W\,(m\,K)^{-1}}$.
The inverse of the coefficient of conductivity is the thermal resistivity.

Note: For a *three-dimensional case*, the vector flux $\vec{J}_{\rm th}$ is connected to the temperature gradient $\vec{\nabla}T$ by a relation similar to (8.48),

$$\vec{J}_{\rm th} = -\mathbf{K}_{\rm th}\vec{\nabla}T\,, \tag{8.49}$$

where now the thermal conductivity $\mathbf{K}_{\rm th}$ is a tensor quantity.

Measurement of thermal conductivity

The coefficient $K_{\rm th}$ can be measured by setting up a constant flux $J_{\rm th}$ along the rod and measuring the temperatures T_1 and T_2 at two different points x_1 and x_2, respectively:

$$|K_{\rm th}| \;=\; \frac{J_{\rm th}\,\Delta x}{\Delta T}\,, \tag{8.50}$$

The limiting case of infinite thermal conductivity (zero thermal resistivity) corresponds to one of the two situations:

1. No temperature variation is measured, $\Delta T = 0$, for a finite flux $J_{\rm th}$.
2. For $\Delta T \neq 0$, the flux $J_{\rm th}$ is only limited by the power released by the heat source.

Thermodynamical considerations

The flow of heat is an irreversible process. A system affected by thermal flow is out of thermodynamical equilibrium.
We consider here only stationary (steady-state) heat flows, $J_{\rm th}$ = constant, without sources or sinks within the rod, so that the input and output fluxes are equal: $J_{\rm in} = J_{\rm out} = J_{\rm th}$.
Regions of relatively small size of the system can be treated as if they were in a state of local thermodynamical equilibrium, so that temperature can be locally defined and measured.

Types of thermal fluxes

Three mechanisms of heat transfer can be distinguished

1. conduction, transfer of heat in material systems without transfer of matter (possible for both solids and fluids),
2. convection, transfer of heat together with transport of matter (possible for fluids, not for solids),
3. radiation, transfer of heat without any material support.

In the following, we consider only the mechanism of heat transfer by conduction.

8.3.2 Thermal conductivity in molecular gases

The formal treatment of thermal conductivity in crystals can be better understood if a preliminary account of thermal conductivity in molecular gases is given. (We use here the term "molecular gases" to distinguish from the "phonon gases", which will be introduced below).

We refer to a gas enclosed in a rod-shaped vessel, whose extrema are maintained at different temperatures $T_{\rm in}$ and $T_{\rm out}$ ($T_{\rm in} > T_{\rm out}$) (Fig. 8.9). The number of molecules inside the vessel is constant and, for a steady state of heat transfer in the absence of convection, there is no drift velocity, say no net movement of matter from left to right or viceversa. We assume conditions of local equilibrium, so that local values of temperature T, number density n and average velocity $\langle v\rangle$ are meaningful.

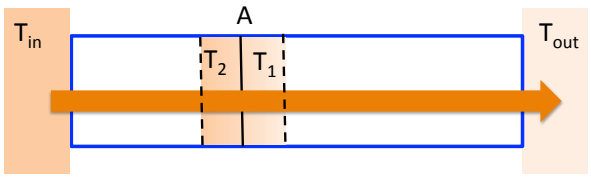

Figure 8.9: A rod-shaped vessel (blue) containing a molecular gas and connected, at the two extrema, to two reservoirs at temperatures $T_{\rm in}$ and $T_{\rm out}$ supports a steady-state flux of heat $J_{\rm th}$ from the hot face to the cold face.

Approximate theory

It is reasonable to assume that:

1. The steady-state flux of heat $J_{\rm in} = J_{\rm th} = J_{\rm out}$ is due to the different amount of average energy carried by the molecules moving from the high-T regions to the low-T regions with respect to the molecules moving from the low-T to the high-T regions.
2. The thermalisation at different temperatures of the different regions along the vessel is guaranteed by the collisions (say scattering processes) between the gas molecules.

Let us consider a cross section A of the vessel and two regions immediately to the left and to the right, with average local temperatures T_2 and T_1, respectively (Fig. 8.9). In conditions of steady-state heat flux, the surface A is crossed by an equal number of molecules N_2 from left to right and N_1 from right to left in the unit time.
Let us further assume, for the moment, that the molecules move along the rod with a velocity only depending on the position (say we neglect for the moment the thermal distribution of velocities). This means that

$$N_2 = N_1 \quad \Rightarrow \quad n_2 v_{2x} = n_1 v_{1x} = n v_x \,, \tag{8.51}$$

where n_1 and n_2 are the number densities to the right and to the left of the surface A, v_{1x} and v_{2x} are the corresponding velocities and nv_x is the common product.
Taking into account the two possible directions of velocity v_x, the heat fluxes from left to right and from right to left crossing the surface A can be expressed as

$$J_1 = \frac{1}{2} n_1 v_{1x}\, \epsilon_1 = \frac{1}{2} n v_x\, \epsilon_1 \,, \qquad J_2 = \frac{1}{2} n_2 v_{2x}\, \epsilon_2 = \frac{1}{2} n v_x\, \epsilon_2 \,, \tag{8.52}$$

where ϵ_1 and ϵ_2 are the average energies of the molecules at the temperatures T_1 and T_2, respectively. The net heat flux crossing the surface A is thus

$$J_{\rm th} = J_2 - J_1 = \frac{1}{2}\, n\, v_x\, \Delta\epsilon \,. \tag{8.53}$$

Introducing the relaxation time τ, say the average time between two consecutive collisions, the energies ϵ_1 and ϵ_2 can be expressed in terms of the distance of the position of the last collision from the position x of the partition A:

$$\begin{aligned} \Delta\epsilon &= \epsilon_2 - \epsilon_1 = \epsilon(x - v_x\tau) - \epsilon(x + v_x\tau) \\ &= -2\,\frac{d\epsilon}{dx}\, v_x\tau = 2\,\frac{d\epsilon}{dT}\left(-\frac{dT}{dx}\right) v_x\tau \,, \end{aligned} \tag{8.54}$$

so that

$$J_{\rm th} = -c_v\, v_x^2\, \tau\, \frac{dT}{dx} \,, \tag{8.55}$$

where c_v is the specific heat per unit volume.
If now we consider the distribution of molecular velocities, so that $\langle v_x^2 \rangle = \frac{1}{3}\langle v^2 \rangle$, the heat flux can be expressed as

$$J_{\rm th} = -\frac{1}{3} c_v\, \langle v^2 \rangle\, \tau\, \frac{dT}{dx} \,. \tag{8.56}$$

Approximate formula

The foregoing treatment leads to the approximate expression

$$K_{\text{th}} = \frac{1}{3} c_v \bar{v} \ell = \frac{1}{3} c_v \langle v^2 \rangle \tau \tag{8.57}$$

where

a) c_v is the specific heat per unit volume (independent of T for monatomic gases); the specific heat connects the temperature gradient to the gradient of energy density;
b) $\bar{v} =$ is the root mean square molecular velocity, $\bar{v} = [\langle v^2 \rangle]^{1/2}$, proportional to $\sqrt{T}$; (for air at room temperature $\bar{v} \simeq 500$ m/s)
c) τ is the relaxation time, say the mean time between two consecutive collisions
d) $\ell = \tau \bar{v}$ is the mean free path between two consecutive collisions (for air at room temperature, $\ell \simeq 600$ Å)

Table 8.2: Coefficient of thermal conductivity for some selected gases.

Gas	H_2	O_2	CO_2	Ar	Kr
K_{th} ($\text{W m}^{-1}\text{K}^{-1}$)	0.176	0.024	0.014	0.016	0.008

The coefficient K_{th} for molecular gases

a) is quite independent of pressure, since when pressure increases the variation of c_v is compensated by the reduction of ℓ,
b) is roughly proportional to $\sqrt{T}$ because of the proportionality to $\bar{v}$,
c) is higher for lighter molecules, whose root mean square velocity $\bar{v}$ is lower, see Table 8.2.

8.3.3 Vibrational thermal conductivity in crystals

Let us now consider the thermal conductivity in crystals.

Inadequacy of the harmonic approximation

In the harmonic theory, the stationary eigenstates of the harmonic Hamiltonian are the normal modes. In thermodynamic equilibrium, the modes $\vec{q}$ and $-\vec{q}$ have the same energy content (the same number of phonons) and no net energy transport takes place. The transmission of thermal energy by normal modes can be accounted for by an unbalance between the number of phonons with wavevector $\vec{q}$ in the direction of heat conduction and with wavevector $-\vec{q}$ in the opposite direction.

One can consider a perfect harmonic crystal connecting two reservoirs at different temperatures. The perfect harmonic crystal would exchange phonons with the two reservoirs but would exhibit an infinite thermal conductivity, say zero thermal resistivity. A flux of heat carried by normal modes, which are intrinsically de-localised, could not cause a temperature gradient along the crystal.
There is some similarity with the energy flux carried by an electromagnetic wave which exchanges photons with a source and an absorber.

Real crystals exhibit finite thermal conductivities: the heat flux is accompanied by a temperature gradient. Besides, thermal conductivities exhibit peculiar temperature dependencies, some examples of which are given in Fig. 8.10.

Beyond the harmonic approximation

In order to account for the real behaviour of crystals, we need a mechanism by which *a)* different regions of the crystals are thermalised at different temperatures and *b)* heat enters from the hot reservoir into the crystal, is transferred along the crystal and finally is released to the cold reservoir. Such a mechanism is based on a particle-like approach in which the quanta of thermal energy (the phonons) are somewhat localised. By considering phonons as localised particles, which can be created or annihilated, one can achieve the following goals:

1. the exchange of energy of the crystal with the hot and cold external reservoirs can be described in terms of phonon creation and annihilation, respectively, within the crystal;
2. the transfer of thermal energy along the crystal can be described in terms of phonons; movement,
3. the local thermalisation can be attributed to collisions between phonons.

It is useful to highlight the similarities and the differences between the phonon gas here introduced and the molecular gas considered above:

a) common properties are the heat transport due to particle movement and the thermalisation due to particle collisions;
b) the main difference is that phonons can be created and annihilated at the boundaries of the crystal and in the collision processes.

Phonons cannot be strictly considered as particles; they only exist as collective excitations of a large number of interacting atoms. For that reason they are qualified as *quasi-particles*.

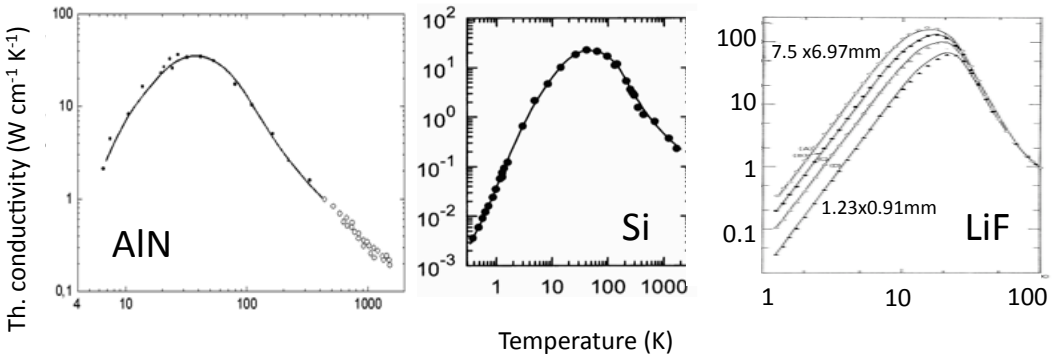

Figure 8.10: Thermal conductivity of different semiconductors as a function of temperature: AlN [G.A. Slack et al., J. Phys. Chem. Solids 48, 641 (1987)], Si [C. Glassbrenner and G. A. Slack, Phys. Rev. 134, A1058 (1964)] and LiF [P. D. Thatcher, Phys. Rev. 156, 975 (1967)]. For LiF, four specimens of different size have been considered.

Phonons as wave packets

The particle-like picture depicted above requires that phonons have a spatial localisation, which can be achieved by considering phonons as wave packets of spatial extent $\Delta\vec{r}$ in the real space and $\Delta\vec{q}$ in the reciprocal space, connected by the uncertainty relation $\Delta x \Delta q_x \simeq 1$.
Without going into details, it is clear that the size of a wave packet size in both the real and reciprocal spaces should undergo the following limitations:

a) in real space: atomic size $\ll$ phonon size $\Delta\vec{r} \ll$ crystal size
b) in reciprocal space: phonon size $\Delta\vec{q} \ll$ first BZ extension

Anharmonicity and phonon collisions

In order to account for phonon collisions, responsible for local thermalisation, it is necessary to consider the anharmonic terms of the crystal potential energy.
In Section 7.4 we have seen that the harmonic Hamiltonian can be expressed in terms of normal coordinates and momenta (eq. 7.113)

$$\mathcal{H} \;=\; \frac{1}{2}\sum_{\vec{q}s}\left[P(\vec{q}s)P^*(\vec{q}s)+\omega^2(\vec{q}s)Q(\vec{q}s)Q^*(\vec{q}s)\right]\,. \tag{8.58}$$

and that normal coordinates and momenta can in turn be expressed in terms of creation and destruction operators, so that (eq. 7.123)

$$\mathcal{H} \;=\; \sum_{\vec{q}s}\left[a^\dagger(\vec{q}s)\,a(\vec{q}s)+\frac{1}{2}\right]\hbar\omega(\vec{q}s) \;=\; \left[n(\vec{q}s)+\frac{1}{2}\right]\hbar\omega(\vec{q}s)\,. \tag{8.59}$$

The matrix elements of the harmonic Hamiltonian are zero between states with different number of phonons for a each normal mode.

By a similar procedure, one can show that the anharmonic terms in the Born-von Karman expansion of the crystal potential can be expressed as a function of normal coordinates; the third-order term becomes

$$V_3 = \frac{1}{3!N^{1/2}}\sum_{\substack{\vec{q}\vec{q}'\vec{q}''\\ ss's''}}\Psi\begin{pmatrix}\vec{q}\,\vec{q}'\,\vec{q}''\\ s\;\;s'\;\;s''\end{pmatrix}\,Q(\vec{q}s)\,Q(\vec{q}'s')\,Q(\vec{q}''s'')\;\;\Delta(\vec{q}+\vec{q}'+\vec{q}'') \tag{8.60}$$

where

$$\Delta(\vec{q}+\vec{q}'+\vec{q}'') \;=\; \frac{1}{N}\sum_{\ell}e^{i(\vec{q}+\vec{q}'+\vec{q}'')\cdot\vec{R}(\ell)} \;=\; \begin{cases}1 & \text{for } \vec{q}+\vec{q}'+\vec{q}''=0 \text{ or } \vec{G}\\ 0 & \text{otherwise}\,.\end{cases} \tag{8.61}$$

and

$$\Psi\begin{pmatrix}\vec{q}\,\vec{q}'\,\vec{q}''\\ s\;\;s'\;\;s''\end{pmatrix} = \sum_{\substack{\alpha\kappa\\ \alpha'\ell'\kappa'\\ \alpha''\ell''\kappa''}}\frac{\Phi_3}{\sqrt{m_\kappa m_{\kappa'}m_{\kappa''}}}\,e^{i\left[\vec{q}'\cdot\vec{R}(\ell'-\ell)+\vec{q}''\cdot\vec{R}(\ell''-\ell)\right]}\,\mathrm{e}_\alpha(\kappa|\vec{q}s)\mathrm{e}_{\alpha'}(\kappa'|\vec{q}'s')\mathrm{e}_{\alpha''}(\kappa''|\vec{q}''s'')\,. \tag{8.62}$$

The product of three normal coordinates of three different normal modes in (8.60) leads to the possibility of creation and destruction of phonons in the three different modes, say to the exchange of energy quanta between different normal modes.

The statistical effects of anharmonicity can be evaluated by a quantum statistical perturbation treatment, taking the harmonic approximation as unperturbed starting point (as was done for the single harmonic oscillator in Section 7.1).
The result of the perturbation theory is the exchange of energy, say of phonons, between normal modes:

- third-order anharmonic terms lead to transitions where three normal-mode occupation numbers are modified through the exchange of phonons (Fig. 8.11, left),
- fourth-order anharmonic terms lead to transitions where four normal-mode occupation numbers are modified through the exchange of phonons (Fig. 8.11, right),
- and so on.

In the particle-like picture of localised phonon wave packets, the processes of phonon exchanges between normal modes are described as *phonon collisions*. Note that the phonon collisions, contrary to molecules collisions, are events where number conservation has not to be fulfilled.
Phonon collisions fulfil the following conservation laws:

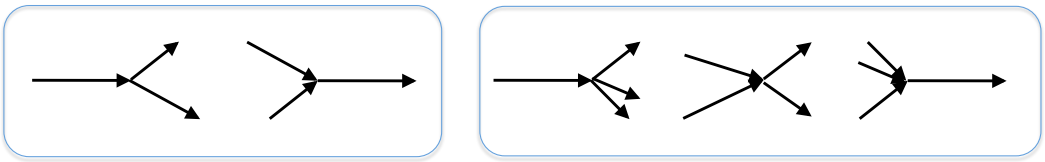

Figure 8.11: Effects of anharmonic perturbations: three-phonons events (left panel) and four-phonons events (right panel).

1. Conservation of energy

$$\sum_{\vec{q}s} \hbar\omega(\vec{q}s)\; n(\vec{q}s) \;=\; \sum_{\vec{q}s} \hbar\omega(\vec{q}s)\; n'(\vec{q}s) \tag{8.63}$$

   where n and n' refer to the number of phonons within a given normal mode before and after the collision.
2. Conservation of momentum

$$\sum_{\vec{q}s} \vec{q}\, n(\vec{q}s) \;=\; \sum_{\vec{q}s} \vec{q}\, n'(\vec{q}s) + \vec{G} \tag{8.64}$$

   where $\vec{G}$ is a reciprocal lattice vector. The presence of $\vec{G}$ is necessary to guarantee that all vectors $\vec{q}$ and $\vec{q}'$ lie in the 1st Brillouin Zone.

Approximate kinetic theory for thermal conductivity in crystals

Once the concepts of phonons as quasi-particles and of phonon collisions have been introduced, it is reasonable to extend the phenomenological theory of thermal conductivity for molecular gases to the phonon gas. To focus only on the most relevant properties of phonon heat conduction, let us here consider an approximate Debye model, where only acoustic modes with the same velocity c_s are present. The thermal conductivity of a crystal, by analogy with (8.57), can be expressed as

$$K_{\text{th}} \;=\; \frac{1}{3}\, c_v\, c_s\, \ell \;=\; \frac{1}{3}\, c_v\, c_s^2\, \tau \tag{8.65}$$

where:

a) The specific heat c_v connects the temperature gradient of (8.50) to the gradient of energy density. The T^3 low-temperature behaviour of c_v characterises the low-temperature behaviour of the thermal conductivity of crystals. At high temperature, c_v is constant.
b) The velocity c_s is temperature independent (Debye model).
d) The mean free path ℓ is inversely proportional to the collision rate $1/\tau$. The evaluation of the effective collision rate, which gives rise to the finite value of thermal conductivity, is far from trivial.

Effective collision rate and crystal momentum conservation

In principle, one can expect the collision rate to be proportional to the number of phonons present in the crystal: the higher the number of phonons, the higher the frequency of collisions and the shorter the mean free path between collisions.
At sufficiently high temperature, the equilibrium phonon occupation number is proportional to temperature:

$$n(\vec{q}s) \;=\; \frac{1}{e^{\hbar\omega(\vec{q}s)/kT} - 1} \quad \rightarrow \quad \frac{k_B T}{\hbar\omega(\vec{q}s)} \tag{8.66}$$

so one would expect that at high temperature the mean free path and the thermal conductivity decrease proportionally to $1/T$. This behaviour is actually observed at high enough temperatures.

When temperature decreases, however, one observes an increase of thermal conductivity by far larger than the one expected on the basis of the $1/T$ dependence.
This behaviour is explained by the peculiarity of the momentum conservation law for phonons (8.64). One has to distinguish two types of phonon collisions:

1. **Normal processes**, where $\vec{G} = 0$ in (8.64).
   In normal processes, the phonon momentum is conserved. An example is shown in the left panel of Fig. 8.12: the sum $\vec{q}_1 + \vec{q}_2$ is inside the 1st B.Z., so that $\vec{q}_1 + \vec{q}_2 = \vec{q}_3$.
   Normal processes correspond to energy transport without local thermalisation, so that they don't contribute to thermal resistivity. They give rise to a thermal energy flux without restrictions.
2. **"Umklapp" processes**, where $\vec{G} \neq 0$ in (8.64).
   In Umklapp processes (after the German word for *flipping over*), the phonon momentum is not conserved: part of it is exchanged with the crystal. An example is shown in the right panel of Fig. 8.12: the sum $\vec{q}_1 + \vec{q}_2$ is outside the 1st B.Z., so that to obtain a resultant vector $\vec{q}_3$ inside the 1st B.Z. is is necessary to subtract a reciprocal lattice vector $\vec{G}$.
   Only umklapp processes contribute to the local thermalisation of the crystal and to the thermal resistivity.

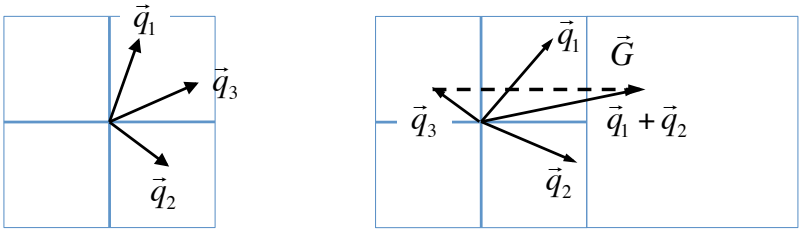

Figure 8.12: *Left:* a normal phonon collision: $\vec{q}_1 + \vec{q}_2 = \vec{q}_3$; the sum $\vec{q}_1 + \vec{q}_2$ is inside the 1st B.Z. *Right:* an umklapp phonon collision: $\vec{q}_1 + \vec{q}_2 = \vec{q}_3 + \vec{G}$; the sum $\vec{q}_1 + \vec{q}_2$ is outside the 1st B.Z. The reciprocal lattice vector $\vec{G}$ is represented by the dashed arrow.

Note: The mechanism of heat transport induced by normal phonon collisions is similar to the mechanism of heat transport by convection in a gas within a vessel open at the low and high-temperature ends.

The balance between normal and umklapp processes depends on temperature.
For example, let us consider a three phonon process (Fig. 8.12). For the process to be of the umklapp type ($\vec{G} \neq 0$), it is necessary that the sum $\vec{q}_1 + \vec{q}_2$ is outside the first BZ, so that the subtraction of $\vec{G}$ is necessary to obtain a $\vec{q}_3$ wavevector inside the first BZ.
Umklapp processes are most frequent for high energy phonons, which are excited at high temperature.

Temperature dependence of the crystal thermal conductivity

To explore the temperature dependence of the lattice thermal conductivity, it is convenient to start from the high temperature side (Fig. 8.10).

- At high temperatures, where umklapp collisions prevail, the dependence on temperature is of the $1/T$ type.
- When temperature decreases and normal collisions progressively prevail over umklapp collisions, the thermal conductivity increases faster that $1/T$, with exponential behaviour.
- When the mean free path for phonon-phonon collisions becomes large enough to be comparable to the mean free path for collisions of phonons with lattice defects or even with

the specimen size, the thermal conductivity ceases to increase. The mean free path contribution becomes temperature independent, and the dependence of thermal conductivity on temperature is determined by the T^3 behaviour of the specific heat.

8.4 Complements and demonstrations

8.4.1 Limits of the single-mode specific heat

Let us consider the expression (8.2) of the single-mode specific heat:

$$c_v = \left(\frac{\partial\langle\epsilon\rangle}{\partial T}\right)_v = k_B \left(\frac{\hbar\omega}{k_B T}\right)^2 \frac{e^{\hbar\omega/k_B T}}{\left(e^{\hbar\omega/kT}-1\right)^2}\,. \tag{8.67}$$

High-temperature limit

The high-temperature limit of (8.67) can be obtained by power expanding the exponentials

$$\begin{aligned} c_v &\simeq k_B \left(\frac{\hbar\omega}{k_B T}\right)^2 \frac{1+\hbar\omega/k_B T+\ldots}{(1+\hbar\omega/k_B T+\ldots-1)^2} \\ &\simeq k_B \left(\frac{\hbar\omega}{k_B T}\right)^2 \left(\frac{k_B T}{\hbar\omega}\right)^2 \left[1+\frac{\hbar\omega}{k_B T}+\ldots\right] \end{aligned} \tag{8.68}$$

so that

$$\lim_{T\to\infty} c_v = k_B\,. \tag{8.69}$$

Low-temperature limit

The low-temperature expression (8.3) can be obtained by multiplying and dividing (8.67) by $\exp(-2\hbar\omega/k_B T)$:

$$\begin{aligned} c_v &= k_B \left(\frac{\hbar\omega}{k_B T}\right)^2 \frac{e^{-2\hbar\omega/k_B T}\, e^{\hbar\omega/k_B T}}{e^{-2\hbar\omega/k_B T}\left(e^{2\hbar\omega/kT}-2e^{\hbar\omega/kT}+1\right)} \\ &= k_B \left(\frac{\hbar\omega}{k_B T}\right)^2 \frac{e^{-\hbar\omega/k_B T}}{1-2e^{-\hbar\omega/kT}+e^{-2\hbar\omega/kT}} \\ &\simeq k_B \left(\frac{\hbar\omega}{k_B T}\right)^2 e^{-\hbar\omega/k_B T} \end{aligned} \tag{8.70}$$

By defining $y = \hbar\omega/k_B T$, the limit for $T \to 0$ is the limit for $y \to \infty$, which can be evaluated by applying the rule of de L'Hôpital twice:

$$\lim_{y\to\infty} \frac{y^2}{e^y} = \lim_{y\to\infty} \frac{2y}{e^y} = \lim_{y\to\infty} \frac{2}{e^y} = 0\,. \tag{8.71}$$

8.4.2 Thermodynamic relations

Useful mathematical identities

Let us consider three variables x, y, z connected by an implicit relation

$$F(x,y,z) = 0\,. \tag{8.72}$$

In the thermodynamical applications here considered, the three variables are pressure, volume and temperature.

The differentials dx and dy can be expressed as:

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz\,, \qquad dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz\,. \tag{8.73}$$

By substituting the expression of dy in the differential of dx, one obtains

$$\left[\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z - 1\right] dx + \left[\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y\right] dz \;=\; 0\,. \tag{8.74}$$

If x and z are chosen as independent variables, (8.74) is valid for whichever values of dx and dz, so that the two expressions in square parentheses must be identically zero.
By equating to zero the first parenthesis of (8.74) one obtains

$$\boxed{\left(\frac{\partial x}{\partial y}\right)_z \;=\; \frac{1}{\left(\dfrac{\partial y}{\partial x}\right)_z}\,.} \tag{8.75}$$

By equating to zero the second parenthesis of (8.74) one obtains

$$\left(\frac{\partial x}{\partial z}\right)_y \;=\; -\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x\,, \tag{8.76}$$

say

$$\boxed{\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y \;=\; -1\,.} \tag{8.77}$$

Coefficient of thermal expansion

The relation (8.77) is now used to demonstrate the relation (8.27) of the coefficient of thermal expansion with other thermodynamic response functions.
As a first step, let us exploit the identity (8.77) to substitute the derivative with respect to temperature and introduce the isothermal compressibility $\chi_T = -(\partial V/\partial p)_T/V$:

$$\beta \;=\; \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p \;=\; -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T \left(\frac{\partial p}{\partial T}\right)_V \;=\; \chi_T \left(\frac{\partial p}{\partial T}\right)_V \tag{8.78}$$

The derivative of pressure with respect to temperature can be further expanded by introducing the internal energy U

$$\left(\frac{\partial p}{\partial T}\right)_V \;=\; \left(\frac{\partial p}{\partial U}\right)_V \left(\frac{\partial U}{\partial T}\right)_V \tag{8.79}$$

so that the heat capacity at constant volume enters in the expression of the coefficient of thermal expansion

$$\beta \;=\; \chi_T \,\frac{C_v}{V}\left[\frac{\partial p}{\partial (U/V)}\right]_V\,. \tag{8.80}$$

The third factor

$$\gamma \;=\; \left[\frac{\partial p}{\partial (U/V)}\right]_V \tag{8.81}$$

is the Grüneisen function.

$$\boxed{\beta \;=\; \gamma(T)\,\chi_T\,\frac{C_v}{V} \;=\; \gamma\,\chi_T\,\frac{c_v}{v}}\,, \tag{8.82}$$

where c_v and v are the molar specific heat and the molar volume, respectively.

8.5 Bibliography of Chapter 8

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions). Chapters 23 (quantum theory of the harmonic oscillator) and 25.
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapter 5 (phonons: thermal properties).
- Mark W. Zemansky: *Heat and Thermodynamics*, MacGraw Hill 1996.
- H.B. Callen: *Thermodynamics and an Introduction to Thermostatistics*, Wiley 1985.
- T. H. K. Barron and G. K. White: *Heat capacity and thermal expansion at low temperatures*, Kluwer Academic - Plenum 1999.
- T. H. K. Barron: *Generalized theory of thermal expansion of solids*, Chapter 1 of *Thermal expansion of solids*, ed. by R. E. Taylor, ASM International 1998.

Chapter 9

Ground-state electronic properties of crystals

In this Chapter 9 we consider the ground-state properties of valence electrons in crystals. The static thermal properties (specific heat) and the transport properties are considered in Chapter 10. We maintain here the distinction, introduced in Section 6.1, between core electrons and valence electrons. By valence electrons we mean the outermost electrons of atoms, which participate to the chemical bond and can give rise to thermal and electric transport properties. In metals, the electrons that are free to move are called conduction electrons.
We begin with the quite crude electron gas model in Section 9.1. In Section 9.2 the general properties of valence electrons in a periodic potential are studied (Bloch states, energy band structure). The nearly-free electron model of Section 9.3 gives a relatively easy explanation of the origin of forbidden energy gaps in the band structure of crystals. In Section 9.4 the origin of semiconducting properties is explained in terms of thermal excitations.

9.1 Non-interacting electron gas: ground state

A very simple approximation consists in considering the conduction electrons of a metal as an ideal gas of free and independent negatively charged particles, embedded in a homogeneous distribution of positive charge (jellium model):

a) free electrons = not interacting with the ion lattice
b) independent electrons = without electron-electron interaction

In spite of its crudeness, the model of non-interacting electron gas can explain the qualitative (and sometimes also quantitative) behaviour of a limited number of properties of metals, such as the electron contribution to the specific heat.
The model is useful to introduce some important concepts, such as the Fermi surface and the Fermi energy; besides, it represents a zero-order approximation on which to build, as higher-order approximations, the effects of the interaction with the ions lattice (Section 9.2) and the effect of electron-electron interaction.
In this Section § 9.1 we consider only the ground state properties of the free electron gas, say the properties for $T = 0$.

9.1.1 Schrödinger equation and its solutions

The Schrödinger equation for a non-interacting electron is

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r}) \;=\; \epsilon\,\Psi(\vec{r})\,. \tag{9.1}$$

The solutions of (9.1) are plane waves

$$\Psi_{\vec{k}}(\vec{r}) \;=\; \frac{1}{\sqrt{V}}\, e^{i\vec{k}\cdot\vec{r}}\,, \tag{9.2}$$

where V is the crystal volume, $\epsilon = \hbar^2 k^2/2m$ and the wavevector $\vec{k}$ is a quantum number.

Periodic boundary conditions

For finite crystals, we impose the periodic boundary conditions

$$\Psi(\vec{r}) \;=\; \Psi(\vec{r} + N_i\vec{a}_i)\,, \quad (i = 1, 2, 3)\,, \tag{9.3}$$

which are fulfilled (Section 3.4) only for wavevectors

$$\vec{k} \;=\; \sum_{i=1}^{3} k_i\vec{b}_i = \sum_{i=1}^{3} \frac{m_i}{N_i}\vec{b}_i\,. \tag{9.4}$$

The tips of vectors $\vec{k}$ represent a grid of equally spaced points filling up the entire reciprocal space.

Density of $\vec{k}$ points

The total number of points per primitive cell of the reciprocal lattice is equal to the number of primitive cells in the real space, $N_1N_2N_3 = N$.
The density of $\vec{k}$ points in reciprocal space is thus

$$\rho(\vec{k}) \;=\; \frac{N}{V_b} \;=\; \frac{N}{(2\pi)^3/V_a} \;=\; \frac{NV_a}{(2\pi)^3} \;=\; \frac{V}{(2\pi)^3}\,, \tag{9.5}$$

where V_a and V_b are the primitive cell volumes of real and reciprocal space, respectively, and V is the total crystal volume. The density $\rho(\vec{k})$ increases when the crystal size (and the number of free electrons) increases.

Linear momentum

One can easily verify that the plane waves (9.2) are eigenfunctions of the linear momentum operator $\vec{p} \;=\; \hbar\vec{k}$ too:

$$-\,i\hbar\,\vec{\nabla} e^{i\vec{k}\cdot\vec{r}} \;=\; \hbar\vec{k}\, e^{i\vec{k}\cdot\vec{r}}\,. \tag{9.6}$$

9.1.2 The Fermi sphere

Due to the Pauli exclusion principle, to each wavevector $\vec{k}$, say to each momentum $\vec{p} = \hbar\vec{k}$, one can associate only two electrons of opposite spin directions.
The ground state of the system of N_e non-interacting electrons is obtained by filling up the $\vec{k}$ states, each one with two electrons, starting from the centre of the reciprocal space and progressively increasing the energy $\epsilon = \hbar^2 k^2/2m$. When all electrons are accomodated, the $\vec{k}$ states filled up are enclosed within the *Fermi sphere*, whose radius is labelled k_F.

The spherical surface enclosing the Fermi sphere is called *Fermi surface*.

Parameters of the Fermi sphere

The volume and the radius of the Fermi sphere are obviously connected by

$$V_F \;=\; \frac{4}{3}\pi\, k_F^3\,. \tag{9.7}$$

The value of k_F can be obtained by considering that the total number N_e of conduction electrons can be expressed as the product of (Fermi sphere volume) $\times$ (density 9.5) $\times$ (2 spin orientations):

$$N_e = V_F \frac{V}{(2\pi)^3} 2 = \frac{k_F^3}{3\pi^2} V , \tag{9.8}$$

so that the Fermi wavevector is given by

$$k_F^3 = 3\pi^2 \frac{N_e}{V} = 3\pi^2 n , \tag{9.9}$$

where $n = N_e/V$ is the number density of conduction electrons.
The Fermi energy is

$$\epsilon_F = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2}{2m} \left(3\pi^2 \frac{N_e}{V}\right)^{2/3} = \frac{\hbar^2}{2m} \left(3\pi^2 n\right)^{2/3} . \tag{9.10}$$

It can be useful to focus once more on the reciprocity property of the Fourier transform:

a) the volume V of the crystal in the real space is proportional to the density of $\vec{k}$ values in the reciprocal space;
b) the volume V_F of the Fermi sphere in the reciprocal space is proportional to the density of conduction electrons in the real space.

From the number density n one can obtain a set of important quantities, which are listed below, together with their typical values:

a) The Fermi wavevector $k_F \simeq 0.7 \div 1.8\,\text{Å}^{-1}$
and the related Fermi momentum $p_F = \hbar k_F$.
b) The Fermi energy $\epsilon_F = \hbar^2 k_F^2/2m \simeq 1.8 \div 13\,\text{eV}$.
The Fermi energy is typically some orders of magnitude larger than the thermal energy $k_B T$ at room temperature.
c) The Fermi temperature $T_F = \epsilon_F/k_B \simeq (2 \div 15) \times 10^4\,\text{K}$.
The Fermi temperature is some orders of magnitude larger than the room temperature.
d) The Fermi velocity $v_F = \hbar k_F/m \simeq (0.8 \div 2) \times 10^6\,\text{m/s}$.
The Fermi velocity is only two orders of magnitude smaller than the velocity of light.

The values for some selected metals are listed in Table 9.1.

Table 9.1: Relevant quantities for some metals: structure, lattice parameter a, conduction electrons per atom Z, volume density of conduction electrons n, Fermi wavevector k_F, energy ϵ_F and temperature T_F.

	struc.	a (Å)	Z	n (10^{22}cm^{-3})	k_F (Å^{-1})	ϵ_F (eV)	T_F (10^4 K)
Na	bcc	4.23	1	2.65	0.92	3.24	3.77
Cu	fcc	3.61	1	8.47	1.36	7.00	8.16
Fe	bcc	2.87	2	17.0	1.71	11.1	13.0

Electron density of states

The number of states included between the energy values ϵ and $\epsilon + d\epsilon$ is $g(\epsilon)\, d\epsilon$, where $g(\epsilon)$ is the electron density of states (DOS). The function $g(\epsilon)$ for non-interacting electrons can be obtained by calculating the number N of states within a sphere of radius $k = \sqrt{2m\epsilon}/\hbar$ and calculating the derivative with respect to ϵ. The density of states per unit volume is

$$g(\epsilon) = \frac{1}{V}\frac{dN}{d\epsilon} = \frac{1}{V}\frac{d}{d\epsilon}\left[\frac{4}{3}\pi \frac{(2m\epsilon)^{3/2}}{\hbar^3} 2 \frac{V}{(2\pi)^3}\right] = \frac{\sqrt{2m^3}}{\pi^2\hbar^3}\sqrt{\epsilon} . \tag{9.11}$$

The density of states is proportional to the square root of the energy.

Problem: Compare the analytical expression and the physical meaning of the DOS of the free electron gas $g(\epsilon)$ and the phonon VDOS of the Debye model $g(\omega)$.

Excitations and Fermi energy

Because of the Pauli principle, only electrons near the Fermi surface can be excited.

9.2 Independent electrons in a periodic potential

The electron gas model is very crude. Actually, valence electrons interact with the positive ions (nuclei plus core electrons) and between themselves. Taking into account these interactions is necessary to account for most of the electronic properties of crystals; in particular, to account for the different behaviour of insulators and metals.

The problem we are facing is a typical many-body problem, which can be solved at different levels of approximation.

9.2.1 Independent electrons approximation

In principle, the electronic quantum state should be described by the total electron wave-function $\Psi_e(r;R)$ introduced in Section 6.1, where r and R are short-hand notations for the full sets of valence electrons and ions coordinates, respectively and the ion coordinates are considered as parameters within the adiabatic approximation.
In the independent electrons approximation, the total electron wavefunction is considered as a Slater determinant of one-electron wavefunctions $\psi(\vec{r})$

$$\Psi_e(\vec{r}_1,\vec{r}_2,\ldots,\vec{r}_N) \;=\; \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\vec{r}_1) & \psi_1(\vec{r}_2) & \cdots & \psi_1(\vec{r}_N) \\ \psi_2(\vec{r}_1) & \psi_2(\vec{r}_2) & \cdots & \psi_2(\vec{r}_N) \\ \cdots & & & \\ \psi_N(\vec{r}_1) & \psi_N(\vec{r}_2) & \cdots & \psi_N(\vec{r}_N) \end{vmatrix} \tag{9.12}$$

with a procedure similar to that used for many-electron atoms.
The one-particle Schrödinger equation for a valence electron is

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right]\psi(\vec{r}) \;=\; \epsilon\,\psi(\vec{r})\,. \tag{9.13}$$

where $U(\vec{r})$ is an effective potential energy for one electron:

1. The potential energy is invariant with respect to the translational symmetry operations of the crystal
$$U(\vec{r}) \;=\; U(\vec{r}+\vec{T}) \tag{9.14}$$
where $\vec{T}$ is any Bravais lattice vector.

2. The term "effective" means here that the potential energy U contains the contribution of the interaction of one electron with all the other valence electrons in terms of a mean periodic field. The potential energy $U(\vec{r})$ and the electron eigenfunctions are calculated by self-consistent procedures.

The evaluation of the potential energy $U(\vec{r})$ for a given system and the solution of the Schrödinger equation (9.13) are difficult tasks, for which different approaches have been developed.

Anyway, some general properties of the wavefunctions $\Psi(\vec{r})$ and of the energy levels ϵ can be determined solely as a consequence of the translational symmetry of the crystal. These general properties will be now analysed.

The case of non-interacting electrons

If $U(\vec{r}) = 0$ in (9.13), the free-electron Schrödinger equation (9.1) is recovered.

9.2.2 Symmetry considerations

One-dimensional chain

In Section 5.4 we studied the one-dimensional linear chain of N points with period a, with periodic boundary conditions. We found that there are N one-dimensional irreducible representations (I.R.) of the group of translations, whose basis functions have the general form (5.64)

$$\psi_k(x) \;=\; e^{ikx} \sum\nolimits_G A_G e^{iGx}\,, \tag{9.15}$$

where G are vectors of the one-dimensional reciprocal lattice. The N irreducible representations are labelled by the N values k

$$k = \frac{2\pi}{a}\,\frac{m}{N} = b\,\frac{m}{N} \qquad (N/2 < m \leq N/2)\,, \tag{9.16}$$

confined within the first Brillouin Zone, where $b = 2\pi/a$ is the period of the one-dimensional reciprocal space and m is an integer number.
A translation $T = sa$ is represented, in the I.R. corresponding to a given value of k, by the complex number

$$\left(e^{iT}\right)^k \;=\; e^{ikT}\,. \tag{9.17}$$

The sum in (9.15) represents any complex periodic function $u(x)$, of period a. There are thus infinite possible basis functions for each one-dimensional irreducible representation k, corresponding to the infinite possible periodic functions $u(x)$. It is customary to characterise the periodic functions as u_{kn}, where k labels the I.R., n labels the basis functions for a given k value.

Three-dimensional crystal

The results of the linear chain can be easily extended to a *three-dimensional* crystal with periodic boundary conditions.
If only translational symmetry is considered, the $N = N_1 N_2 N_3$ generalised basis functions are

$$\Psi_{\vec{k}}(\vec{r}) \;=\; e^{i\vec{k}\cdot\vec{r}} \sum\nolimits_{\vec{G}} B_{\vec{G}} e^{i\vec{G}\cdot\vec{r}}\,, \tag{9.18}$$

where $\vec{G}$ are vectors of the reciprocal lattice; the vectors $\vec{k}$, of components

$$k_i = b_i\,\frac{m_i}{N_i} \qquad (N_i/2 < m_i \leq N_i/2)\,, \tag{9.19}$$

form a grid of equally spaced points within the three-dimensional first Brillouin Zone and b_i is the magnitude of the $\vec{b}_i$ primitive vector of reciprocal space.
The vectors $\vec{k}$ are the quantum labels of the N irreducible representations of the translation group, often called crystal momentum quantum numbers.
When the crystal size increases, the number $N = N_1 N_2 N_3$ of primitive cells in the real space increases, and the density of $\vec{k}$ points in the reciprocal space increases accordingly.
A translation $\vec{T}$ is represented, in the I.R. corresponding to the $\vec{k}$ crystal momentum, by the complex number

$$e^{i\vec{k}\cdot\vec{T}}\,. \tag{9.20}$$

The sum in (9.18) represents any complex periodic function $u(\vec{r})$ with the periodicity of the lattice. There are infinite possible basis functions for each one-dimensional irreducible representation $\vec{k}$,

corresponding to the infinite possible periodic functions $u(\vec{r})$. It is customary to characterise the periodic functions as $u_{\vec{k}n}$, where $\vec{k}$ labels the I.R., n labels the basis functions for a given k value.

Note: Similar considerations on irreducible representations of the translation group of a finite crystal with periodic boundary conditions led, in Section 7.3, to express the spatial periodicity of the eigenvectors of the dynamical matrix in terms of a set of wavevectors $\vec{q}$ that form a grid in the first Brillouin Zone and to the definition of a Fourier transformed dynamical matrix for each of the $\vec{q}$ values.

9.2.3 Bloch's theorem

As it was shown in Section 5.6, as a consequence of the invariance of the Hamiltonian operator with respect to symmetry transformations (in this case lattice translations), the eigenvalues ϵ of the Schrödinger equation (9.13), say the energy levels, are also invariant with respect to the symmetry transformations.

Let us, for the moment, focus our attention on the eigenfunctions of the Schrödinger equation.

Each one-electron eigenfunction $\psi(\vec{r})$ of (9.13), defined to within a phase factor, is a basis of a one-dimensional irreducible representation of the translation group. According to (9.18), the eigenfunctions of the Schrödinger equation (9.13) can be expressed (Bloch theorem) as

$$\boxed{\psi_{\vec{k}n}(\vec{r}) \;=\; e^{i\vec{k}\cdot\vec{r}}\, u_{\vec{k}n}(\vec{r})\,,} \tag{9.21}$$

where $u_{\vec{k}n}(\vec{r})$ is periodic with respect to the Bravais lattice,

$$u_{\vec{k}n}(\vec{r}) \;=\; u_{\vec{k}n}(\vec{r}+\vec{T})\,. \tag{9.22}$$

In the I.R. represented by the vector $\vec{k}$, a translation $\vec{T}$ is represented by (9.20), so that the eigenfunctions $\Psi(\vec{r})$ transform according to

$$\boxed{\psi_{\vec{k}n}(\vec{r}+\vec{T}) \;=\; e^{i\vec{k}\cdot\vec{T}}\,\psi_{\vec{k}n}(\vec{r})} \tag{9.23}$$

The probability density has the same periodicity of the lattice: $|\psi_{\vec{k}}(\vec{r}+\vec{T})|^2 = |\psi_{\vec{k}}(\vec{r})|^2$.

For each I.R. characterised by the wavevector k, infinite basis functions are possible, labelled by the index n.

Eq. (9.23) can be directly derived from (9.21). Proof:

$$\begin{aligned}
\psi_{\vec{k}n}(\vec{r}+\vec{T}) \;&=\; e^{i\vec{k}\cdot\vec{r}}\, e^{i\vec{k}\cdot\vec{T}}\, u_{\vec{k}n}(\vec{r}+\vec{T}) \\
&=\; e^{i\vec{k}\cdot\vec{T}}\, e^{i\vec{k}\cdot\vec{r}}\, u_{\vec{k}n}(\vec{r}) \\
&=\; e^{i\vec{k}\cdot\vec{T}}\,\psi_{\vec{k}}(\vec{r})\,,
\end{aligned} \tag{9.24}$$

The case of non-interacting electrons

For a non-interacting electron, say when $U(\vec{r}) = 0$ in (9.13), the eigenfunctions of the Schrödinger equation (9.1) are the plane waves (9.2), defined for all $\vec{k}$ values of the reciprocal space that satisfy the periodic boundary conditions.

Even for the free electron case, however, one can consider the effects of lattice periodicity and express the wavefunctions in Bloch form, restricting the $\vec{k}$ values to within the first Brillouin zone. Here we label $\vec{k}_{\rm in}$ and $\vec{k}_{\rm out}$ the wavevectors that lie inside and outside the 1st B.Z., respectively. Let us consider a wavevector $\vec{k}_{\rm out}$ outside the 1st B.Z. and let $\vec{G}$ be a reciprocal lattice vector such that $\vec{k}_{\rm out} - \vec{G} = \vec{k}_{\rm in}$. The wavefunction can be modified as follows

$$\psi_{\vec{k}_{\rm out}}(\vec{r}) \;=\; \frac{1}{\sqrt{V}}\, e^{i\vec{k}_{\rm out}\cdot\vec{r}} \;=\; \frac{1}{\sqrt{V}}\, e^{i(\vec{k}_{\rm out}-\vec{G}+\vec{G})\cdot\vec{r}} \;=\; \frac{1}{\sqrt{V}}\, e^{i\vec{k}_{\rm in}\cdot\vec{r}}\, e^{i\vec{G}\cdot\vec{r}}\,. \tag{9.25}$$

The right-hand expression has the form of a Bloch wavefunction, where the wavevector defining the plane wave is inside the 1st B.Z. and $u_{k_{\rm in}n}(\vec{r}) = e^{i\vec{G}\cdot\vec{r}}$ has the periodicity of the lattice.

Reduced-zone scheme and extended-zone scheme

When only the $\vec{k}_{\text{in}}$ values within the 1st B.Z. are considered, one speaks of reduced zone scheme. When all the possible values of $\vec{k}$ are considered, inside and outside the 1st B.Z., one speaks of extended zone scheme. See Figure 9.1, left and centre, for the one-dimensional case of the non-interacting electrons.

9.2.4 Consequences of the Bloch's theorem

Let us consider here some of the most relevant consequences of the Bloch's theorem.

Meaning of the wavevector $\vec{k}$

For non-interacting electrons, the plane waves $\exp(i\vec{k}\cdot\vec{r})$, solutions of the Schrödinger equation (9.1), are also eigenfunctions of the linear momentum operator, with

$$\vec{p} = \hbar\vec{k}\,. \tag{9.26}$$

For electrons in a periodic potential, the solutions of the Schrödinger equation (9.13), say the Bloch functions

$$e^{i\vec{k}\cdot\vec{r}}\,u_{\vec{k}n}(\vec{r})\,, \tag{9.27}$$

are not eigenfunctions of the linear momentum operator.
Actually (omitting the indices $\vec{k}n$ to simplify notation)

$$\begin{aligned} -i\hbar\vec{\nabla}\left[e^{i\vec{k}\cdot\vec{r}}\,u(\vec{r})\right] &= \left(-i\hbar\vec{\nabla}e^{i\vec{k}\cdot\vec{r}}\right)u(\vec{r}) - i\hbar\,e^{i\vec{k}\cdot\vec{r}}\,\vec{\nabla}u(\vec{r}) \\ &= \hbar\vec{k}\,e^{i\vec{k}\cdot\vec{r}}\,u(\vec{r}) \;-\; i\hbar\,e^{i\vec{k}\cdot\vec{r}}\,\vec{\nabla}u(\vec{r})\,. \end{aligned} \tag{9.28}$$

The quantity $\hbar\vec{k}$ is named the quasi-momentum (or crystal momentum) of the electron. It is constant for a stationary electron state within a periodic potential energy. It can be modified only by the action of external forces (e.g. and external electric field).

The quantity $\hbar\vec{k}$ cannot be identified with the linear momentum $\vec{p}$, which is not constant under the action of the periodic potential, say of internal forces.

Note: For a free electron, the relation (9.26) holds for plane waves, not for Bloch functions (9.25), where $\vec{k}$ is restricted to the 1st B.Z.

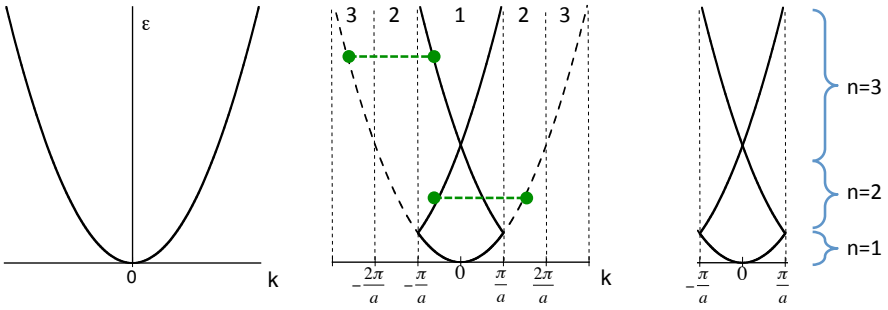

Figure 9.1: *Left:* Dispersion relation $\epsilon = (\hbar k)^2/2m$ for a non-interacting electron in one dimension. *Centre:* The same dispersion relation in the extended-zone scheme (dashed line) and in the reduced-zone scheme (continuous line). The vertical dashed lines separate the first three Brillouin zones; the horizontal lines connect points corresponding to values of k differing by a reciprocal lattice unit vector $G = 2\pi/a$. *Rigth:* Reduced-zone scheme; the first three bands labelled by the index n.

First Brillouin Zone

According to (9.19), only the wavevectors $\vec{k}$ included within the First Brillouin Zone (or within any single primitive cell of the reciprocal space) correspond to independent irreducible representations. This means that any two wavevectors $\vec{k}$ and $\vec{k}'$ differing by a reciprocal lattice vector $\vec{G}$

$$\vec{k}' = \vec{k} + \vec{G}\,, \tag{9.29}$$

correspond to the same irreducible representation.
Let $\vec{k}$ be in the 1st B.Z. and let

$$\psi_{\vec{k}n} \;=\; e^{i\vec{k}\cdot\vec{r}}\, u_{\vec{k}n}(\vec{r}) \tag{9.30}$$

be a Bloch solution of the Schrödinger equation for the wavevector $\vec{k}$.
By substituting $\vec{k} = \vec{k}' - \vec{G}$ one obtains

$$\psi_{\vec{k}n}(\vec{r}) \;=\; e^{i(\vec{k}'-\vec{G})\cdot\vec{r}}\, u_{\vec{k}n}(\vec{r}) \;=\; e^{i\vec{k}'\cdot\vec{r}}\, e^{-i\vec{G}\cdot\vec{r}}\, u_{\vec{k}n}(\vec{r}) \;=\; e^{i\vec{k}'\cdot\vec{r}}\, u_{\vec{k}'n}(\vec{r}) \;=\; \psi_{\vec{k}'n}(\vec{r})\,. \tag{9.31}$$

The function $u_{\vec{k}'n}(\vec{r}) = \exp(-i\vec{G}\cdot\vec{r})\, u_{\vec{k}n}(\vec{r})$ is periodic with respect to lattice translations.
Actually, $u_{\vec{k}'n}(\vec{r}) = u_{\vec{k}'n}(\vec{r}+\vec{T})$ because $\exp(-i\vec{G}\cdot\vec{T}) = 1$.
Eq. (9.31) shows that the same Bloch function $\psi_{\vec{k}n}(\vec{r}) = \psi_{\vec{k}'n}(\vec{r})$ corresponds to two different wavevectors $\vec{k}$ and $\vec{k}' = \vec{k} + \vec{G}$.
The different wavelengths of the plane waves $e^{i\vec{k}\cdot\vec{r}}$ and $e^{i\vec{k}'\cdot\vec{r}}$ are compensated by the phase factor $\exp(-i\vec{G}\cdot\vec{r})$ which multiplies the periodic function $u_{\vec{k}n}(\vec{r})$.

Again on the crystal momentum

A better understanding of the relation between the crystal momentum $\hbar\vec{k}$ and the true momentum $\vec{p}$ can be gained by the following considerations.
According to (9.18), the periodic function $u_{\vec{k}n}(\vec{r})$ can be expanded as a Fourier series including only the vectors $\vec{G}$ of the reciprocal lattice:

$$u_{\vec{k}n}(\vec{r}) \;=\; \sum\nolimits_{\vec{G}} B_{\vec{G}} e^{i\vec{G}\cdot\vec{r}}\,, \tag{9.32}$$

so that the Bloch functions (9.21) can be written as

$$\psi_{\vec{k}n}(\vec{r}) \;=\; e^{i\vec{k}\cdot\vec{r}}\, u_{\vec{k}}(\vec{r}) \;=\; \sum\nolimits_{\vec{G}} B_{\vec{G}}\, e^{i(\vec{k}+\vec{G})\cdot\vec{r}}\,, \tag{9.33}$$

say as a sum of plane-waves of different wavenumbers $\vec{k} + \vec{G}$.
As already observed, the true linear momentum $\vec{p}$ is not constant as the electron moves in the periodic potential. Should we precisely measure the linear momentum of the electron in its state $\psi_{\vec{k}}$, we would obtain one of the discrete set of values:

$$\vec{p} = \hbar(\vec{k} + \vec{G})$$

.

9.2.5 Dispersion relations and energy bands

Up to now, we focussed our attention on the Bloch eigenfunctions of the Hamiltonian.
It is now time to consider the energy eigenvalues corresponding to the Bloch eigenfunctions.
To that purpose, it is convenient to start from the simpler case of a free electron in a one-dimensional lattice.
We will again consider the extended and reduced zone schemes introduced above.

Energy bands: the one-dimensional case of non-interacting electrons

If the potential energy $U(\vec{r})$ is switched off in (9.13), we recover the non-interacting electron case considered in Section 9.1, but the concept of Brillouin Zones, which is connected to the presence of the ion lattice, is maintained.
For a non-interacting electron, the relation between energy and momentum (dispersion relation) is parabolic

$$\epsilon = \frac{(\hbar k)^2}{2m} \tag{9.34}$$

and is naturally represented in the extended zone scheme (Fig. 9.1, left), where the wavefunction is $\psi_k(x) = e^{ikx}$ and to each value of k there corresponds a value $\epsilon(k)$ according to (9.34).
In the reduced zone scheme, it was previously found that $\psi_{kn}(x) = e^{ikx}\, u_{kn}(x)$, where k is restricted to the 1st B.Z. and $u_{kn}(x) = e^{iGx}$. The dispersion relation is now limited to the 1st B.Z.: to each value of k there correspond infinite values of energy ϵ_n (Fig. 9.1, centre, continuous lines). The energy values corresponding to a given value of n represent the n-th energy band (Fig. 9.1, right). The bands $n = 1, 2, 3, \ldots$ correspond to different Brillouin zones (BZ) in the extended zone scheme (Fig. 9.1, centre).
The ground state of a system of N_e non-interacting electrons is obtained by filling up the available one-electron states in the bands, in order of increasing energy. The highest energy is the Fermi energy. For the one-dimensional system, the Fermi energy is located in one of the energy bands, depending on the number of electrons to be accomodated.

Energy bands: the two-dimensional case of non-interacting electrons

The forms of the energy bands become more complex when two o three-dimensional systems are considered, not only for physical reasons but even for the difficulty of graphical representations.
To grasp some of the basic concepts, it is useful to consider the two-dimensional case of non-interacting electrons (Fig. 9.2). For simplicity, we consider a square lattice in both real and reciprocal spaces; the 1st B.Z. is a square.
For a non-interacting electron moving on a plane in the real space, the dispersion relation

$$\epsilon = \frac{(\hbar k)^2}{2m} = \frac{\hbar^2}{2m}\,(k_x^2 + k_y^2) \tag{9.35}$$

is represented by a rotation paraboloid in the the reciprocal space (Fig. 9.2, left). The constant-energy curves projected on the (k_x, k_y) plane are circles, whose radius increases with increasing energy (Fig. 9.2, centre). One circle corresponds to the Fermi energy.
The band structure is different for different directions in the reciprocal space, since the constant-energy lines are circles and the 1st B.Z. is a square. It is convenient to plot the dispersion relations along selected high-symmetry directions, in the present case the [10] and [11] directions. In Fig. 9.2, right, one can see the difference of the band structure along the two [10] and [11] directions; in the case here considered, the Fermi energy is within the second band in the [11] direction, within the third band in the [10] direction.
The situation is more complicate in the three-dimensional cases, where the 1st B.Z. can present complex forms. It is anyway customary to represent the dispersion curves along directions of high symmetry, as for the two-dimensional case.

Energy bands: general case

Let us now generalise the concept of energy bands to the three-dimensional case of interacting electrons.
For each wavevector $\vec{k}$ within the First Brillouin Zone, different one-electron energy eigenvalues and corresponding eigenfunctions are possible, which are labelled by the index n.
Let us consider a wavevector $\vec{k}$. According to the Bloch theorem, the wavefunction has the form

$$\psi_{\vec{k}n} = e^{i\vec{k}\cdot\vec{r}}\, u_{\vec{k}n}(\vec{r})\,. \tag{9.36}$$

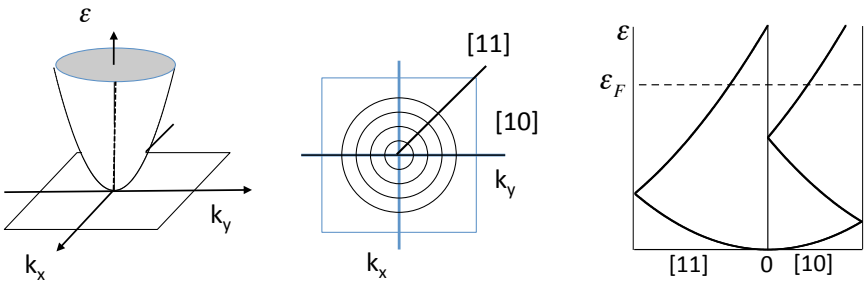

Figure 9.2: Reciprocal space for the free-electron model in two dimensions. *Left:* paraboloid of the dispersion relation. *Centre:* equal-energy contours projected on the k_x, k_y plane. *Right:* Dispersion curves along the high symmetry directions [10] and [11]; the horizontal dashed line represents a possible value of Fermi energy.

Inserting the Bloch function (9.36) into the Schrödinger equation (9.13), the exponential $e^{i\vec{k}\cdot\vec{r}}$ can be eliminated, after evaluation of the effect of the ∇^2 operator on it. One obtains a new differential equation acting only un the $u_{\vec{k}n}(\vec{r})$ functions

$$\left[\frac{\hbar^2}{2m}(-i\vec{\nabla}+\vec{k})^2+U(\vec{r})\right]u_{\vec{k}n}(\vec{r}) \;=\; \epsilon_{\vec{k}n}\,u_{\vec{k}n}(\vec{r})\,. \tag{9.37}$$

The eigenfunctions $u_{\vec{k}n}(\vec{r})$ of (9.37) have the periodicity of the crystal lattice. One can thus restrict our attention within one real-space primitive cell. The finite-volume problem for a bound system leads to a set of discretely spaced eigenvalues $\epsilon_{\vec{k}n} = \epsilon_n(\vec{k})$, where $n = 1, 2, 3, \ldots \infty$.

The values $\epsilon_n(\vec{k})$ for a given n, considered as quasi-continuous functions of the quasi-continuous wavevector values $\vec{k}$, represent the n-th energy band.

The energy bands can be represented both in a reduced zone scheme or in an extended zone scheme. The two alternatives have been illustrated in Fig. 9.1, left panel and right panel, respectively, for the simple one-dimensional case of non-interacting electrons.

The representations are much more complex in two and three dimensions. The effect of a finite periodic potential $U(\vec{r})$ modifies the band structure with respect to the band structure of free electrons; the topic is treated in introductory way in Section 9.3.

Group velocity of electrons

The Bloch functions are de-localised. In order to describe the movement of an electron in the real space, it is necessary to localise it by constructing a wavepacket. If the $\vec{k}$ values are narrowly distributed around a central value $\vec{k}_0$ in the reciprocal space, the corresponding group velocity for an electron in the band n is given by

$$\vec{v}_n(\vec{k}_0) \;=\; \frac{1}{\hbar}\left(\frac{d\epsilon_n(\vec{k})}{d\vec{k}}\right)_{\vec{k}_0}\,. \tag{9.38}$$

The group velocity is easily calculated fro the non-interacting electron case. The case of electrons in a non-zero periodic potential is considered below.

If the distribution of k values in reciprocal space is sufficiently narrow so that the extension of the wavepacket in the real space is much larger than the interatomic distances, the dynamics of the wavepacket is largely independent of the periodic potential $U(\vec{r})$ and is influenced by the presence of external fields (such as externally apllied electric or magnetic fields).

9.3 Nearly-free electron model

The Bloch theorem gives some general properties of the solution of the Schrödinger equation for independent electrons in a periodic potential, in particular the existence of energy bands $\epsilon_n(\vec{k})$. Some further information on the properties of energy bands, and in particular the existence and the origin of forbidden energy gaps, can be gained by considering approximate models of the periodic potential.
Particularly useful, to this purpose, it the quite simple nearly-free electron model. In this model, the periodic crystal potential is considered as a weak perturbation with respect to the Hamiltonian of a non-interacting electron.
In the following, we first consider in some detail the one-dimensional case and then in a more qualitative way the three-dimensional case.

9.3.1 One-dimensional case

Let us consider a linear chain of equal atoms, and let a be the nearest-neighbours distance. Correspondingly, the First Brillouin Zone extends from $k = -\pi/a$ to $k = \pi/a$. The second zone corresponds to the sum of the intervals $-2\pi/a < k < \pi/a$ and $\pi/a < k < 2\pi/a$. And so on.
In view of the weak interaction between electron and ions, the electron wavefunctions are only slightly different from sinusoidal waves $\exp(ikx)$.

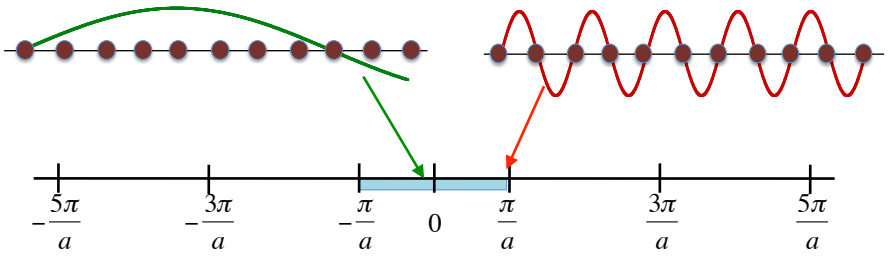

Figure 9.3: *Top:* One-dimensional lattice with spacing a and two electron wavefunctions corresponding to a a central region and the border of the 1st B.Z., respectively. *Bottom:* Corresponding reciprocal space, with 1st B.Z. extending from $-\pi/a$ ti $+\pi/a$.

Let us first consider an electron whose wavevector k lies near the BZ centre, $k \simeq 0$. The wavelength λ is large with respect to the lattice spacing a. The periodic potential energy $U(x)$ is averaged over electron wavelength (Fig. 9.3, top left). As a consequence, the $\epsilon(k)$ dispersion relation is with good approximation parabolic, as for non-interacting electrons.

This behaviour is modified when the electron wavelength λ decreases and the wavenumber $k = 2\pi/\lambda$ approaches the boundaries of the first BZ or of the outer zones, $\pm n\pi/a$ ($n = 1, 2, 3, \ldots$), say when the periodicity of the electron wavefunction approaches the periodicity of the lattice, say of the potential energy $U(x)$.
Let us consider the case of the first BZ (Fig. 9.3, top right): for

$$k \to \pi/a\,, \qquad \lambda \to 2\pi/k = 2a\,, \tag{9.39}$$

the electron wave approaches the condition of being in phase with the lattice. For $\lambda = 2a$, the Bragg condition for elastic scattering of the electron wave by the one-dimensional lattice is fulfilled.

A generalisation of the Bragg condition to all Brillouin Zones is $n\lambda = 2a$. This expression can be recast as the Brillouin condition $k = n\pi/a = G/2$, where G is any reciprocal lattice distance.

If the Bragg (or Brillouin) condition is fulfilled, the electron wavefunction is back and forth reflected by the periodic potential.

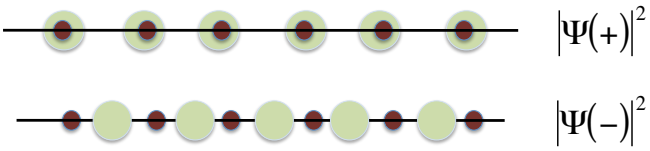

Figure 9.4: For the wavefunction (9.40), the density of electronic charge is maximum in correspondence of the ionic positions (top). For the wavefunction (9.41), the density of electronic charge is maximum between the ionic positions (bottom).

For $k = \pi/a$, the two waves $\exp(\pm i\pi x/a - i\omega t)$ (direct and reflected) can be linearly combined to give rise to the two stationary waves with different phases:

$$\Psi(+) \;=\; \left[e^{i\pi x/a} + e^{-i\pi x/a}\right] e^{-i\omega t} \;=\; 2\cos(\pi x/a)\, e^{-i\omega t}\,, \tag{9.40}$$

$$\Psi(-) \;=\; \left[e^{i\pi x/a} - e^{-i\pi x/a}\right] e^{-i\omega t} \;=\; 2i\sin(\pi x/a)\, e^{-i\omega t}\,. \tag{9.41}$$

The electron probability density is correlated with the ion positions. The two stationary waves (9.40) and (9.41) correspond to two different distributions of electronic charge:

- For $|\Psi(+)|^2 \propto \cos^2(\pi x/a)$, the negative charge density is maximum in correspondence of the ionic positions (Fig. 9.4 , top), giving rise to a decrease of the electrostatic potential energy.

- For $|\Psi(-)|^2 \propto \sin^2(\pi x/a)$, the negative charge density is maximum between the ionic positions (Fig. 9.4, bottom), giving rise to an increase of the electrostatic potential energy.

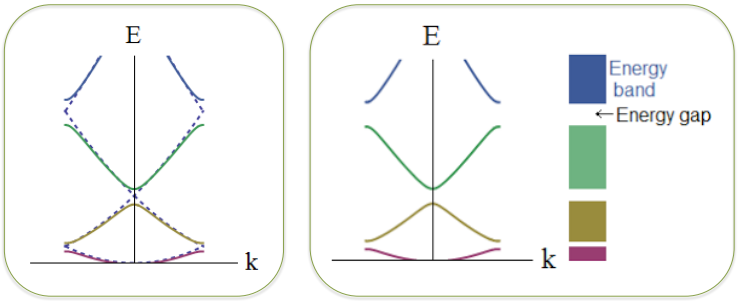

Figure 9.5: Origin of energy gaps in the one-dimensional case. Taken from http://skepticsplay.blogspot.it/2011/06/what-is-electronic-band-structure.html

Similar considerations hold for the borders of the higher-order Brillouin zones.
Actually, since the potential energy is not delta-like in correspondence of the ion positions, the deviation from the free-electron behaviour is not limited to the borders of the Brillouin zones $k = \pm n\pi/a$, but progressively increases when $k \to \pm n\pi/a$; the stronger the periodic potential, the larger is the extent of the deviation from the free-electron behaviour.
For $k \to \pm n\pi/a$, the free-electron parabolic behaviour of the $\epsilon(k)$ dispersion relation is modified, and two different values of ϵ are found for $k = \pm n\pi/a$, lower and higher than $(\hbar k)^2/2m$, respectively. This behaviour is represented in Fig. 9.5, left, in the reduced zone scheme, where the different Brillouin zones correspond to fifferent energy bands.. The differences between the two values of energy ϵ of two contiguous bands at the zone boundary represent forbidden energy gasp (Fig. 9.5, right).
In correspondence of the BZ boundaries, the dispersion curves become horizontal, and the group velocity becomes zero: $d\epsilon/dk = 0$.

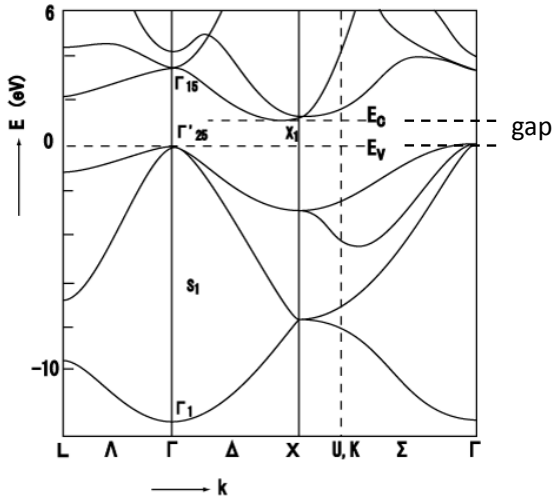

Figure 9.6: Electron energy bands for Silicon. As usual, Γ labels the centre of the 1st B.Z. The energy gap between the filled valence bands and the empty conduction bands is evidenced.

9.3.2 Three-dimensional case

In three dimensions, at the origin of the reciprocal space (Γ point), where the wavelength λ is large with respect to the lattice parameters, the dependence of the energy ϵ on k is parabolic, as for the non-interacting electron model.
In the vicinity of the border of a B.Z., however, the electron wavelength becomes comparable with an inter-planar distance of the crystal and the Brillouin (or Bragg) condition for electron scattering is fulfilled. For the three-dimensional case, the Brillouin condition is

$$\vec{k} \cdot \hat{G} = G/2 \ : \tag{9.42}$$

where $\vec{G}$ is a vector of the reciprocal lattice. The Brillouin condition is fulfilled when the tip of the wave-vector k is on a plane of the reciprocal lattice (Bragg plane) perpendicular to a reciprocal lattice vector $\vec{G}$ through its middle point. The Bragg planes form the borders of the Brillouin Zones.

In correspondence of the BZ boundaries, the electron stationary waves are in phase with the potential energy periodicity. As for the one-dimensional case, two values of energy, higher and lower than the free-electron value, give rise to an energy gap. The extent of the deviation from the free-electron behaviour depends on the peculiarities of the interaction potential $U(\vec{r})$.

The representation of the $\epsilon_n(\vec{k})$ bands in the 1st B.Z. in a two or three-dimensional reciprocal space is far from trivial. As for the case of the energy bands of non-interacting electrons (Fig. 9.2)and of phonon dispersion curves, the energy bands are conventionally represented in two dimensional plots $\epsilon_n(k)$, where k is measured along the directions of highest symmetry, such as [100], [110] or [111].
The example of Silicon is shown in Fig. 9.6. One can notice the different form of the bands along the different high symmetry directions. The most striking peculiarity is the presence of an energy gap which is common to all the 1st B.Z., say an interval of energy values for which there are no available electron states.

In three dimensions, the shape of the equal-energy surfaces in the reciprocal space can be quite complicated, as well as the relation between the equal-energy surfaces and the first B.Z. borders.

9.3.3 A two-dimensional example

We have previously considered the band structure of a two-dimensional system of non-interacting electrons (see Fig. 9.2).
It is now interesting to consider how the presence of a weak periodic potential $U(\vec{r})$ modifies the constant energy curves and the dispersion relations in the two-dimensional reciprocal space.

Let us again consider an ion lattice of square symmetry, so that the first BZ is a square (Fig. 9.7, left). The presence of the periodic potential energy distorts the paraboloid of the dispersion relation $\epsilon(\vec{k})$ of the non-interacting model.

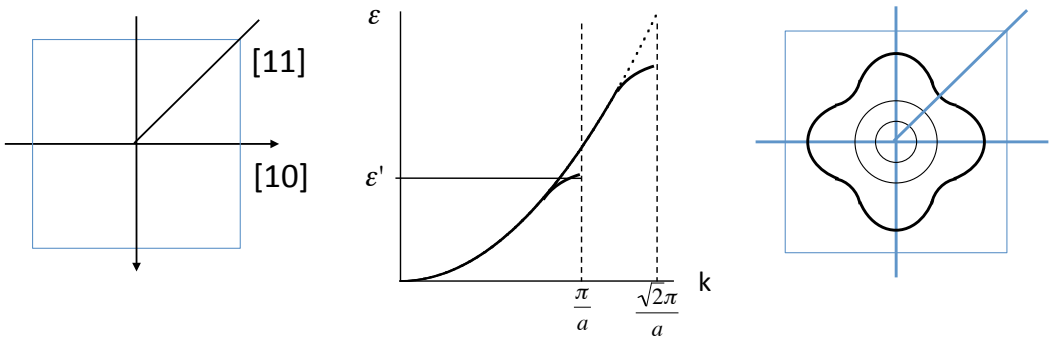

Figure 9.7: Left: 1st Brillouin Zone for a two-dimensional system of square symmetry. Centre: dispersion relations along the two [10] and [11] directions. Left: equal-energy contours within the 1st B.Z.

To study the distortion induced on the paraboloid by the periodic potential energy, let us focus again our attention on the two high-symmetry directions [10] and [11]. Along the [10] directions, the BZ boundary is at $k = \pi/a$; along the [11] direction the BZ bondary is at $k = \sqrt{2}\pi/a$ (Fig. 9.7, centre).
The gap effect is thus obtained for smaller k values (and smaller energies ϵ) along the [10] than along the [11] direction. Let us consider an energy ϵ' for which the gap effect is present along the [10] direction but not along the [11] direction. The energy ϵ' corresponds to a larger k value along the [10] than along the [11] direction (Fig. 9.7, centre).
Let us now consider again the projection of the constant-energy curves on the (k_x, k_y) plane (Fig. 9.7, right). The innermost curves are circles, as for the free-electron case. The outermost curves are deformed: for example, the $\epsilon = \epsilon'$ value corresponds to a higher k value along the [10] than along the [11] direction; as a consequence, the ϵ' constant energy curve exhibits a lobe extending towards the [10] direction (Fig. 9.7, right). The lobes extend in the [10] directions and contract in the [11] directions.
For high enough energies, the constant-energy lines intersect the BZ borders at right angles.

9.4 Filling-up of electron bands

As for the non-interacting electron gas model, also for the electrons in a non-zero periodic potential $U(\vec{r})$ the energy levels are filled up according to the Pauli exclusion principle. The presence of the band structure and of the energy gaps accounts for the important distinction between conductors and insulators at $T = 0$.

9.4.1 One-dimensional case

It is instructive to begin with the simpler one-dimensional model, although it cannot account or all the peculiarities of the more realistic two and three-dimensional systems.

Let us consider a one-dimensional lattice of N atoms, so that the 1st B.Z. contains N points k. For non-interacting electrons, the first band has a parabolic shape (Fig. 9.1) and there are no gaps between the bands; for electrons interacting with a weak potential, energy gaps open at the B.Z. borders (Fig. 9.5). In any case, the first band can accomodate $2N$ electrons (two for each value of k, spin up and spin down).
Let us now consider two ground-state cases:

1. Each atom contributes one valence electron. The total number of electrons to be accomodated in the first band is N: only half of the available states in the first band are occupied. The Fermi energy ϵ_F is defined as the energy separating the higher occupied levels from the lowest unoccupied levels. The Fermi energy is within the first band. Empty levels are available immediately above the Fermi energy, and can be occupied by excited electrons, however weak is the excitation.

2. Each atom contributes two valence electrons. The total number of electrons to be accomodated in the first band is $2N$, say equal to the number of available states. The first band is completely filled.
   The Fermi energy is undefined. No empty levels are available to excited electrons, unless the excitation is sufficient to overcome the energy gap between the first and the second band.

9.4.2 Conductors and insulators

The relatively simple situation depicted for the one-dimensional model becomes much more complicated for two or three-dimensional systems, where the dispersion relations are different along different directions of the reciprocal space and the constant-energy surfaces have complex relationships with the 1t B.Z. borders.

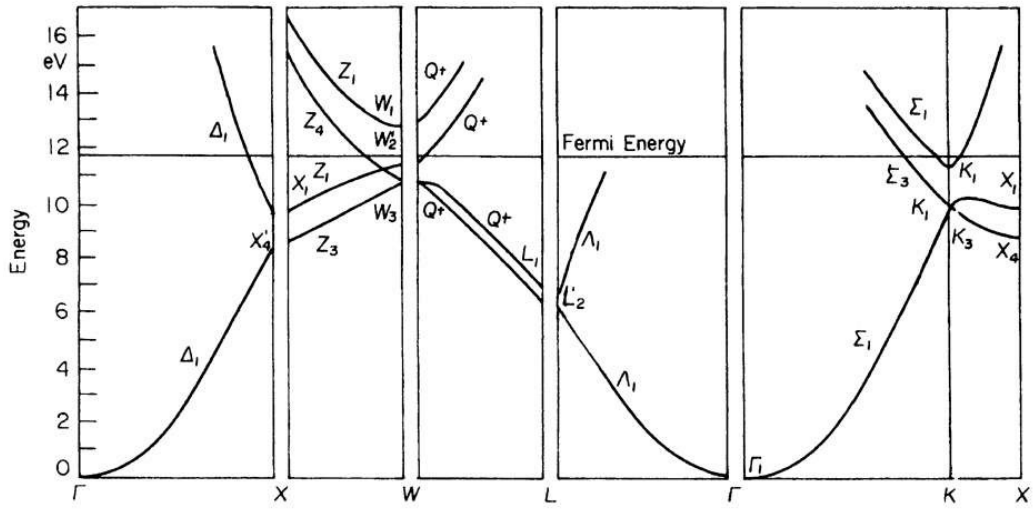

Figure 9.8: Electron energy bands for Aluminium. As usual, Γ labels the centre of the 1st B.Z. The Fermi energy is evidenced.

One can anyway again distinguish two ground-state cases (for $T = 0$), conductors and insulators.

1. In *conductors*, there is at least one energy band not completely filled. The Fermi energy ϵ_F corresponds to the energy level separating the filled levels of highest energy from the empty levels of lowest energy. The band structure of Aluminium is shown in Fig. 9.8: the similarity with the band structure of non-interacting electrons supports the validity of the nearly-free electrons model, at least for metals where no d levels are involved.
   The Fermi energy defines a Fermi surface in the three-dimensional reciprocal space. The Fermi surface is generally composed of several branches corresponding to different partially filled energy bands. The presence of a Fermi surface means that empty levels are available to excited electrons, however weak is the excitation: the system is a good conductor e electricity.

2. In *insulators*, a number of energy bands is completely filled, the other bands being completely empty. The energy difference between the highest filled band and the lowest empty band

corresponds to the energy gap. The band structure of Silicon is shown in Fig. 9.6. The values of the energy gaps for some elements are listed in Table 9.2. If the energy gap is different from zero, at $T = 0\,\mathrm{K}$ the crystal is an insulator. Filled bands are called valence bands, empty bands are called conduction bands. The Fermi energy is undefined.

Table 9.2: Energy gap E_g and occupation index at the bottom of CB at 300 K for selected insulators.

	Diamond	Silicon	Germanium
E_g (eV)	5.33	1.14	0.67
$\epsilon_i - \epsilon_F$ (eV)	2.65	0.5	0.33
n_i/g_i (@ 300 K)	10^{-44}	10^{-9}	10^{-6}

Each energy band is composed by N levels, corresponding to N wavevectors $\vec{k}$, where N is the number of primitive cells in the real space. Each level can host two electrons of opposite spin. Necessary (not sufficient) condition for the presence of an energy gap between valence and conduction bands is that the number of electrons per primitive cell is even.

The most complete description of the electronic structure is given by the dispersion curves. In insulators and semiconductors, the top of the valence band and the bottom of the conduction band (whose energy distance represent the energy gap) can correspond to the same value $\vec{k}$ (direct gap) or to different values of $\vec{k}$ (indirect gap). In Silicon (Fig. 9.6) the gap is indirect.

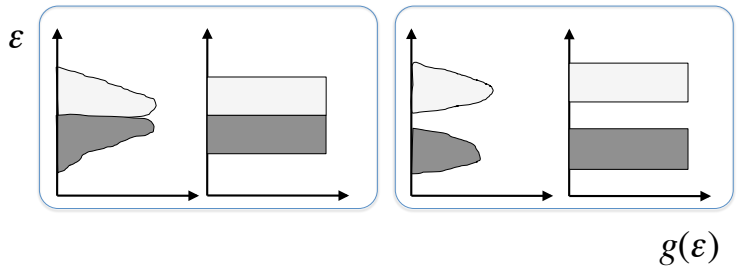

Figure 9.9: Typical density of states for conductors (left panel) and for insulators (right panel). Dark grey represents filled levels, light grey empty levels. In each panel, the real DOS $g(\epsilon)$ (left) and its schematic representation (right) are shown.

Density of states

Frequently, the conduction properties of a crystal are summarised in terms of the density of states (DOS) $g(\epsilon)$.

For a given band n,

$$g_n(\epsilon) \;=\; \int \frac{d\vec{k}}{4\pi^3}\,\delta\left[\epsilon - \epsilon_n(\vec{k})\right]. \tag{9.43}$$

The total density of states is

$$g(\epsilon) \;=\; \sum\nolimits_n g_n(\epsilon)\,. \tag{9.44}$$

In conductors, the conduction band is only partially filled up to the Fermi energy and there is no gap in the density of states (Fig, 9.9, left panel).

In insulators (at $T = 0$), the completely filled valence band and the empty conduction band are separated by a gap, which corresponds to a region where $g(\epsilon) = 0$ (Fig, 9.9, right panel).

Sometimes, when the details of the dispersion curves and of the density of states are not relevant and one focusses only on the gap, the densities of states are schematically represented as rectangles (Fig, 9.9, right part of both panels).

9.4.3 Thermal excitation. Insulators and semiconductors

In insulators, for $T = 0$ there is no electronic conduction.
For $T > 0$ there is the finite probability that the thermal energy $k_B T$ can lead to the promotion of some electrons from the filled valence band (VB) to the empty conduction band (CB), giving rise to a weak conductivity; the system is said to be a *semiconductor*.
In order to evaluate the probability of promotion from the VB to the CB making use of the Fermi-Dirac distribution, it is convenient to define a Fermi energy ϵ_F exactly in the midpoint of the gap between VB and CB, so that the energy ϵ_i at the bottom of the conduction band is related to the width E_g of the energy gap by

$$\epsilon_i - \epsilon_F = E_g/2\,. \tag{9.45}$$

Let n_i be the number of electrons promoted by thermal energy to the bottom of the CB, with energy ϵ_i. The occupation index, corresponding to the probability of finding an electron at the bottom of the CB, is

$$\frac{n_i}{g_i} \;=\; \frac{1}{e^{\beta(\epsilon_i-\epsilon_F)}+1} \;\simeq e^{-\beta(\epsilon_i-\epsilon_F)}\,, \tag{9.46}$$

where g_i is the degeneracy of level i and the last equality is justified by the fact that at room temperature for most systems of interest $\beta(\epsilon_i - \epsilon_F) \gg 1$, say $\epsilon_i - \epsilon_F \gg k_B T$ (Table 9.2).
The values of n_i/g_i are shown in Table 9.2 for selected insulators at room temperature ($k_B T \simeq 0.025$ eV).
The number of atoms is of the order of 10^{22} per cubic centimeter; of the same order of magnitude is the number of electrons in the VB. Each one of these electrons has the probability n_i/g_i of being promoted to the CB.
The probability that one electron is promoted form the VB to the CB by thermal energy is exceedingly low for diamond, which is a good insulator at room temperature; it is instead not negligible for silicon and germanium, which exhibit a weak conductivity at room temperature, and are called semiconductors.
Since n_i/g_i, say the number of carriers, increases with temperature, the electrical conductivity of semiconductors increases with temperature

$$\sigma \;\simeq\; \sigma_0\, e^{-E_g/k_B T}\,. \tag{9.47}$$

Semimetals

In semimetals, such as Bismuth, the band structure tends to be similar to that of semiconductors. However, semimetals are distinguished from semiconductors in that the conduction band minimum lies slightly below the valence band maximum. As a consequence, in the ground state a small number of conduction states are occupied. The electronic properties are those of a metal, but the small number of electrons in the conduction band leads to e very low conductivity.

9.5 Bibliography of Chapter 9

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions). Chapters 1, 2 8.
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapters 6 and 7.
- W. A. Harrison: *Solid state theory*, Dover Publications 1980.
- N. B. Brandt and S. M. Chudinov: *Electronic structure of metals*, MIR Publishers Moskow 1973.

Chapter 10

Electron thermal and transport properties

10.1 Electronic contribution to the specific heat

In metals, electrons contribute to the specific heat in addition to the lattice contribution. An electron contribution to specific heat is present only in systems where a Fermi surface exists. The free electron model is a good starting point to account for the main characteristics of the electronic specific heat, provided the Fermi-Dirac distribution is considered.
To account for the electronic contribution to the specific heat, a complete knowledge of the $\epsilon(\vec{k})$ dispersion relations is not necessary; the specific heat only depends on the vibrational density of states $g(\epsilon)$. (A similar situation was found for the lattice contribution.)

10.1.1 Fermi-Dirac distribution

The effects of temperature on the free electron gas are described by the Fermi-Dirac (FD) distribution. The FD probability distribution is generally expressed in terms of the occupation index n_i/g_i of the energy level i:

$$\frac{n_i}{g_i} = \frac{1}{e^{\alpha+\beta\epsilon_i}+1} \tag{10.1}$$

where

- n_i is the number of electrons with energy ϵ_i,
- g_i is the degeneracy of level i, determined by the number of $\vec{k}$ wavevectors corresponding to the energy value ϵ_i and by the two possible spin states
- the occupation index $n_i/g_i = 0$ or 1, for the Pauli exclusion principle
- α and β are the Lagrange multipliers of the procedure for maximizing the distribution probability, corresponding to the conditions $\sum n_i = N$ and $\sum \epsilon_i n_i = E$, respectively.

The β parameter is connected to the temperature, $\beta = 1/k_B T$.
One can show that the α parameter can be expressed as

$$\alpha = \frac{\partial \ln Z}{\partial N} = -\frac{1}{k_B T}\left(\frac{\partial F}{\partial N}\right) = -\frac{\mu}{k_B T} \tag{10.2}$$

where $Z = \sum_j \exp -\beta E_j$ is the canonical partition function of the entire system of N electrons, F is the Helmholtz function and μ is the chemical potential.
The FD distribution can thus be rewritten as

$$\frac{n_i}{g_i} = \frac{1}{e^{(\epsilon-\mu)/kT}+1}\,, \tag{10.3}$$

Notice that the chemical potential is actually temperature dependent, $\mu(T)$.

Fermi function and density of states

It is often convenient to consider a continuous distribution of energy values, and introduce the continuous function (Fermi function)

$$f(\epsilon) \;=\; \frac{1}{e^{(\epsilon-\mu)/kT}+1}\,. \tag{10.4}$$

The probability of finding an electron within the energy interval between ϵ and $\epsilon + d\epsilon$ is given by the product of the Fermi function (10.4) and the electron density of states

$$f(\epsilon)\, g(\epsilon)\, d\epsilon\,. \tag{10.5}$$

For the free electron gas model, the density of states per unit volume is given by (9.11), so that

$$f(\epsilon)\, g(\epsilon)\, d\epsilon \;=\; \frac{1}{e^{(\epsilon-\mu)/kT}+1}\;\frac{\sqrt{2m^3}}{\pi^2\hbar^3}\;\sqrt{\epsilon}\;d\epsilon\,. \tag{10.6}$$

For a real crystal, the density of states $g(\epsilon)$ depends on the band structure.

Chemical potential and Fermi energy

For $T \to 0\,\mathrm{K}$, the FD distribution (10.4) is

$f(\epsilon) = 1$ for $\epsilon < \mu$,
$f(\epsilon) = 0$ for $\epsilon > \mu$.

Remembering the definition of the Fermi energy ϵ_F, this corresponds to

$$\lim_{T\to 0} \mu \;=\; \mu_0 \;=\; \epsilon_F\,. \tag{10.7}$$

In the ground state, for $T = 0$, the chemical potential can thus be identified with the Fermi energy, $\mu_0 = \epsilon_F$. The $f(\epsilon)$ function is characterised by a sharp step for $\epsilon = \mu_0$.

When the temperature increases, the chemical potential $\mu(T)$ progressively deviates from the value ϵ_F of the Fermi energy, and the sharp step of the $f(\epsilon)$ function is progressively smoothed.
At any temperature, anyway, for $\epsilon = \mu$ the Fermi function is $f(\mu) = 1/2$.
For temperatures much smaller than the Fermi temperature, e.g. room temperatur, the difference between Fermi energy ϵ_F and chemical potential μ can generally be neglected, $\mu \simeq \epsilon_F$.

10.1.2 Specific heat, qualitative approach

At relatively low temperatures, $T \ll T_F$, as a consequence of the Pauli exclusion principle only the electrons near the surface of the Fermi sphere can be thermally excited.
The interested energy interval, where the Fermi function $f(\epsilon)$ significantly deviates from the sharp step behaviour of the ground-state, is of the order of

$$\Delta\epsilon \;\simeq\; |\epsilon - \mu| \;\simeq\; k_B T\,. \tag{10.8}$$

The energy density per unit volume depends on temperature as

$$u(T) \;=\; u_0 + \Delta u(T)\,, \tag{10.9}$$

where u_0 is the ground state energy density. The temperature variation of the energy density is

$$\begin{aligned} \Delta u(T) \;&\simeq\; [\text{number of excited electrons}] \times [\text{average excitation energy}] \\ &\simeq\; g(\mu)\,\Delta\epsilon \;\times\; k_B T \\ &\simeq\; g(\mu)\, k_B T \;\times\; k_B T \\ &=\; g(\mu)\,(k_B T)^2 \end{aligned} \tag{10.10}$$

The electronic contribution to the specific heat per unit volume (at constant volume)

$$c_v^{\mathrm{el}} \;=\; \left(\frac{\partial u}{\partial T}\right)_v \;=\; \gamma T \tag{10.11}$$

is thus linearly dependent on temperature.

A more accurate derivation (see below) confirms the linear dependence on the temperature and on the density of states at the Fermi energy. The specific heat per unit volume at constant volume is

$$\boxed{c_v^{\mathrm{el}} \;\simeq\; \frac{\pi^2}{3}\, k_B^2\, g(\mu_0)\, T \;=\; \frac{\pi^2}{3}\, k_B^2\, g(\epsilon_F)\, T} \tag{10.12}$$

For the free electron gas model, inserting the density of states (9.11) and taking account of the value (9.10) of the Fermi energy, one can show that the specific heat per unit volume (at constant volume) becomes

$$c_v^{\mathrm{el}} \;\simeq\; \frac{\pi^2}{2}\, \frac{n k_B^2}{\epsilon_F}\, T\,. \tag{10.13}$$

Lattice and electronic contributions to specific heat

The total specific heat of metals is the sum of the electronic and vibrational contributions,

$$c_v \;=\; c_v^{\mathrm{el}} \;+\; c_v^{\mathrm{vib}} \;=\; \gamma T \;+\; c_v^{\mathrm{vib}}\,. \tag{10.14}$$

The electronic contribution to specific heat is generally much smaller than the lattice contribution. It can be appreciated at low temperatures, where the linear dependence on temperature prevails over the T^3 dependence, and at high temperatures, where it contributes, together with anharmonicity, to the deviation from the Dulong and Petit constant behaviour.

At low temperatures, where

$$c_v \;=\; \gamma T \;+\; A T^3\,, \tag{10.15}$$

a plot of the measured values c_v/T against T^2 allows the experimental evaluation of γ as the intercept fot $T = 0$ of a straight line.

The agreement between the values of γ measured and evaluated through the free electron approximation is satisfactory for a number of metals, less good or even bad for other metals. The agreement depends on the reliability of the density of states $g(\epsilon)$ calculated through the free electrons model.

Example: In copper, $\gamma \simeq 6.7 \times 10^{-4}\,\mathrm{J\,mol^{-1}\,K^{-1}}$.

At the Debye temperature $\Theta_D = 315$ K, $c_v^{\mathrm{vib}} \simeq 100\, c_v^{\mathrm{el}}$.

The electronic contribution prevails over the lattice contribution only for $T < 15\,\mathrm{K}$.

10.1.3 Detailed calculation of $c_v^{\mathrm{el}}(T)$ and $\mu(T)$

To evaluate the specific heat, we need to know the temperature dependence of the total energy of the conduction electrons. The energy per unit volume is

$$u \;=\; \frac{U}{V} \;=\; \int_0^\infty \epsilon\, g(\epsilon)\, f(\epsilon)\, d\epsilon\,. \tag{10.16}$$

Since the Fermi function $f(\epsilon)$ depends on the chemical potential $\mu(T)$, which in turn depends on temperature, we need a second equation to previously determine $\mu(T)$:

$$n \;=\; \frac{N}{V} \;=\; \int_0^\infty g(\epsilon)\, f(\epsilon)\, d\epsilon\,, \tag{10.17}$$

where N is the total number of conduction electrons.

When T increases, the Fermi function is progressively smoothed around $\epsilon = \mu$, so that the density function $g(\epsilon)$ is differently weighted. To maintain constant the value N of the integral (10.17), the value of μ has to shift with temperature.

Evaluation of the integrals

The integrals of (10.16) and (10.17) are of the general type

$$\int_0^\infty \phi(\epsilon)\, f(\epsilon)\, d\epsilon\,, \tag{10.18}$$

where $\phi(\epsilon) = \epsilon\, g(\epsilon)$ for (10.16) and $\phi(\epsilon) = g(\epsilon)$ for (10.17).
For ordinary temperatures, the Fermi function $f(\epsilon)$ differs from the ground-state step-like behaviour only in a short interval $\Delta\epsilon \simeq k_B T$ around the value $\epsilon = \mu$. This property is exploited to evaluate the dependence on temperature of the integrals (10.18) by a Taylor expansion around the value $\epsilon = \mu$. In most cases, the expansion can be limited to the second order term.
The details of the procedure can be found in § 10.3.1.
The final result, for a second order approximation in Taylor expansion, is

$$\int_0^\infty \phi(\epsilon)\, f(\epsilon)\, d\epsilon \;=\; \int_0^\mu \phi(\epsilon)\, d\epsilon \;+\; \frac{\pi^2}{6}\,(k_B T)^2 \left(\frac{\partial \phi}{\partial \epsilon}\right)_\mu \tag{10.19}$$

The first integral on the right can be splitted:

$$\int_0^\infty \phi(\epsilon)\, f(\epsilon)\, d\epsilon \;=\; \int_0^{\mu_0} \phi(\epsilon)\, d\epsilon \;+\; \int_{\mu_0}^\mu \phi(\epsilon)\, d\epsilon \;+\; \frac{\pi^2}{6}\,(k_B T)^2 \left(\frac{\partial \phi}{\partial \epsilon}\right)_\mu \tag{10.20}$$

The first term in (10.20) corresponds to the ground-state behaviour, where the Fermi function is step-like. The second term accounts for the variation of the chemical potential. The third term is the lowest-order perturbation, which depends on T^2.
For $k_B T \ll \mu$, there is no significative difference if the derivative of the perturbation term is calculated with respect to $\mu_0 = \epsilon_F$ instead of μ.

Chemical potential

Let us come back to (10.17), and substitute $\phi(\epsilon) = g(\epsilon)$ in (10.20). Taking into account that the density of states $g(\epsilon)$ can be considered as a constant $g(\mu_0)$ in the restricted region here considered where $f'(\epsilon) \neq 0$, one finds

$$\begin{aligned} n \;&=\; \int_0^\infty g(\epsilon)\, f(\epsilon)\, d\epsilon \\ &\simeq\; \int_0^{\mu_0} g(\epsilon)\, d\epsilon \;+\; \int_{\mu_0}^\mu g(\epsilon)\, d\epsilon \;+\; \frac{\pi^2}{6}\,(k_B T)^2 \left(\frac{\partial g(\epsilon)}{\partial \epsilon}\right)_{\mu_0} \\ &\simeq\; N \;+\; g(\mu_0)\,(\mu - \mu_0) \;+\; \frac{\pi^2}{6}\,(k_B T)^2\, g'(\mu_0) \end{aligned} \tag{10.21}$$

whence,

$$g(\mu_0)\,(\mu - \mu_0) \;+\; \frac{\pi^2}{6}\,(k_B T)^2\, g'(\mu_0) \;=\; 0 \tag{10.22}$$

and substituting $\mu_0 = \epsilon_F$,

$$\mu \;=\; \epsilon_F - \frac{\pi^2}{6}\,(k_B T)^2\, \frac{g'(\epsilon_F)}{g(\epsilon_F)}\,. \tag{10.23}$$

For the free electron gas model, where the density of states for unit volume is given by (9.11),

$$\mu \;=\; \epsilon_F - \frac{\pi^2}{6}\,(k_B T)^2\, \frac{1}{2\epsilon_F} \;=\; \epsilon_F \left[1 - \frac{1}{12}\left(\frac{\pi k_B T}{\epsilon_F}\right)^2\right]. \tag{10.24}$$

The extent of the deviation of μ at room temperature from the ground-state value ϵ_F can be easily evaluated, by taking into account that the Fermi energy is of the order of the eV, while at $k_B T \simeq 0.025$ eV at 300 K.

Specific heat

Let us now come back to (10.16), and substitute $\phi(\epsilon) = \epsilon\, g(\epsilon)$ in (10.20):

$$\begin{aligned} u &= \int_0^\infty \epsilon\, g(\epsilon)\, f(\epsilon)\, d\epsilon \\ &\simeq \int_0^{\mu_0} \epsilon\, g(\epsilon)\, d\epsilon \;+\; \int_{\mu_0}^{\mu} \epsilon\, g(\epsilon)\, d\epsilon \;+\; \frac{\pi^2}{6}\,(k_B T)^2 \left(\frac{\partial[\epsilon g(\epsilon)]}{\partial \epsilon}\right)_{\mu_0} \\ &\simeq u_0 \;+\; \mu_0 \left[g(\mu_0)\,(\mu-\mu_0) + \frac{\pi^2}{6}\,(k_B T)^2\, g'(\mu_0)\right] \;+\; \frac{\pi^2}{6}\,(k_B T)^2\, g(\mu_0) \end{aligned} \tag{10.25}$$

According to (10.22), the term in square parentheses in (10.25) is zero, so that

$$u \;=\; u_0 \;+\; \frac{\pi^2}{6}\,(k_B T)^2\, g(\mu_0)\,, \tag{10.26}$$

whence the expression (10.12) for the specific heat is derived ($\mu_0 = \epsilon_F$)

$$\boxed{c_v^{\rm el} \;\simeq\; \frac{\pi^2}{3}\, k_B^2\, g(\mu_0)\, T \;=\; \frac{\pi^2}{3}\, k_B^2\, g(\epsilon_F)\, T} \tag{10.27}$$

For the free electron gas model, one obtains the expression (10.13) for the specific heat per unit volume at constant volume.

10.2 Electric conductivity in metals

Let us consider only the effect of the electric field $\vec{E}$, neglecting the magnetic effects, so that the force is $\vec{F} = -e\vec{E}$.

10.2.1 Classical free-electron model (Drude model)

For a gas of free classical electrons, the effect of the electric field $\vec{E}$ would be a continuous increase of the electron momentum:

$$d\vec{p}(t) \;=\; \hbar d\vec{k}(t) \;=\; -e\vec{E}\, dt\,, \qquad \Delta\vec{p} \;=\; -e\vec{E}\, \Delta t\,. \tag{10.28}$$

The Drude model hypothesises that the origin of electrical resistivity is due to the collisions of electrons with the nuclei of the crystal lattice. A basic quantity is thus the average time between collisions (or relaxation time) τ. The average momentum gained between any two collisions is

$$\langle \Delta\vec{p} \rangle \;=\; -e\vec{E}\,\tau\,, \tag{10.29}$$

proportional to the electric field and to the relaxation time.
The drift velocity of electrons is

$$\vec{v}_{\rm dr} \;=\; \langle \vec{v} \rangle \;=\; \frac{\langle \Delta\vec{p} \rangle}{m} \;=\; -\frac{e\vec{E}\tau}{m}\,. \tag{10.30}$$

The current density is

$$\vec{J} \;=\; -ne\vec{v}_{\rm dr} \;=\; \frac{ne^2\tau}{m}\,\vec{E}\,, \tag{10.31}$$

where n is the number density of conduction electron, so that the conductivity is

$$\sigma \;=\; \frac{ne^2\tau}{m}\,. \tag{10.32}$$

By inserting the measured conductivity, one finds that the order of magnitude of the relaxation time is $\tau \simeq 10^{-14}\,$s.

10.2.2 Beyond the classical model

Electrons in a periodic potential

As we have seen, the quantum treatment of electrons in a periodic potential leads to stationary Bloch's wavefunctions

$$\Psi_{\vec{k}n} \;=\; e^{i\vec{k}\cdot\vec{r}}\, u_{\vec{k}n}(\vec{r})\,, \tag{10.33}$$

solutions of the time-independent Schrödinger equation. Two drawbacks of the classical models are evident:

1. The presence of ions is included in the periodic potential energy of the Schrödinger equation. Electrical resistivity cannot be accounted for by collisions of electrons with static ions.
2. Only electrons near the Fermi surface can contribute to conduction; the Fermi surface is present only in crystals where some band is not completely filled.

Particle approach to transport properties

The particle approach is most suitable to describe the transport properties.
Bloch's wavepackets should fulfil the following requirements:

- $\Delta\vec{k}$ small with respect to the size of the 1st Brillouin Zone.
- $\Delta\vec{r}$ large with respect to the lattice parameters, but small with respect to the wavelengths of the external fields.

Origin of resistivity

The origin of resistivity (finite value of conductivity) is connected to collisions of electrons with

- Lattice defects, impurities, etc (contribution independent of temperature)
- Thermal vibrations of ions, say electron-phonon collisions.
  This contribution depends on temperature; when T increases, the phonon density increases. As for thermal resistivity due to phonon-phonon collisions, also for electrical resistivity due to electron-phonon collisions umklapp processes play a fundamental role.

Fermi-Dirac statistics, the Lifschitz model

As for specific heats, only electrons near the Fermi surface contribute to electrical conduction.
A simple model due to E. Lifschitz can help to find a quantitative estimate of conductivity. The model is based on a system where the electron dynamics is described by a Fermi sphere (Fig. 10.1).

If no external field is present (Fig. 10.1, left), opposite values of wavevectors are balanced within the Fermi sphere, and no net charge drift is possible.
When an external electric field $\vec{E}$ is applied along the x direction, the electrons near the Fermi surface undergo a wavevector modification $\hbar\Delta k_x = -eE_x\tau_x$. As a consequence, the shape of the Fermi sphere is modified (Fig. 10.1, centre and right):

1. A layer 1 is created and filled up on the left (negative values of k_x).
2. Correspondingly, a layer 2 of equal volume is emptied on the right.
3. The original Fermi sphere (dashed line), having lost the volume 2, is now unbalanced, and a corresponding volume 3 on the left has to be counted to evaluate the net contribution to electron drift.

The contribution of a volume element of regions 1 and 3 of the reciprocal space

$$dV_k \;=\; (k\,d\theta)\;(k\,\sin\theta\,d\phi)\;dk \;=\; k^2\,\sin\theta\,d\theta\,d\phi\,dk \tag{10.34}$$

to the current density is

$$dJ \;=\; e\,\frac{2\,dV_k}{(2\pi)^3}\,v_F\,\cos\theta\,, \tag{10.35}$$

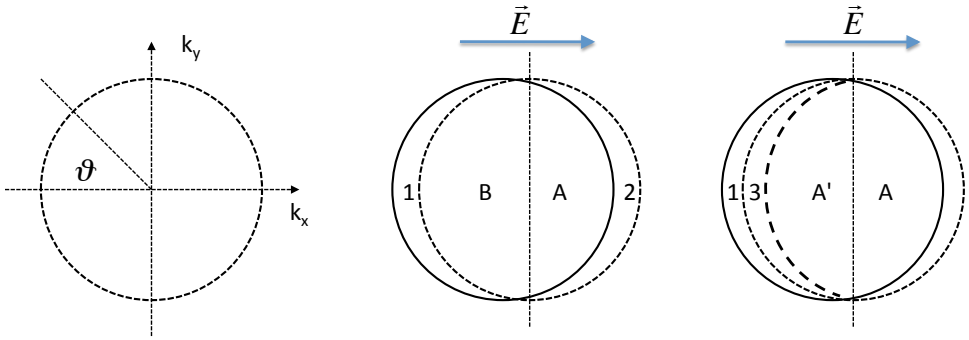

Figure 10.1: Lifschitz model for electrical conduction (two-dimensional projection). Left: Fermi sphere with no external electric field. Centre and right: modifications of Fermi sphere in the presence of an electric field $\vec{E}$.

where $2dV_k/(2\pi)^3$ is the density of conduction electrons and $v_F \cos\theta$ is the x component of the Fermi velocity.
Integrating over the two regions 1 and 3 (factor 2), one finds the current density

$$J = 2\,\frac{2ev_F}{(2\pi)^3}\int_0^{\pi/2} d\theta \int_0^{2\pi} d\phi \int_{k_F}^{k_F+\Delta k_F\cos\theta} k^2\,\sin\theta\,\cos\theta\,dk = \frac{8\pi e v_F k^2 \Delta k}{3(2\pi)^3}\,. \tag{10.36}$$

Substituting $S_F = 4\pi k_F^2$ (Fermi surface) and $\Delta k = eE\tau/\hbar = eE\ell/v_F\hbar$ one obtains the final expression for the electrical conductivity

$$\sigma = \frac{2e^2 S_F \ell}{3\hbar\,(2\pi)^3}\,. \tag{10.37}$$

By expressing the Fermi surface in terms of the Fermi volume $S_F = 3V_F/k_F$, introducing the number density $n = V_F/(2\pi)^3$ and substituting $\ell/\hbar k_F = \tau/m$, one sees that the Lifschitz conductivity corresponds to the conductivity of the Drude model:

$$\sigma = \frac{2e^2 S_F \ell}{3\hbar\,(2\pi)^3} = \frac{e^2 n\tau}{m}\,. \tag{10.38}$$

The relaxation time τ obtained from experimental measurements of σ is the same for both models. The main difference between the two models can be retraced to the evaluation of the mean free path $\ell = v\tau$.

- In the classical model all electrons contribute to conductivity, and the drift velocity, determined by the Maxwell distribution, is of the order of 10^5 m/s, so that the mean free path is $\ell \simeq 10$ Å.
- In the Lifschitz model, only the electrons near the Fermi surface contribute to conduction, the Fermi velocity is of the order of 10^6 m/s and the mean free path is $\ell \simeq 100$ Å.

10.3 Complements and demonstrations

10.3.1 Evaluation of the specific heat integrals

We want to integrate the integrals (10.16) and (10.17), which are of the general type

$$\int_0^\infty \phi(\epsilon)\,f(\epsilon)\,d\epsilon\,, \tag{10.39}$$

where

1. $\phi(\epsilon) = \epsilon\, g(\epsilon)$ for (10.16) and $\phi(\epsilon) = g(\epsilon)$ for (10.17),
2. $f(\epsilon)$ is the Fermi function

$$f(\epsilon) \;=\; \frac{1}{e^{(\epsilon-\mu)/kT}+1}\,. \tag{10.40}$$

It is convenient to integrate (10.39) by parts,

$$\int_0^\infty \phi(\epsilon)\, f(\epsilon)\, d\epsilon \;=\; |\psi(\epsilon)\, f(\epsilon)|_0^\infty \;-\; \int_0^\infty \psi(\epsilon)\, f'(\epsilon)\, d\epsilon\,, \tag{10.41}$$

where

$$\psi(\epsilon) \;=\; \int_0^\epsilon \phi(\epsilon')\, d\epsilon'\,, \qquad \phi(\epsilon) = \frac{d\psi(\epsilon)}{d\epsilon}\,. \tag{10.42}$$

The first term of the right-hand side of (10.41) is zero, because $\psi(\epsilon) = 0$ for $\epsilon = 0$ and $f(\epsilon) = 0$ for $\epsilon \to \infty$.
As for the integral of the second term of the right-hand side of (10.41), the first derivative $f'(\epsilon) = df(\epsilon)/d\epsilon$ is negative and different from zero only in an interval of size $\simeq k_B T$ around $\epsilon = \mu$. Only this interval gives a contribution to the integral.
When $k_B T \ll \mu$, one can expand $\psi(\epsilon)$ around μ

$$\psi(\epsilon) \;=\; \psi(\mu) \;+\; \left(\frac{\partial\psi}{\partial\epsilon}\right)_\mu (\epsilon-\mu) \;+\; \frac{1}{2}\left(\frac{\partial^2\psi}{\partial\epsilon^2}\right)_\mu (\epsilon-\mu)^2 \;+\; \cdots \tag{10.43}$$

Let's now insert (10.43) into (10.41) and consider the first three terms:

First term

$$-\,\psi(\mu)\int_0^\infty f'(\epsilon)\, d\epsilon \;=\; \psi(\mu) \tag{10.44}$$

because

$$\int_0^\infty f'(\epsilon)\, d\epsilon \;=\; f(\epsilon)\big|_0^\infty \;=\; 0-1\,. \tag{10.45}$$

Second term

$$-\left(\frac{\partial\psi}{\partial\epsilon}\right)_\mu \int_0^\infty (\epsilon-\mu)\, f'(\epsilon)\, d\epsilon \;=\; 0 \tag{10.46}$$

since all terms $(\epsilon-\mu)^{2n+1}$ are odd functions and $f'(\epsilon)$ is an even function in the short energy interval considered.

Third term

To evaluate the third term

$$-\,\frac{1}{2}\left(\frac{\partial^2\psi}{\partial\epsilon^2}\right)_\mu \int_0^\infty (\epsilon-\mu)^2\, f'(\epsilon)\, d\epsilon \tag{10.47}$$

it is convenient and to introduce the variable

$$x \;=\; \beta\,(\epsilon-\mu)\,, \tag{10.48}$$

and to explicitly express the derivative of the Fermi function (10.40), so that (10.47) becomes

$$\frac{1}{2}\left(\frac{\partial^2\psi}{\partial\epsilon^2}\right)_\mu \int_0^\infty (\epsilon-\mu)^2\, \frac{\beta\, e^{\beta(\epsilon-\mu)}}{[e^{\beta(\epsilon-\mu)}+1]^2}\, d\epsilon \;=\; \frac{1}{2}\left(\frac{\partial^2\psi}{\partial\epsilon^2}\right)_\mu \int_{-\beta\mu}^\infty \frac{1}{\beta^2}\, \frac{x^2\, e^x}{[e^x+1]^2}\, dx\,. \tag{10.49}$$

Since $\beta\mu = \mu/k_B T >> 1$ and for $x = -\beta\mu$ the integrand is negligible, the lower integration limit can be substituted by $-\infty$ and (10.49) becomes

$$\frac{1}{2}\left(\frac{\partial^2\psi}{\partial\epsilon^2}\right)_\mu (k_b T)^2 \int_{-\infty}^{\infty} \frac{x^2\, e^x}{[e^x+1]^2}\, dx\,. \tag{10.50}$$

By studying the mathematical properties of the integrand in (10.50), one can show [see Ashcroft-Mermin or Reif for details] that

$$-\frac{1}{2}\left(\frac{\partial^2\psi}{\partial\epsilon^2}\right)_\mu \int_0^{\infty} (\epsilon-\mu)^2\, f'(\epsilon)\, d\epsilon \;=\; \frac{\pi^2}{6}\,(k_B T)^2 \left(\frac{\partial^2\psi}{\partial\epsilon^2}\right)_\mu \tag{10.51}$$

Final result

As a final result, (10.41) becomes

$$\int_0^{\infty} \phi(\epsilon)\, f(\epsilon)\, d\epsilon \;=\; \psi(\mu) \;+\; \frac{\pi^2}{6}\,(k_B T)^2 \left(\frac{\partial^2\psi}{\partial\epsilon^2}\right)_\mu \;+\; \cdots \tag{10.52}$$

10.4 Bibliography of Chapter 10

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions). Chapters 1, 2, 12, 13.
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapters 6 and 7.
- W. A. Harrison: *Solid state theory*, Dover Publications 1980.
- N. B. Brandt and S. M. Chudinov: *Electronic structure of metals*, MIR Publishers Moskow 1973.
- F. Reif: *Statistical and thermal physics*, McGraw Hill 1988, Chapter 9 (Fermi-Dirac statistics and electronic specific heat).

Chapter 11

Structural probes

The experimental study of the physical properties of matter – structural, electronic and vibrational – is based on the interaction with suitable probes: electromagnetic radiation, electrons, neutrons, positrons, ions...
Information on the structural properties of bulk matter is obtained by diffraction of suitable probes: X-rays, electrons, neutrons. Diffraction means here elastic scattering plus interference of the scattered waves.
After an introduction to the wave-particle relations for X-rays, electrons and thermal neutrons (§ 11.1) and to the cross section concept (§ 11.2), the attention is focused on the elastic scattering of X-rays from single electrons (§ 11.3) and from atoms (§ 11.4). The atomic cross sections for electrons and thermal neutrons are introduced in § 11.5.
Once the intrinsic cross sections for the different probes have been introduced, the unified treatment of scattering from atomic aggregates is introduced in § 11.6, distinguishing the cases of molecules, non-crystalline systems and crystals.
§ 11.8 is dedicated to the EXAFS technique (Extended X-ray Fine Structure), which can give information on the local structure around selected atomic species.

11.1 Introduction

To investigate the bulk structure of matter by elastic scattering, the probe must fulfil two requirements:

1. Wavelength smaller than the inter-atomic distances, say $\lambda \leq 1$ Å.
2. High enough penetration depth, say sufficiently weak interaction of the probe with matter .

11.1.1 X-rays, electrons, neutrons

The most important structural probes of the bulk structure of matter are

1. X-rays, discovered in 1895 by W.K. Röntgen in Würtzburg. The first diffraction of X-rays from crystals was obtained in 1912 by M. von Laue in Münich. The utilisation of X-ray for the quantitative study of the atomic arrangement began in 1913, thanks to the pioneering work of W.H. Bragg and W.L. Bragg in Cambridge.
2. Electrons, discovered in 1897 by J.J. Thomson in Cambridge. The first electron diffraction experiments were performed in 1927, confirming the wave-particle duality, contemporarily by
   a) C. Davisson and L.H. Germer at Bell Labs. (USA), using electrons of 30-600 eV
   b) G.P. Thomson and A. Reid at the University of Aberdeen, using electrons of 10-60 keV
3. Neutrons, discovered by J. Chadwick in 1932. The first diffraction pattern from NaCl was obtained in 1946 at the Oak-Ridge reactor in Tennessee.

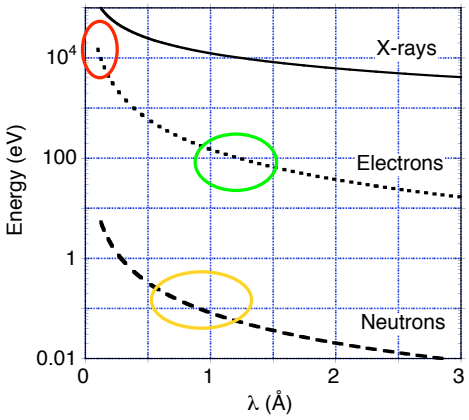

Figure 11.1: Relation between energy and wavelength for X-rays, electrons and thermal neutrons. The ellipses show the regions of use.

11.1.2 Particle and wave properties

The complementarity of the wave and particle behaviour is relevant for the structural probes. Different phenomena are described by different quantum approaches: particle-like approach for production, absorption, scattering, wave-like approach for interference.

The **wave properties** of a structural probe are described in terms of the plane-wave parameters:

1. the angular frequency $\omega = 2\pi\nu = 2\pi/T$, where ν is the frequency and T is the period,
2. the wavevector $\vec{k} = (2\pi/\lambda)\hat{s}$, where λ is the wavelength $\hat{s}$ is a unit vector

Real waves are actually a superposition of plane waves of different frequencies and wavevectors. The degree of monochromaticity, say the width of the distribution of frequencies ω, corresponds to the longitudinal coherence of the beam. The degree of parallelism of the wavevectors $\vec{k}$ corresponds to the transverse coherence.

The **particle properties** are the (kinetic) energy E and the linear momentum $\vec{p}$, connected by the relativistic expression

$$E^2 = (pc)^2 + (m_0c^2)^2\,, \tag{11.1}$$

where $c \simeq 3{\times}10^8$ m s^{-1} is the velocity of light and m_0 is the rest mass.

- For massless particles, such as photons, $E = pc = \hbar\omega$
- For massive particles, such as neutrons and electrons, the non-relativistic kinetic energy is $E_k = p^2/2m$

Wave-particle connection. The wave and particle approaches are connected by the concept of wave-packet. The general relations hold

$$\vec{p} = (h/\lambda)\hat{s} = \hbar\vec{k}\,, \qquad E = \hbar\omega\,, \tag{11.2}$$

where $h \simeq 4.13 \times 10^{-15}$ eV s is the Planck constant and $\hbar = h/2\pi \simeq 6.58 \times 10^{-16}$ eV s.
Here $\vec{k}$ represents the central value of the distribution of wave-vectors characterising the wave-packet.

Different **dispersion relations** $\omega(k)$ characterise massless and massive particles:

- for massless particles (photons), the dispersion relation is linear: $\omega = ck$,
- for massive particles, the dispersion relation is not linear, $\omega = \hbar k^2/2m$; the particle velocity is the wave group velocity, $v = d\omega/dk$.

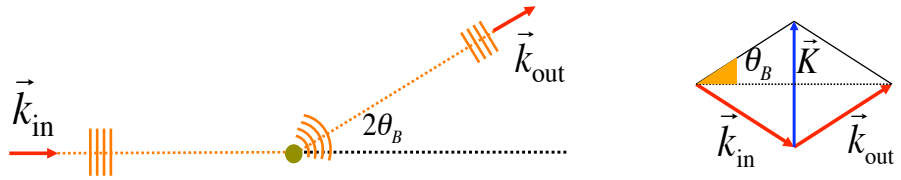

Figure 11.2: Nomenclature of scattering. Left: incoming and outgoing wave-vectors and scattering angle $2\theta_B$. Right: scattering vector $\vec{K}$ for elastic scattering.

The **relation between energy and wavelength** is different for X-rays, electrons and thermal neutrons:

X-rays:	$E\,[\mathrm{keV}] = 12.4/\lambda\,[\mathrm{Å}]$	so that	$\lambda = 1\,\mathrm{Å}$	$\Rightarrow$	$E = 12.4\,\mathrm{keV}$
Electrons:	$E_k\,[\mathrm{eV}] = 150/\lambda^2\,[\,\mathrm{Å}^2]$		$\lambda = 1\,\mathrm{Å}$	$\Rightarrow$	$E = 150\,\mathrm{eV}$
Thermal neutrons:	$E_k\,[\mathrm{meV}] = 82/\lambda^2\,[\mathrm{Å}^2]$		$\lambda = 1\,\mathrm{Å}$	$\Rightarrow$	$E = 82\,\mathrm{meV}$

The relations are shown in Fig. 11.1. Electrons of relatively high energy (50-100 keV, corresponding to wavelengths $\lambda \simeq 0.05\,\mathrm{Å}$) are generally used for diffraction experiments. The high energy allows a relatively high penetration depth. Low energy electron diffraction (LEED, energies 10-300 eV) is used for structural studies of surfaces.

Note: We have used above the relation for non-relativistic massive particles $E_k = p^2/2m$, which gives $\lambda = h/[2mE_k]^{1/2}$. To take into account relativistic effects, from (11.1) one obtains $p^2 = 2m_0E_k\,[1 + E_k/2m_0c^2]$, whence

$$\lambda = \frac{h}{2m_0E_k}\,\frac{1}{\sqrt{1 + E_k/2m_0c^2}}$$

For our present purposes, the relativistic effects can be considered negligible.

11.2 Scattering cross-section

11.2.1 Nomenclature of scattering

Let us consider a narrow beam of particles (photons, electrons, neutrons) impinging on a sample along a fixed direction (Fig. 11.2, left). We want to measure the intensity of the beam scattered by the sample within a given narrow solid angle $d\Omega$ as a function of the output direction. The output direction is characterised by two angles:

- a polar angle, or scattering angle $\vartheta = 2\theta_B$ with respect to the incoming direction; in diffraction problems one generally refers to the Bragg angle $\theta_B = \vartheta/2$;
- an azimutal angle ϕ that measures the rotation around the incidence direction.

The input and output beams are characterised by:

- wavevectors $\vec{k}_{\rm in}$ and $\vec{k}_{\rm out}$ $(k = 2\pi/\lambda)$;
- energies $E_{\rm in} = \hbar\omega_{\rm in}$ and $E_{\rm out} = \hbar\omega_{\rm out}$;
- polarisation unit vectors $\hat{\epsilon}_{\rm in}$ and $\hat{\epsilon}_{\rm out}$; for X-rays the polarization is given by the electric field direction, for neutrons and electrons by their spin.

We define

- the exchanged energy $E = E_{\rm out} - E_{\rm in}$,
- the exchanged momentum $\hbar\vec{K} = \hbar(\vec{k}_{\rm out} - \vec{k}_{\rm in})$, where $\vec{K}$ is the scattering vector.

The scattering is said to be

a) elastic if $E = 0$, say if $E_{\rm out} = E_{\rm in}$, say $|\vec{k}_{\rm out}| = |\vec{k}_{\rm in}|$;
b) inelastic if $E \neq 0$, say if $E_{\rm out} \neq E_{\rm in}$, so that $|\vec{k}_{\rm out}| \neq |\vec{k}_{\rm in}|$.

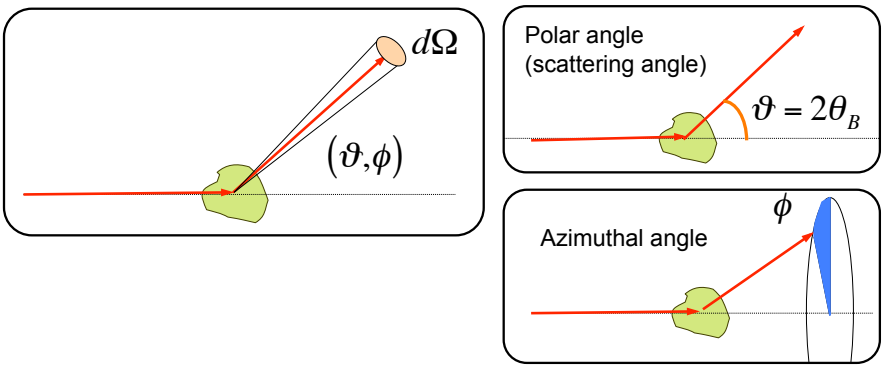

Figure 11.3: Scattering solid angle (left); polar and azimutal angles (right).

Scattering vector

The *scattering vector* is defined as the difference between outgoing and incoming wavevectors:

$$\vec{K} = \vec{k}_{\text{out}} - \vec{k}_{\text{in}} \,. \tag{11.3}$$

In elastic scattering, both beams have the same wavelength, so that both wavevectors have the same magnitude $k_{\text{in}} = k_{\text{out}} = 2\pi/\lambda$.
From simple geometrical considerations (Fig. 11.2, right), one can see that the magnitude of the scattering vector for elastic scattering is

$$K = 4\pi \frac{\sin\theta_B}{\lambda} \,. \tag{11.4}$$

For a given wavelength λ, the scattering vector magnitude is limited to $K_{\text{max}} = 4\pi/\lambda$.

11.2.2 Definition of scattering cross section

The intensity of the collimated incoming beam is measured by the flux Φ_{in}, expressed for example as the number of particles per unit area and per unit time.
The intensity of the divergent outgoing beam is measured by the number dn_{out} of particles per unit time scattered within the solid angle $d\Omega$ with direction (ϑ, ϕ) (Fig. 11.3).

The differential cross section $\sigma(\vartheta, \phi)$ (or $d\sigma/d\Omega$) for *elastic scattering* is defined by the relation

$$dn_{\text{out}} \;=\; \Phi_{\text{in}}\, \sigma(\vartheta, \phi)\, d\Omega \;=\; \Phi_{\text{in}} \frac{d\sigma}{d\Omega}\, d\Omega \,. \tag{11.5}$$

The total cross section (for elastic scattering) is

$$\sigma_{\text{tot}} = \int \sigma(\vartheta, \phi)\, d\Omega \,. \tag{11.6}$$

The double differential cross section $\sigma(\vartheta, \phi, \omega)$ for *inelastic scattering* is defined by the relation

$$dn_{\text{out}} \;=\; \Phi_{\text{in}}\, \sigma(\vartheta, \phi, E)\, d\Omega\, dE \;=\; \Phi_{\text{in}} \frac{d^2\sigma}{d\Omega\, dE}\, d\Omega\, dE \,. \tag{11.7}$$

Both elastic and inelastic cross sections generally depend on the energy of the incoming beam.

Alternative definition

The cross sections can be defined also within the wave-like approach. Instead of considering the flux of particles, one can consider the flux of energy.
The differential scattering cross section is the ratio between the angular density of power emitted in a given direction and the incoming power per unit area of the beam.

Factorization of the scattering cross sections

The differential cross section for **elastic scattering** can be factorized as

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_0 S(\vec{K}) \,, \tag{11.8}$$

where $\vec{K}$ is the scattering vector ($\hbar\vec{K}$ is the exchanged momentum).
The differential cross section for **inelastic scattering** can be factorized as

$$\frac{\partial^2\sigma}{\partial\Omega\,\partial E} = \left(\frac{d\sigma}{d\Omega}\right)_0 S(\vec{K}, E) \,, \tag{11.9}$$

where $\hbar\vec{K}$ is the exchanged momentum and E is the exchanged energy.

The first factor on the right-hand side of both equations, $(d\sigma/d\Omega)_0$, is the intrinsic cross section, which depends on the coupling between the probe and the sample. The intrinsic cross section depends on the nature of the probe, it is different for X-rays, neutrons or electrons; it is instead independent of the structure of the sample.

The second factors on the right-hand sides of (11.8) and (11.9) are, respectively:

- the static scattering function (or static structure function) $S(\vec{K})$
- the dynamic scattering function (or dynamic structure function) $S(\vec{K}, E)$

The scattering functions contain the information on the structure and the possible excitations of the sample. They are independent of the peculiar type of interaction of the probe with matter.

11.2.3 Cross section and scattering amplitude

The scattering cross section is the quantity which is measured in scattering experiments.
To extract structural and dynamical information from the measured cross sections, one needs a suitable theoretical model. In general, the output of theoretical models are scattering amplitudes $f(\vartheta, \phi)$, which are connected to the experimental cross sections by the relation

$$\sigma(\vartheta, \phi) \;=\; |f(\vartheta, \phi)|^2 \,. \tag{11.10}$$

The expressions of the scattering amplitudes are different for different types of interaction between probe and matter (e.g. electromagnetic interaction, scattering from a central potential, strong nuclear interaction).

Example: A particle freely moving in space in the z direction can be described as a wave-packet, superposition of plane waves $\exp[i(kz - \omega t)]$, with a given distribution of k values.
A particle scattered by a central potential $V(r)$ can be described in a similar way, provided each stationary plane wave $\Psi_k^{\text{free}} = \exp(ikz)$ is substituted by the stationary scattering wave-function

$$\Psi_k^{\text{scatt}}(\vec{r}) \;\simeq\; e^{ikz} \;+\; f_k(\vartheta, \phi)\,\frac{e^{ikr}}{r} \,. \tag{11.11}$$

The quantity $f_k(\vartheta, \phi)$ is the scattering amplitude from the central potential and is connected to the experimental scattering cross section by $\sigma(\vartheta, \phi) \;=\; |f_k(\vartheta, \phi)|^2$.

In the following § 11.3, we will study the scattering amplitude and the intrinsic scattering cross section for X-rays from *single electrons*. In most applications it is more convenient to consider the intrinsic cross section of X-rays from *atoms* (§ 11.4). The atomic intrinsic cross sections for electrons and neutrons scattering are introduced in § 11.5 and compared with the X-ray cross section.
Once the intrinsic cross sections for the different probes have been compared, the static scattering function $S(\vec{K})$ is studied in § 11.6 for different types of atomic aggregates (molecules, crystalline solids, liquids and non-crystalline solids).

11.3 X-ray elastic scattering

To better understand diffraction effects, it is convenient to focus the attention on X-rays and consider first the interaction with one point-like electron, then the interaction with a set of point-like electrons and finally the interaction with a distribution of electron charge.

11.3.1 Interaction of X-rays with matter

In the range of energies of interest here (1 to 100 keV), there are two channels of interaction of X-rays with matter: photoelectric absorption or emission and scattering, the first one being predominant. In § 11.3 and 11.4 the scattering channel is considered, in § 11.8 photoelectric absorption is considered.
The interaction of X-rays with matter can be treated at different levels of approximation.

Classical approximation

The classical Thomson theory considers the interaction of an X-ray electromagnetic wave with a free electron. The electric field induces an oscillatory motion of the electron; the accelerated electron in turn emits electromagnetic radiation of the same frequency, with a π phaseshift. The Thomson model is straightforwardly extended to the treatment of continuous distributions of charge (or of probability distributions, such as quantum orbitals).
The free-electron approximation is justified by the high energy of X-rays with respect to the binding energy of the majority of atomic electrons. As we will see, only for very loosely bound electrons can the inelastic quantum Compton effect be relevant and anyway it is automatically taken into account by the Thomson treatment of the scattering from atomic orbitals.
The Thomson theory is effective in explaining the main properties of X-ray elastic scattering from atoms, and it is sufficient for our present purposes.
In the classical approximation, the average density of energy carried by a monochromatic X-ray plane wave is measured by the amplitude of the electric field, or alternatively by the amplitude of the vector potential:

$$\langle u \rangle \;=\; \epsilon_0 \, E_0^2/2 \;=\; \epsilon_0 \omega^2 A_0^2/2 \tag{11.12}$$

Quantum treatment

As it is customary to simplify notation in the description of the electromagnetic field, we consider a cubic box of volume V in the real space. The wave-vectors of the electromagnetic plane waves are limited, by the periodic boundary conditions, to a three-dimensional lattice in the reciprocal space, whose density (number of wave-vectors for unit reciprocal space volume) is $V/8\pi^3$. To each wave-vector, two independent polarisation directions are connected.

The quantum system considered is the sum of an atom and the electromagnetic field. A stationary state of the system is the product

$$|\Psi\rangle \;=\; |\Psi^{\mathrm{at}}\rangle \;\otimes\; |\{n_{\vec{k}s}\}\rangle \,, \tag{11.13}$$

where $|\Psi_{\rm at}\rangle$ is the state of the atom and $|\{n_{\vec{k}s}\}\rangle$ is the state of the electromagnetic field, quantised in terms of the number $n_{\vec{k}s}$ of photons belonging to each normal mode $(\vec{k}, s)$, where $\vec{k}$ is the wavevector and s labels the polarisation.

To account for the interaction of electrons with the electromagnetic field, the atomic Hamiltonian (in the Coulomb gauge $\vec{\nabla} \cdot \vec{\mathcal{A}} = 0$) is modified as

$$H \;=\; \sum_j \left\{ \frac{1}{2m} \left[\vec{P}_j + e\vec{\mathcal{A}}(\vec{r}_j, t) \right]^2 \right\} \;+\; V(\vec{r}_1 \ldots \vec{r}_n) \tag{11.14}$$

where the sum is over the electrons, $\vec{P}_j = m\vec{v}_j - e\vec{\mathcal{A}}$ is the generalised momentum operator, $\vec{\mathcal{A}}(\vec{r}, t)$ is the vector potential of the electromagnetic field calculated at the electron position $\vec{r}_j$ and e is the (positive) elementary charge.

By expanding the square in (11.14) and introducing the Hamiltonian of the radiation field $H_{\rm rad}$, one can express the total Hamiltonian as the sum

$$H_{\rm tot} \;=\; H_{\rm at} \;+\; H_{\rm rad} \;+\; H_{\rm int} \tag{11.15}$$

where

$$H_{\rm at} \;=\; \sum_j P_j^2/2m + V(\vec{r}_1 \ldots \vec{r}_n) \tag{11.16}$$

$$H_{\rm rad} \;=\; \sum_{\vec{k}s} \hbar\omega_{\vec{k}s} \left[a^\dagger_{\vec{k}s} a_{\vec{k}s} + \frac{1}{2} \right] \tag{11.17}$$

$$H_{\rm int} \;=\; \frac{e}{m} \sum_j \vec{\mathcal{A}}(\vec{r}, t) \cdot \vec{P}_j \;+\; \frac{e^2}{2m} \sum_j \mathcal{A}^2(\vec{r}_j, t) \tag{11.18}$$

The vector potential operator is

$$\vec{\mathcal{A}}(\vec{r}_j, t) \;=\; \sum_{\vec{k}s} \sqrt{\frac{\hbar}{2\epsilon_0 V \omega_{\vec{k}s}}} \left[a_{\vec{k}s}(t) e^{-i\vec{k}\cdot\vec{r}} + a^\dagger_{\vec{k}s}(t) e^{i\vec{k}\cdot\vec{r}} \right] \hat{\epsilon}_{\vec{k}s} \tag{11.19}$$

and

$$a_{\vec{k}s}(t) = a_{\vec{k}s}(0)\, e^{i\omega t}\,, \quad a^\dagger_{\vec{k}s}(t) = a^\dagger_{\vec{k}s}(0)\, e^{-i\omega t} \tag{11.20}$$

are the annihilation and creation operators for modes $(\vec{k}s)$.

It is worth stressing a fundamental difference between the classical and the quantum description. In quantum mechanics, the energy content of the electromagnetic field is given by the state vector (11.13); the electric and magnetic fields, as well as the vector potential, are operators – linear combinations of annihilation and creation operators – acting on the state vector.

The term (11.18) in the total Hamiltonian describes the interation of the electromagnetic field with matter. The interaction effect is generally evaluated within the framework of the time-dependent perturbation theory; to first order (Fermi golden rule) the transition probability per unit time from an initial to a final stationary state of the system (atom+field) is

$$w_{fi} \;=\; \frac{2\pi}{\hbar} \left| \langle \Psi_f | H_I | \Psi_i \rangle \right|^2 g(E_f)\,, \tag{11.21}$$

where $g(E_f)$ is the density of final states.

An elastic scattering event corresponds to the transition from an initial stationary state to a final stationary state, both expressed as (11.13), where the final and initial states of the atom are identical, while the population of the different photon states is modified: one photon of the initial state is annihilated and one photon is created in the final state; the process is a two-photons process.

The elastic scattering of X-rays (two-photons process) is accounted for by the second term in (11.18), which is quadratic in $\vec{\mathcal{A}}$ and thus contains products of creation and annihilation operators.

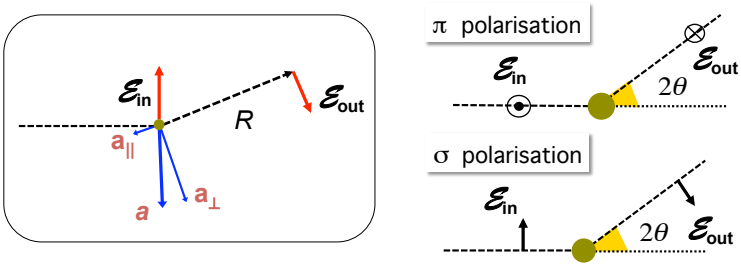

Figure 11.4: Left: relation between the electric field $\vec{\mathcal{E}}_{\rm in}$ acting on an electron, the electron acceleration $\vec{a}$ and the emitted electric field $\vec{\mathcal{E}}_{\rm out}$. Right: electric field perpendicular (top, π polarisation) and parallel (bottom, σ polarisation) to the scattering plane.

Semi-classical approximation

The semiclassical approximation is based on a quantum treatment of the atom, represented by a vector in Hilbert space $|\Psi_{\rm at}\rangle$, and a classical treatment of the electromagnetic field. The interaction Hamiltonian is still expressed by (11.18) and the transition probability is still evaluated by (11.21), but the vector potential is considered as a classical external field.
The semiclassical approximation is often used to describe the absorption and emission of single photons, where the attention is focussed only on the atom and the first term of (11.18) is relevant. We will rely on the semi-classical approximation when dealing with the spectroscopyc EXAFS technique in § 11.8.

11.3.2 Thomson scattering from a point-like free electron

Let us consider an electron at a position $\vec{r}$ with respect to a preset reference frame. The beam impinging on the electron is described by an oscillating electric field

$$\vec{\mathcal{E}}_{\rm in} = \vec{\mathcal{E}}_0 \, \cos\left(\omega t - \vec{k}_{\rm in} \cdot \vec{r}\right) \tag{11.22}$$

The effect of the electric field is the electron acceleration (Fig. 11.4, left)

$$\vec{a} = (-e/m)\, \vec{\mathcal{E}}_{\rm in} \, . \tag{11.23}$$

According to classical electrodynamics, an accelerated charge emits electromagnetic radiation. The electron emits an electromagnetic wave with the same frequency of the impinging wave.

Under the following assumptions (the so called *dipole approximation*):

a) the velocity of the charged particle is much smaller that the velocity of light, $v \ll c$,
b) the size of the particle is much smaller than the X-ray wavelength,
c) the emitted electric field is measured at a distance from the charged particle much larger than the X-rays wavelength,

the outgoing electric field at position $\vec{R}$ (with respect to the scattering electron) and time t is

$$\vec{\mathcal{E}}_{\rm out}(\vec{R}, t) = \frac{e}{4\pi\epsilon_0} \, \frac{\vec{a}_\perp(t')}{Rc^2} \, , \tag{11.24}$$

where $\vec{a}_\perp$ is the projection of the electron acceleration perpendicular to $\vec{R}$ and the difference $t - t'$ takes into account the finite propagation time of the electromagnetic wave.
Inserting (11.23) into (11.24) and taking into account the different projections of the acceleration for different electric field polarisations (within the plane of scattering and perpendicular to it), one

finds the following expression for the magnitude of the electric field at position $\vec{R}$ field (Fig. 11.4, right)

$$\mathcal{E}_{\rm out}(\vec{R},t) = -\frac{r_e}{R}\,\mathcal{E}_{\rm in}(t') \;\times\; \begin{cases} 1 & \text{for } \pi \text{ polarization} \\ & \text{(perpendicular to scattering plane)} \\ \cos(2\theta) & \text{for } \sigma \text{ polarization} \\ & \text{(in the scattering plane)} \end{cases} \tag{11.25}$$

The quantity

$$r_e \;=\; \frac{e^2}{4\pi\epsilon_0 c^2 m} \;=\; 2.8\times10^{-5}\ \text{Å} \tag{11.26}$$

is the Thomson scattering length and measures of the strength of the elastic scattering of X-rays by an electron. The quantity r_e is also called the classical electron radius (see § 11.10.1).

Concerning the polarisation dependence of (11.25), it is worth noting that:

a) X-ray beams emitted by laboratory sources are un-polarised, say an equal mixture of the two polarisations considered in (11.25).
b) X-rays emitted by synchrotron radiation sources are strongly polarised within the horizontal plane of the electron orbit; in a vertical scattering plane, the X-rays are π polarised; if scattering within all space is considered (all values of the azimutal angle ϕ are considered), X-rays are on the average unpolarised.

Problem: The force exerted on the electron by the electromagnetic field is the Lorentz force; when is the effect of the magnetic field negligible? We consider only the effect of the electric field on the electrons; show that the effect on the positive nuclear charge can be neglected.

11.3.3 Scattering amplitude and scattering intensity

Let us develop a formalism suitable for X-ray diffraction experiments, say for connecting the relative position of electrons to the diffraction patterns. To this purpose, it is convenient to consider the waves at a *fixed time* and relate their phase variations only to the position coordinates. To simplify the notation, we first consider only polarisation perpendicular to the scattering plane (π polarisation in Fig. 11.4).

Input amplitude

The electric field at the position $\vec{r}$ of an electron (Fig. 11.5) is

$$\mathcal{E}_{\rm in} = \text{Re}\left\{\mathcal{E}_0\, e^{-i\vec{k}_{\rm in}\cdot\vec{r}}\right\} = \text{Re}\left\{A_0\right\}. \tag{11.27}$$

The quantity within curled brackets is called *input amplitude* A_0.
The input flux, measured by the power carried by the impinging beam, is (energy density) $\times$ (velocity of light), say (in W/m^2)

$$\Phi_{\rm in} \;=\; \frac{1}{2}\epsilon_0\mathcal{E}_0^2 c \;=\; \frac{1}{2}\epsilon_0\,|A_0|^2\,c\,. \tag{11.28}$$

The input flux, measured by the number of photons per second and unit beam section is

$$\Phi_{\rm in} \;=\; \frac{\epsilon_0\mathcal{E}_0^2 c}{2\hbar\omega}\,. \tag{11.29}$$

Example: At the ID15 high-energy scattering beamline of the Synchrotron Radiation Facility ESRF (Grenoble), a flux of 10^{11} photons/second impinges on a 1×1 mm^2 area within an energy interval of 0.1% bandwidth around 50 keV.

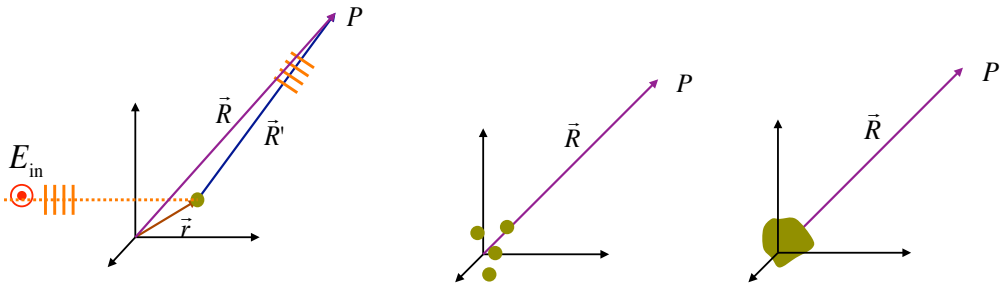

Figure 11.5: Left: scattering from a point-like electron at position $\vec{r}$. Center: scattering from a set of point-like electrons. Right: scattering from a continuous distribution of charge.

Scattering amplitude

The electric field at a point P, at position $\vec{R}$ from the reference origin ($\vec{R}'$ from the scattering electron), is the real part of the *scattering amplitude* A:

$$\mathcal{E}_{\rm out} = \mathrm{Re}\left\{\mathcal{E}_0 e^{-i\vec{k}_{\rm in}\cdot\vec{r}}\,(-r_e)\,\frac{e^{-ik_{\rm out}R'}}{R'}\right\} = \mathrm{Re}\,\{A\}\ . \tag{11.30}$$

If $R' \simeq R \gg r$, one can express R' (Fig. 11.5, left) as

$$R' \ \simeq\ R\ -\ r\,(\hat{k}_{\rm out}\cdot\hat{r}) \tag{11.31}$$

so that

$$\frac{e^{-ik_{\rm out}R'}}{R'} \quad\to\quad \frac{e^{-ik_{\rm out}R}}{R}\,e^{+i\vec{k}_{\rm out}\cdot\vec{r}} \tag{11.32}$$

The (complex) scattering amplitude A can now be written as

$$A \ =\ \underbrace{-\mathcal{E}_0\,r_e\,\frac{e^{-ik_{\rm out}R}}{R}}_{A_{\rm el}}\ \underbrace{e^{i(\vec{k}_{\rm out}-\vec{k}_{\rm in})\cdot\vec{r}}}_{\rm phase}\ =\ A_{\rm el}\ e^{i\vec{K}\cdot\vec{r}}\,, \tag{11.33}$$

say as the product of two factors

1. a one-electron constant $A_{\rm el}$, which depends on the scattering length r_e and on the distance R but not on the position $\vec{r}$ of the electron nor on the scattering vector $\vec{K}$,
2. a phase factor $\exp(i\vec{K}\cdot\vec{r})$ that depends on the electron position $\vec{r}$ and on the scattering vector $\vec{K} = \vec{k}_{\rm out} - \vec{k}_{\rm in}$.

For X-rays, it is customary to define the scattering amplitude in electron units (e.u.)

$$A_{\rm e.u.}(\vec{K}) \ =\ \frac{A(\vec{K})}{A_{\rm el}}\ =\ e^{i\vec{K}\cdot\vec{r}}\,. \tag{11.34}$$

Scattering intensity

The outgoing electric field $\mathcal{E}_{\rm out} = \mathrm{Re}\{A(\vec{K})\}$ cannot be directly measured, not only for the experimental impossibility of dealing with its exceedingly high frequency, of the order of 10^{19} Hz), but even for a more fundamental quantum reason: the better is defined the energy of the photons, the less is defined the phase of the electric field.

One actually measures the power carried by the outgoing beam, which is proportional to the square of the electric field. Otherwise stated, one measures the intensity

$$I(\vec{K}) \ =\ |A(\vec{K})|^2 \ =\ |A_{\rm el}(\vec{K})|^2 \ =\ \frac{\mathcal{E}_0^2}{R^2}\,r_e^2\,. \tag{11.35}$$

For an unpolarised beam, taking into account (11.25), eq. (11.35) becomes

$$I(\vec{K}) \;=\; \frac{\mathcal{E}_0^2}{R^2}\, r_e^2 \left[\frac{1}{2} + \frac{\cos^2(2\theta_B)}{2}\right], \tag{11.36}$$

where the quantity in square brackets is the polarisation factor.
In electronic units, the intensity scattered by one electron is

$$I_{\text{e.u.}}(\vec{K}) \;=\; \frac{I(\vec{K})}{I_{\text{el}}(\vec{K})} \;=\; 1\,, \tag{11.37}$$

where I_{el} includes the polarisation factor.

Thomson cross section

The intensity (11.36) is collected by a detector at distance R and corresponds to a power (in W) emitted in the solid angle $d\Omega$

$$\mathcal{P}(\vec{K})\, d\Omega \;=\; \frac{1}{2}\epsilon_0\, \mathcal{E}_0^2\, c\, r_e^2 \left[\frac{1}{2} + \frac{\cos^2(2\theta_B)}{2}\right] d\Omega \;=\; \Phi_{\text{in}}\, r_e^2 \left[\frac{1}{2} + \frac{\cos^2(2\theta_B)}{2}\right] d\Omega\,. \tag{11.38}$$

The Thomson cross section for X-ray scattering from one electron is thus

$$\sigma(2\theta_B, \phi) \;=\; r_e^2 \left[\frac{1}{2} + \frac{\cos^2(2\theta_B)}{2}\right], \tag{11.39}$$

where $r_e^2 = 7.84 \times 10^{-10}$ Å^2.
Apart from the angular dependence due to the polarisation factor, the scattering cross section doesn't depend on the scattering direction nor on the X-ray wavelength. The physical reason is that the size of the electron, measured by the classical radius r_e, is much smaller than the X-ray wavelength, so that all elements of the electron scatter in phase.

The total scattering cross section for Thomson scattering from one electron is obtained by integrating (11.39) over the full solid angle (see § 11.10.2):

$$\sigma_{\text{Th}} \;=\; \int \sigma(2\theta_B, \phi)\, d\Omega \;=\; \frac{8}{3}\,\pi\, r_e^2 \;=\; 66.6 \times 10^{-10\,2} \tag{11.40}$$

and is independent of the X-ray wavelength.

Example: Let us consider again an X-ray flux of 10^{11} photons/(s mm^2). The total Thomson cross section is 66.6×10^{-23} mm^2. This means that a given electron scatters an average of 6.66×10^{-12} photons per second.

11.3.4 Scattering from Z point-like electrons

Let us now consider the scattering by a number Z of point-like electrons, at positions $\vec{r}_i$, where $i = 1, 2, ...Z$ (Fig. 11.5, centre). The complex scattering amplitude $A(\vec{K})$ is the sum of the scattering amplitudes of all electrons:

$$A(\vec{K}) \;=\; A_{\text{el}} \sum_{i=1}^{Z} e^{i\vec{K}\cdot\vec{r}_i}\,, \qquad A_{\text{e.u.}}(\vec{K}) \;=\; \sum_{i=1}^{Z} e^{i\vec{K}\cdot\vec{r}_i}\,. \tag{11.41}$$

The measured intensity is

$$I_{\text{e.u.}}(\vec{K}) \;=\; \left|\sum_{i=1}^{Z} e^{i\vec{K}\cdot\vec{r}_i}\right|^2. \tag{11.42}$$

The sum of the phase factors of each electron gives rise to the interference effect. One can easily verify (see eq. 11.179 of § 11.10.3) that

$$I_{\text{e.u.}}(\vec{K}) \;=\; Z \;+\; \sum_i \sum_{j\neq i} \cos(\vec{K}\cdot\vec{r}_{ij})\,, \qquad \text{where} \;\; \vec{r}_{ij} = \vec{r}_j - \vec{r}_i\,. \tag{11.43}$$

The first term Z on the right side of (11.43) is the effect of independent scattering from Z electrons, the sum of the remaining $Z(Z-1)$ terms is the effect of interference. Note that the interference effect depends only on the relative positions $\vec{r}_{ij}$ of electrons, not on their absolute positions $\vec{r}$.
As a consequence of interference, the scattering intensity now depends on the scattering vector $\vec{K}$, say on scattering angle and X-rays wavelength.

Note: Equations (11.41) and (11.42) are basic for diffraction from atomic aggregates, in spite of their present derivation for the quite unrealistic model of point-like electrons. Actually, they account for more realistic cases provided the Thomson scattering length r_e is substituted by the neutron scattering lengths (or thermal neutron scattering) or by the atomic form factors (for X-rays or electron scattering) (see below, § 11.5).

11.3.5 Scattering from a continuous electron charge distribution

A more elaborate picture of the electronic structure of matter is represented by a continuous charge distribution, measured by the number density $\rho_{\text{el}}(\vec{r})$. The simplest example is the distribution of electronic charge within an atom (see § 11.4), but one can consider the distribution of electronic charge within a large aggregate of atoms, such as a crystal.

Scattering amplitude

The scattering amplitude of X-rays from an electron distribution is obtained from (11.41) substituting the sum by an integral

$$A(\vec{K}) = A_{\text{el}}(\vec{K}) \int \rho_{\text{el}}(\vec{r})\, e^{\mathrm{i}\vec{K}\cdot\vec{r}}\, dV\,, \qquad A_{\text{e.u.}}(\vec{K}) = \int \rho_{\text{el}}(\vec{r})\, e^{\mathrm{i}\vec{K}\cdot\vec{r}}\, dV\,. \tag{11.44}$$

The scattering amplitude in electronic units $A(\vec{K})$, defined in the reciprocal space, is the *Fourier transform of the electron density* $\rho_{\text{el}}(\vec{r})$, defined in the real space.

Equation (11.44) is basic for X-ray diffraction. If one could measure the scattering amplitude $A(\vec{K})$ in a conveniently large region of the reciprocal space, one could recover, by Fourier inversion, the density of electronic charge $\rho_{\text{el}}(\vec{r})$ with the required accuracy, and get information on the electron structure of atoms, molecules and condensed systems. This in turn would mean to get information on atomic positions (structural information) and on electron distribution among atoms (information on chemical bonds).

Scattering intensity

As already observed, the amplitude $A(\vec{K})$ cannot be directly measured. One can measure the intensity $I(\vec{K}) = |A(\vec{K})|^2$, from which the amplitude can be obtained only to within a phase factor. Some information content is thus generally lost. The so called "phase problem" is of central importance for the determination of complex structures.
Let us explicitly evaluate the scattering intensity:

$$\begin{aligned} I_{\text{e.u.}}(\vec{K}) \;&=\; \int \rho(\vec{r}_1)\, e^{i\vec{K}\cdot\vec{r}_1} dV_1 \int \rho(\vec{r}_2)\, e^{-i\vec{K}\cdot\vec{r}_2} dV_2 \\ &=\; \int \left[\int \rho(\vec{r}_1)\, \rho(\vec{r}_1+\vec{r})\, dV_1\right] e^{-i\vec{K}\cdot\vec{r}} dV\,, \end{aligned} \tag{11.45}$$

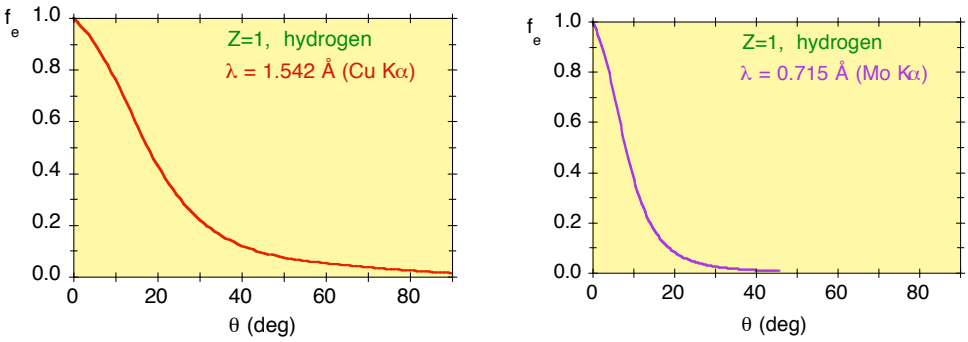

Figure 11.6: Atomic scattering factor for Hydrogen ($Z = 1$) as a function of the Bragg angle θ_B for two different X-rays wavelengths: λ=1.542 Å (left) and λ=0.715 Å (right).

where $\vec{r} = \vec{r}_2 - \vec{r}_1$. The quantity in square parentheses in (11.45) is the density-density autocorrelation function (2.26) introduced in § 2.6.

$$C(\vec{r}) \;=\; \langle\, \rho(\vec{r}\,')\, \rho(\vec{r}\,' + \vec{r})\, \rangle \;.$$

The measured intensity $I(\vec{K})$ is thus the *Fourier transform of the density-density autocorrelation function.*

Interaction strength

The quality and quantity of structural information available from X-ray diffraction experiments is limited, in addition to the phase problem, by the relative weakness of the interaction of X-rays with matter. For example, obtaining good information on atoms or molecules in the gas phase is difficult also with present-day synchrotron radiation sources. For gases, better results are obtained by electron diffraction, thanks the the much higher strength of the interaction (see below, § 11.5). This difficulty is to a good extent overcome when atoms or molecules are arranged in a crystal lattice. In this case, very intense X-ray diffraction is obtained for some points of the reciprocal space, corresponding to reciprocal lattice points (§ 11.6).
The high intensity and collimation of the X-ray Free Electron Lasers (XFEL), which are now becoming available for structural research, can overcome the difficulties related to the interaction weakness.

11.4 X-ray scattering from atoms

Le us now focus the attention on the scattering of X-rays from atoms. The quantum description of the electronic structure of an atom is made in terms of a continuous probability distribution $\rho(\vec{r}) = |\Psi(\vec{r})|^2$, where $\Psi(\vec{r})$ is the total electronic wave-function. We can thus formally use the same expression of the amplitude of scattering $f(\vec{K})$ introduced in (11.44) for a classical distribution of charge.

11.4.1 The X-ray atomic scattering factor

Let us consider an atom with Z electrons. The scattering amplitude $A(\vec{K})$ in electronic units

$$A_{\text{e.u.}}(\vec{K}) \;=\; \int \rho(\vec{r})\, e^{\,\mathrm{i}\vec{K}\cdot\vec{r}}\, dV \;=\; f_0(\vec{K}, Z)\,, \tag{11.46}$$

is named *atomic scattering factor* $f_0(\vec{K}, Z)$. The atomic scattering factor $f_0(\vec{K}, Z)$ is the Fourier transform of the electron density $\rho(\vec{r})$.

Note: When comparing the scattering strength of different probes (X-rays, electrons and neutrons), as will be done in § 11.5, it is convenient to consider the X-ray scattering length $f_X(\vec{K}, Z) =$

$r_e f_0(\vec{K}, Z)$.

The size of the electron orbitals of an atom is comparable with the X-rays wavelength.
The superposition of waves scattered by the different regions of the electron cloud gives rise to interference effects, which modulate the scattered amplitude as a function of both the X-rays wavelength λ and the scattering angle $\vartheta = 2\theta_B$; this effect is depicted in Fig. 11.6 for the one-electron hydrogen atom. The dependencies on λ and θ_B can be synthesised in the dependence on the scattering vector magnitude $K = 4\pi \sin\theta_B/\lambda$ (Fig. 11.7, left).

The scattering intensity is

$$I_{\text{e.u.}}(\vec{K}, Z) \;=\; |A_{\text{e.u.}}(\vec{K}, Z)|^2 \;=\; |f_0(\vec{K}, Z)|^2 \,. \tag{11.47}$$

Example: The two plots of Fig. 11.6 show the atomic scattering factor of hydrogen as a function of the Bragg angle θ_B for two different wavelengths of incident X-rays. The left plot of Fig. 11.7 shows the atomic scattering factor of hydrogen as a function of the scattering vector magnitude $K = 4\pi \sin\theta_B/\lambda$.
The maximum value of scattering vector magnitude $K_{\max} = 4\pi/\lambda$ depends on the wavelength. Decreasing the wavelength allows one to explore a larger portion of Fig. 11.7. The two cases of Fig. 11.6 are considered in Table 11.1:

Table 11.1: Maximum achievable values of scattering vector for X-rays wavelengths corresponding to the Cu and Mo Kα lines. The last column shows the energy lost by Compton scattering, eq. (11.61), for $\theta_B = 20°$.

	λ (Å)	$\hbar\omega$ (keV)	$K_{\max}$ (Å^{-1})	ΔE_{com} (eV)
Cu Kα	1.5	8.0	8.37	30.4
Mo Kα	0.7	17.4	17.95	142

Evaluation of the atomic scattering factor

The atomic scattering factor $f_0(\vec{K}, Z)$ is generally evaluated from ab-initio calculations of the density of electronic charge. In most cases, assuming a spherical symmetry is a good approximation. For *spherical symmetry*, one can show (§ 11.10.4) that

$$f_0(\vec{K}, Z) \;=\; 4\pi \int_0^\infty r^2\, \rho(r)\, \frac{\sin(Kr)}{Kr}\, dr \;=\; f_0(K, Z) \;: \tag{11.48}$$

the atomic scattering factor is a real quantity, only depending on the magnitude K of the scattering vector $\vec{K}$ (say on the radial coordinate in reciprocal space).
In the following, we consider only spherical charge distributions.

Properties of the atomic scattering factor

The properties of the atomic scattering factor, illustrated in Fig. 11.7, can be summarised as follows.

1. In the forward direction, say for $\theta_B = 0$ ($K = 0$) the interference is always completely constructive. For $K = 0$, the atomic scattering factor is equal to the atomic number Z and the intensity (in electronic unit) is equal to Z^2, as one can verify from (11.48):

$$A_{\text{e.u.}}(0) \;=\; Z\,, \qquad\qquad I_{\text{e.u.}}(0) \;=\; |A_{\text{e.u.}}(0)|^2 \;=\; Z^2\,. \tag{11.49}$$

2. When the scattering angle increases (K increases), the scattering amplitude and intensity are progressively reduced due to interference effects. The shorter the wavelength, the stronger is the effect.

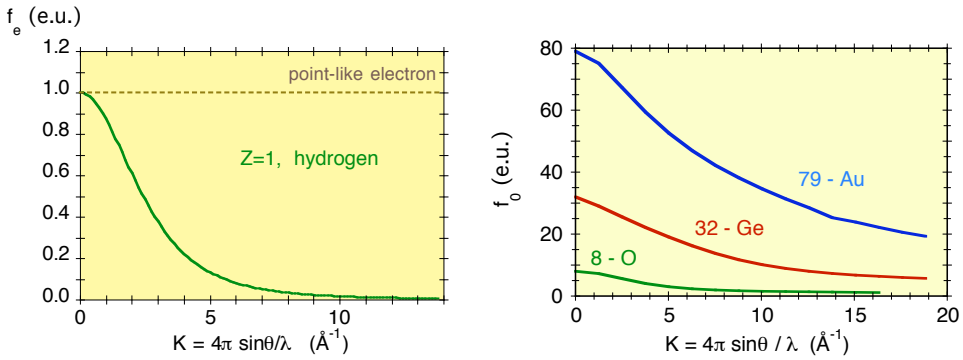

Figure 11.7: Atomic scattering factors for different elements, as a function of the scattering wavevector K.

3. The atomic scattering factor depends on the atomic number Z for any value of K, the heavier atoms giving a stronger contribution to diffraction patterns than the lighter atoms.

Problem: Evaluate the relative intensities of the contributions to X-ray diffraction patterns of the elements H, C, Ti, Zn

.

Quantum treatment of elastic scattering

Let us calculate the matrix element of (11.21) for the case of an elastic scattering, where

a) the atom is left unchanged, $|\Psi^{\rm at}_{\rm fin}\rangle = |\Psi^{\rm at}_{\rm fin}\rangle$;
b) a photon $(\vec{k}\hat{\epsilon})$ id annihilated and a photon $(\vec{k}\,'\hat{\epsilon}')$ is created; the two photons share the same energy $\hbar\omega$.

The matrix element of (11.21) can be simplified by considering only the photons involved in scattering and the term of the interaction Hamiltonian quadratic in $\mathcal{A}$, which can support the two-photon process of scattering:

$$\left\langle \Psi^{\rm at}; 1(\vec{k}\,'\hat{\epsilon}'), 0(\vec{k}\hat{\epsilon}) \;\middle|\; \frac{e^2}{2m}\,\mathcal{A}^2 \;\middle|\; 0(\vec{k}\,'\hat{\epsilon}'), 1(\vec{k}\hat{\epsilon});\, \Psi^{\rm at} \right\rangle \tag{11.50}$$

Let us now reduce the expression (11.19) of the vector potential operator to the case of two photons and one scattering electron:

$$\vec{\mathcal{A}} \;=\; \sqrt{\frac{\hbar}{2\epsilon_0 V\omega}}\,\left[\hat{\epsilon}\,a_{\vec{k}}e^{-i\vec{k}\cdot\vec{r}} + \hat{\epsilon}\,a^\dagger_{\vec{k}}e^{i\vec{k}\cdot\vec{r}} \;+\; \hat{\epsilon}'\,a_{\vec{k}\,'}e^{-i\vec{k}\,'\cdot\vec{r}} + \hat{\epsilon}'\,a^\dagger_{\vec{k}\,'}e^{i\vec{k}\cdot\vec{r}}\right]. \tag{11.51}$$

The square of $\mathcal{A}$ contains 16 terms, of which only the terms corresponding to the annihilation of the $(\vec{k}\hat{\epsilon})$ photon and to the creation of the $(\vec{k}\,'\hat{\epsilon}')$ photon should be retained. The corresponding matrix element is

$$\frac{\hbar}{2\epsilon_0 V\omega}\,\frac{e^2}{2m}\,(\hat{\epsilon}\cdot\hat{\epsilon}')\left\langle \Psi^{\rm at}; 1(\vec{k}\,'\hat{\epsilon}'), 0(\vec{k}\hat{\epsilon}) \;\middle|\; a^\dagger_{\vec{k}\,'}a_{\vec{k}}e^{i(\vec{k}\,'-\vec{k})\cdot\vec{r}} + a_{\vec{k}}a^\dagger_{\vec{k}\,'}e^{i(\vec{k}\,'-\vec{k})\cdot\vec{r}} \;\middle|\; \Psi^{\rm at}; 0(\vec{k}\,'\hat{\epsilon}'), 1(\vec{k}\hat{\epsilon})\right\rangle \tag{11.52}$$

The action of the annihilation and creation operators transform the initial state of the electromagnetic field into the final state, and we are left with

$$\frac{\hbar}{2\epsilon_0 V\omega}\,\frac{e^2}{2m}\,(\hat{\epsilon}\cdot\hat{\epsilon}')\;2\;\left\langle \Psi^{\rm at} \;\middle|\; e^{i\vec{K}\cdot\vec{r}} \;\middle|\; \Psi^{\rm at}\right\rangle \tag{11.53}$$

where $\vec{K} = \vec{k}\,' - \vec{k}$ is the scattering vector.

In the coordinate representation, the remaining matrix element corresponds to the atomic scattering factor

$$\left\langle \Psi^{\mathrm{at}} \,\middle|\, e^{i\vec{K}\cdot\vec{r}} \,\middle|\, \Psi^{\mathrm{at}} \right\rangle \;=\; \int |\Psi^{\mathrm{at}}|^2\, e^{i\vec{K}\cdot\vec{r}}\, dV \;=\; f_0(\vec{K}, Z)\,. \tag{11.54}$$

It is easy to verify that the polarisation term $\hat{\epsilon}\cdot\hat{\epsilon}'$ is equal to 1 for π polarisation and to $\cos(2\theta_B)$ for σ polarisation, in agreement with the classical approach.
The probability of transition per unit time from the initial to the final states (11.21)

$$w_{fi} \;=\; \frac{2\pi}{\hbar}\, |\langle \Psi_f | H_I | \Psi_i \rangle|^2\, g(E_f) \tag{11.55}$$

is proportional to the density of final states available to the scattered photon within the angular acceptance $\Delta\Omega$ of the detector. The number of available states for a photon of given polarisation $\hat{\epsilon}'$ is

$$dN \;=\; \frac{V}{8\pi^3}\, k^2\, dk\, \Delta\Omega \;=\; g(E_f)\, dE_f\,, \tag{11.56}$$

whence

$$g(E_f) \;=\; \frac{V}{8\pi^3}\, \frac{\omega^2}{\hbar c^3}\, \Delta\Omega \tag{11.57}$$

so that

$$w_{fi} \;=\; \frac{c}{V} \left(\frac{e^2}{4\pi\epsilon_0 c^2 m} \right)^2 (\hat{\epsilon}\cdot\hat{\epsilon}')^2\, |f_0(\vec{K}, Z)|^2\, \Delta\Omega\,. \tag{11.58}$$

In the cubic box normalisation condition, the one-photon flux is $\Phi_{\mathrm{in}} = c/V$, so that the intrinsic cross section for elastic scattering from an atom is

$$\left(\frac{d\sigma}{d\Omega} \right)_0 \;=\; \frac{w_{fi}}{\Phi_{\mathrm{in}}}\, \frac{1}{\Delta\Omega} \;=\; r_e^2\, (\hat{\epsilon}\cdot\hat{\epsilon}')^2\, |f_0(\vec{K}, Z)|^2\,, \tag{11.59}$$

in agreement with the classical result.

11.4.2 Limits of the Thomson theory for X-ray scattering

Once the distribution of electronic charge is allowed for, the classical Thomson theory is valid, unless:

1. electrons are so weakly bound that X-rays can undergo inelastic Compton scattering,
2. the X-ray energy is in the vicinity of the binding energy of some electrons, and X-rays undergo resonant scattering.

In the following, we consider only the first point (Compton effect).

Compton scattering

A perfectly free electron hit by a photon gives rise to Compton scattering. From the equation for energy and linear momentum conservation of the system (photon+electron), one can find the variation of the photon wavelength

$$\Delta\lambda \;=\; \lambda_{\mathrm{out}} - \lambda_{\mathrm{in}} \;=\; \frac{h}{m_e c}\, (1 - \cos 2\theta_B) \;=\; \lambda_c\, (1 - \cos 2\theta_B)\,, \tag{11.60}$$

where $\lambda_c = 0.02426$ Å is the (constant) Compton wavelength.

- In the forward direction, say for $\theta_B = 0$, there is no wavelength modification, $\Delta\lambda = 0$: the scattering is elastic (unmodified Thomson scattering).
- For $\theta_B \neq 0$, the wavelength is modified, $\Delta\lambda > 0$; the scattering is inelastic (modified Comtpon scattering); the shift $\Delta\lambda$ depends on the scattering angle but does not depend on λ_{in}.

The energy lost by a photon undergoing Compton scattering,

$$\Delta E_{\mathrm{com}} \;=\; \hbar\omega_{\mathrm{in}} - \hbar\omega_{\mathrm{out}} \simeq\; \hbar\omega_{\mathrm{in}}\, \frac{\Delta\lambda}{\lambda_{\mathrm{in}}}, \tag{11.61}$$

is gained by the recoiling electron, and increases when the scattering angle increases.

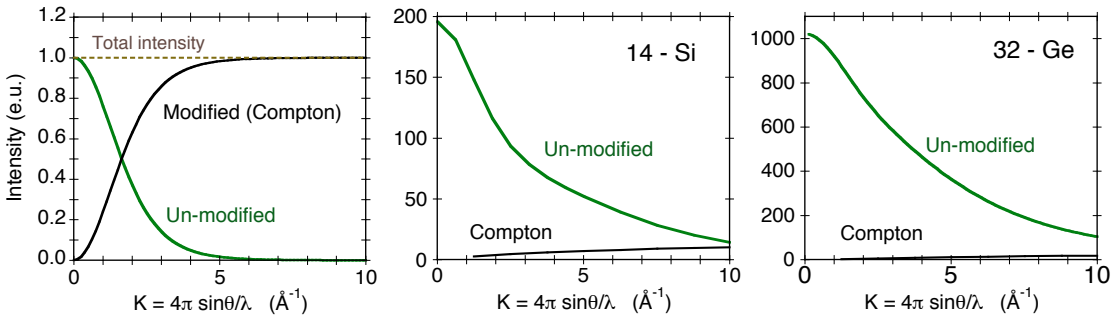

Figure 11.8: Contributions of Thomson (unmodified, elastic) and Compton (modified, inelastic) scattering to the scattering intensities for Hydrogen (left), Silicon (centre) and Germanium (right).

Modified versus unmodified scattering

The statistical balance between modified (Compton) and unmodified (Thomson) scattering from an atomic electron is given by the relation between the binding energy E_b of the electron and $\Delta E_{\rm com}$:

- if $\Delta E_{\rm com} > E_b$, modified Compton scattering prevails,
- if $\Delta E_{\rm com} < E_b$, unmodified Thomson scattering prevails.

Example: With reference to the two plots of Fig. 11.6, let us evaluate the energy lost by Compton scattering $\Delta E_{\rm com}$ for the Bragg angle $\theta_B = 20°$ (scattering angle $2\theta_B = 40°$), for the two X-rays wavelengths corresponding to the Cu and Mo Kα lines. In both cases, $\Delta\lambda = 0.0057$ Å. The values of $\Delta E_{\rm com}$ are shown in the last column of Table 11.1 at page 262. They should be compared with the binding energy $E_{\rm b} = 13.6$ eV. From the two plots of Fig. 11.6 one can evaluate the relative weights of unmodified and modified scattering.

A quantitative evaluation of the relative weights of modified and unmodified scattering can be obtained by refined quantum mechanical calculations.
One finds that for one electron (see Fig. 11.8, left):

a) the total scattered intensity in electronic units (modified plus unmodified) is $I_{\rm tot} = 1$

$$I_{\rm tot}(K) \;=\; I_{\rm mod}(K) + I_{\rm unmod}(K) \;=\; \left[1 - f_0^2(K)\right] + \; f_0^2(K) \;=\; 1\,; \qquad (11.62)$$

b) the modified (Compton) intensity is the complement to 1 of the unmodified intensity, $I_{e,\rm mod} = 1 - f_0^2$; it is zero for $\theta = 0$, and progressively increases when K increases.

Let us now consider an atom containing Z electrons (in Fig. 11.8, centre and right). The contributions of the different electrons have to be summed up, but different procedures have to be used for modified and unmodified scattering.

a) *Unmodified scattering.* The unmodified scattering from the different electrons is coherent, and gives rise to interference (there is no way of determining which one of the atomic electrons scatters an X-ray photon).
The scattering amplitudes of different electrons are first summed up, to give the atomic scattering factor $f_0(K, Z)$. The intensity is the squared modulus of the sum of amplitudes. The maximum value of the total intensity is Z^2 for $K = 0$ (see eq. 11.43 with $K = 0$).

b) *Modified Compton scattering.* The scattering contributions of different electrons are incoherent (in principle, it is possible to determine which one of the atomic electrons scatters an X-ray photon and is ejected from the atom).

The intensities (not the amplitudes) have to be summed up. The maximum value of the total intensity for modified scattering is Z for $K \to \infty$. The modified intensity can be calculated from the knowledge of the scattering factors for unmodified scattering from the different electrons:

$$I_{\mathrm{mod}}(K) \;=\; Z \;-\; \sum_{i=1}^{Z} |f_i(\vec{K})|^2 \tag{11.63}$$

As a consequence, when the atomic number Z increases, the Thomson effect becomes progressively predominant over the Compton effect. The Compton effect can be relevant only at high K values.

11.5 X-rays, electrons and thermal neutrons

The approach depicted above for the scattering of X-rays from atoms can be generalised to the scattering of neutrons and electrons. The interaction mechanisms are different for the three probes, and the intrinsic cross sections are different accordingly.

11.5.1 Intrinsic scattering cross section for X-rays

The interaction of X-rays with matter can be treated at different levels of approximation (classical, semi-classical, quantistic). One always finds that the intrinsic cross section for the interaction of an unpolarised X-ray beam with a point-like electron is

$$\left(\frac{d\sigma}{d\Omega}\right)_0 \;=\; r_e^2 \;\left(\frac{1+\cos^2(2\theta_B)}{2}\right), \tag{11.64}$$

where $r_e = 2.8 \times 10^{-5}$ Å is the Thomson scattering length. The intrinsic Thomson scattering cross section from one point-like electron doesn't depend on the X-ray wavelength, and depends on the scattering angle $\vartheta = 2\theta_B$ only through the polarisation factor.

The intrinsic X-rays cross section from an atom (with spherical symmetry) is

$$\left(\frac{d\sigma}{d\Omega}\right)_0 \;=\; r_e^2 \;|f_0(K,Z)|^2 \;\left(\frac{1+\cos^2(2\theta_B)}{2}\right), \tag{11.65}$$

where $f_0(K,Z)$ is the atomic scattering factor for X-rays.

For a meaningful comparison with electrons and neutrons, it is convenient to introduce the *atomic scattering length* for X-rays

$$f_X(K,Z) \;=\; r_e f_0(K,Z)\,. \tag{11.66}$$

The X-ray atomic scattering length for germanium is shown in the lower panel of Fig. 11.9.

11.5.2 Intrinsic cross section for electrons

Electrons are scattered by the Coulomb potential generated by both the positive nuclear charge and the negative electron charge of the atoms. According to (11.10), the intrinsic cross section for elastic scattering of electrons from atoms is the squared modulus of the scattering amplitude:

$$\left(\frac{d\sigma}{d\Omega}\right)_{0,\mathrm{el}} \;=\; |f_{\mathrm{el}}(\vec{K},Z)|^2\,. \tag{11.67}$$

Within the Born approximation for scattering from a central potential $\Phi(r)$, the scattering amplitude is

$$f_{\mathrm{el}}(\vec{K},Z) \;\propto\; \int \Phi(r)\, e^{i\vec{K}\cdot\vec{r}}\, dV\,. \tag{11.68}$$

The formal similarity with (11.46) is evident: the atomic scattering amplitude for electrons is the Fourier transform of the interaction potential, as the atomic scattering amplitude for X-rays is the Fourier transform of the electron density.

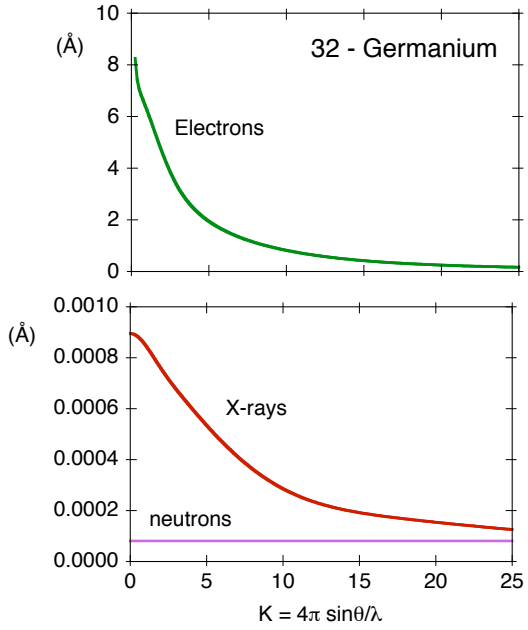

Figure 11.9: Scattering lengths for the interaction of electrons, X-rays and neutrons with a Ge atom as a function of the scattering vector magnitude $K = 4\pi \sin\theta_B/\lambda$. The scattering lengths for electrons and X-rays decrease when K increases. The scattering length for neutrons is instead constant. The vertical scale allows to appreciate the relative strengths of the interaction mechanisms.

In the following we consider only spherically symmetric potentials $\Phi(r)$ (see also § 11.10.5). The electrostatic potential $\Phi(r)$ is connected to the charge density $\rho(r)$ by the Poisson equation

$$\nabla^2\Phi(r) \;=\; -\frac{1}{\epsilon_0}\,[\,\rho_+(r) - \rho_-(r)\,]\;. \tag{11.69}$$

In order to find a convenient expression for the atomic scattering amplitude for electrons, let us re-express the three terms of (11.69) as Fourier back-transforms.

1. Inverting (11.68), the potential $\Phi(r)$ can be expressed as the Fourier back-transform of the electron scattering amplitude $f_{\rm el}(K,Z)$; the laplacian is

$$\nabla^2\Phi(r) \;\propto -\int K^2\, f_{\rm el}(K,Z)\, e^{-iKr}\, dK\;. \tag{11.70}$$

2. The positive nuclear charge can be expressed as

$$\rho_+(r) \;=\; Z\,\delta(r) \;\propto\; \int Z\, e^{-iKr}\, dK \tag{11.71}$$

3. The negative electron charge can be expressed as the Fourier back-transform of the X-ray scattering factor $f_0(K,Z)$

$$\rho_-(r) \;\propto\; \int f_0(K,Z)\, e^{-iKr}\, dK \tag{11.72}$$

By inserting (11.70), (11.71) and (11.72) into the Poisson equation (11.69) and considering the arguments of the integrals, one finds the Mott-Bethe formula for the electron scattering amplitude:

$$f_{\rm el}(K,Z) \;=\; \frac{me^2}{2\pi\hbar^2\epsilon_0}\,\frac{Z - f_0(K,Z)}{K^2}\;. \tag{11.73}$$

In Fig. 11.9 the scattering amplitude (or scattering length) for electrons is compared with the scattering length for X-rays; the scattering length is much stronger and the intensity is more forward peaked for electrons than for X-rays.

11.5.3 Intrinsic cross section for thermal neutrons

Thermal neutrons are scattered by matter via two mechanisms, of approximately comparable strength.

1. Thermal neutrons are scattered by atomic nuclei. Contrary to the case of the Thomson theory for X-rays, there is no approximate classical model that accounts for nuclear scattering, and the quantum treatment of the strong nuclear interaction can be only very approximate.
   The strong nuclear interaction responsible for neutron scattering has a very short range, of the order of a few 10^{-5} Å, much shorter than the wavelength of thermal neutrons. As a consequence, neutron scattering from a single nucleus is isotropic. The situation is reminiscent of X-ray scattering by point-like electrons.
2. The neutron spin can interact with the unpaired electrons of magnetic atoms. This kind of interaction is possible only for systems where magnetic atoms are present, and gives information on the magnetic structure of matter.
   Since the magnetic interaction involves the electron cloud, magnetic neutron scattering is anisotropic like X-ray scattering from atoms.

Let us consider here only the elastic scattering of thermal neutrons from the nuclei of non-magnetic materials (case 1).
The interaction of thermal neutrons with the nuclei is described by the Fermi pseudo-potential

$$V(\vec{r}) \;\propto\; b\,\delta(\vec{r})\,, \tag{11.74}$$

where b is a constant. Making again use of the Born approximation, one finds that the scattering length for thermal neutrons is

$$f_{\mathrm{n}} \;=\; b\,. \tag{11.75}$$

The strength of the interaction between neutrons and the different nuclei is measured by the constant scattering length b, which is independent of the scattering vector $\vec{K}$.
Contrary to the X-ray and electron scattering lengths, the neutron scattering length b has no regular dependence on the atomic number Z. Besides, the scattering length b can be dramatically different for different isotopes of the same atomic species (Table 11.2).

Since different isotopes of the same atomic species are randomly distributed within the material, the intensity of coherent neutron scattering from an aggregate of atoms is proportional to the average value of the square of the scattering length

$$\langle b_i b_j \rangle \;=\; \langle b^2 \rangle \;=\; b^2_{\mathrm{coh}}\,, \tag{11.76}$$

where i and j label two atomic sites (including the case $i = j$), so that

$$\left(\frac{d\sigma}{d\Omega}\right)_{0,n} \;=\; b^2_{\mathrm{coh}}\,. \tag{11.77}$$

The scattering length of neutrons is compared in Fig. 11.9 with the scattering lengths of electrons and X-rays for the case of germanium.

11.5.4 Comparisons

1. The diffraction of X-rays and electrons can give information on the electron density inside matter, neutron diffraction is sensitive only to the nuclear positions.
   If the calculated electronic densities of single atoms are subtracted from the electronic density of an atomic aggregate determined by X-ray diffraction, one can evaluate the electronic densities in the regions between atoms, and get information on the nature of the chemical bonds.

Table 11.2: Some examples of neutron scattering lengths. Positive value correspond to a π phase change on scattering, negative values to no phase change.

Isotope	Z	A	b (10^{-15} m)	**Isotope**	Z	A	b (10^{-15} m)
Hydrogen	1	1	-3.74	^{28}Si	14	28	+4.11
Deuterium	1	2	+6.67	^{70}Ge	32	70	+8.4
Tritium	1	3	+4.94	^{139}La	57	139	+8.2
				^{197}Au	79	197	+7.63

2. The interaction of electrons with atoms is much stronger than the interaction of X-rays and neutrons (see the vertical scales in Fig. 11.9). Considerations on the interaction strength:
   a) The weaker the interaction, the higher is the penetrating power. Electrons have a small penetrating power; neutrons are more penetrating than X-rays. This property reflects on the size of the samples that can be studied. To increase the penetrating power of electrons, high energy (small wavelength) is used.
   b) Diffraction can be interpreted by the kinematic theory (each beam is scattered only once inside matter) only if the interaction is sufficiently weak with respect to the sample size. For X-rays, the kinematic theory is suitable for powders and mosaic crystals, where the size of the coherent scattering regions are of the order of the micrometer. For electrons of 50-100 keV, the kinematical theory is suitable for samples of 100-300 Å size.

3. X-ray diffraction is strongly and monotonically dependent on the atomic number Z. Electron diffraction is less strongly dependent on Z. Neutron diffraction randomly depends on atomic and mass number of isotopes.

4. Different types of interaction are generally present, in addition to elastic scattering, such as inelastic scattering and absorption. Photoelectric absorption is particularly strong in the case of X-rays, and gives the most important contribution to the beam attenuation.

When a unified treatment of scattering from atomic aggregates is sought, it is convenient to express the strength of the interaction with an atom in terms of an atomic factor $f_a(K)$ (whose dimension is length) for all probes. For a given atomic species, the atomic factor is different for different probes.
For X-rays

$$f_a(K) \;=\; f_X \;=\; r_e\, f_0(K) \tag{11.78}$$

For electrons

$$f_a(K) \;=\; f_{\rm el} \;=\; \frac{me^2}{2\pi\hbar^2\epsilon_0}\,\frac{Z - f_0(K)}{K^2} \tag{11.79}$$

For neutrons

$$f_a \;=\; f_{\rm n} \;=\; b_{\rm coh} \tag{11.80}$$

Once the peculiarities of a given probe are taken into account by the atomic factor f_a, one can calculate the diffraction effects due to scattering from aggregates of atoms by the same procedure for the different probes (see below).

11.5.5 Temperature effects

The amplitude of atomic vibrations is finite even at zero temperature (zero point energy) and increases when temperature increases. Atomic vibrations are accounted for by a probability density $w(\vec{r})$ of finding the atomic nucleus at the position $\vec{r}$ with respect to the equilibrium position.

Within the adiabatic approximation, the displacement of a nucleus is accompanied by the displacement of the electronic cloud without appreciable delay: for each instantaneous position the nucleus is surrounded by the same electron distribution.
The electron density $\rho(\vec{r})$ of a static atom has to be replaced by the average electron cloud of a vibrating atom $\rho_T(\vec{r})$, which is the convolution of $\rho(\vec{r})$ and $w(\vec{r})$:

$$\rho_T(\vec{r}) \;=\; \rho(\vec{r}) * w(\vec{r}) \;=\; \int \rho(\vec{r}-\vec{r}\,')\, w(\vec{r}\,')\, dV_{\vec{r}\,'}\,. \tag{11.81}$$

Even the point-like picture of static nuclei is substituted by the probability density $w(\vec{r})$.

Let us now consider the situation in the reciprocal space.
The Fourier transform of the probability density $w(\vec{r})$ is the thermal factor

$$f_T(\vec{K}) \;=\; \int w(\vec{r})\, \exp(i\vec{K}\cdot\vec{r})\, dV\,. \tag{11.82}$$

According to the convolution theorem (see § 2.4), the atomic factors f_a (11.78)–(11.80) are substituted by atomic thermal factors f_{aT}, product of the atomic factors f_a and the thermal factors f_T:

$$f_{aT}(\vec{K}) \;=\; f_a(\vec{K})\, f_T(\vec{K})\,. \tag{11.83}$$

The enlargement of the electron cloud due to vibrational motion influences the scattering of X-rays and of electrons. For neutrons, the constant scattering length b is substituted by $f_{aT}(\vec{K}) = b\, f_T(\vec{K})$.

The vibrational atomic displacement is conventionally named $\vec{u}$.
For isotropic harmonic vibrations, the probability density $w(u)$ has the simple gaussian form

$$w(u) \;=\; \frac{1}{[2\pi\langle u^2\rangle]^{3/2}}\, \exp\left[-\frac{u^2}{2\langle u^2\rangle}\right]\,, \tag{11.84}$$

where $\langle u^2\rangle$ is the variance of the distribution of atomic displacements. Correspondingly, the thermal factor f_T only depends on the magnitude K of the scattering vector and has a gaussian shape:

$$f_T(K) \;=\; \exp[-K^2\,\langle u^2\rangle/2]\,. \tag{11.85}$$

Problem: The atomic factors for X-rays, electrons and neutrons are shown in Fig. 11.9 for the case of germanium. Depict qualitatively the modification induced by the thermal factor.

11.6 Elastic scattering from atomic aggregates

Let us now consider the scattering from atomic aggregates. For concreteness, we follow the formalism for X-ray scattering, based on the amplitude expressed in atomic units. The adaptation to the other probes is trivial.

To calculate the amplitude of X-ray scattering from a set of N atoms, we use the same formal expression (11.41), where single electrons are now substituted by atoms; the electronic structure of atoms and the effect of thermal vibrations are globally taken into account by the atomic thermal factor $f_m = f_0\, f_T$ in electronic units. The scattering amplitude is

$$A_{\mathrm{e.u.}}(\vec{K}) \;=\; \sum_{m=1}^{N} f_m(\vec{K})\, e^{\mathrm{i}\vec{K}\cdot\vec{R}_m} \tag{11.86}$$

where m labels the atoms and $\vec{R}_m$ are the equilibrium positions.
If the system is monatomic, all the atomic scattering factors are equal. If not, different atomic scattering factors have to be considered for different atomic species. Heavier atoms give a stronger contribution to the X-ray scattering amplitude.

The scattered intensity is

$$\boxed{I_{\text{e.u.}}(\vec{K}) \;=\; |A_{\text{e.u.}}(\vec{K})|^2 \;=\; \sum_{m=1}^{N}\sum_{n=1}^{N} f_m(\vec{K})\, f_n^*(\vec{K})\, e^{\mathrm{i}\vec{K}\cdot\vec{R}_{mn}}} \tag{11.87}$$

where $\vec{R}_{mn} = \vec{R}_m - \vec{R}_n$ is the distance between atom m and atom n (note that also the terms $m = n$ are included).

Eq. (11.87) has *general validity*, for any aggregate of atoms.

Example: Planar distribution of N atoms of the same species.

For atoms of the same species all scattering factors are equal, $f(\vec{K})$. In analogy with (11.43) for point-like electron scattering, eq. (11.87) for atoms can be elaborated as

$$I_{\text{e.u.}}(\vec{K}) \;=\; N\,|f(\vec{K})|^2 \;+\; |f(\vec{K})|^2 \sum_{m}\sum_{n\neq m} \cos(\vec{K}\cdot\vec{R}_{mn})\,, \tag{11.88}$$

where the first term on the right-hand side is the effect of independent scattering from N atoms, the sum of the remaining $N(N-1)$ terms is the effect of interference.

If the N atoms are arranged in a plane (two dimensional sheet), the interference contribution is maximum when $\vec{K}$ is perpendicular to the plane (specular reflection).

Different types of atomic aggregates

Depending on the different types of atomic aggregate, (11.87), different approaches have been developed:

1. An approach useful for samples made by a large number of randomly oriented systems of relatively small size: molecular gases, nanoclusters, sometimes also crystalline powders (§ 11.6.1). The relative atomic positions can be exactly defined within a single system (molecule, cluster, crystallite), but each system is randomly oriented with respect to the others. This approach leads to the so-called Debye formula.

2. An approach useful for liquids and-non crystalline systems (§ 11.6.2), where the structure is statistically described in terms of a one-dimensional pair distribution function (§ 2.6).

3. An approach specifically devised for crystalline systems, whose structure can be considered as the convolution of a lattice function and a basis function (§ 11.6.3). This approach leads to the well known Laue (or Bragg or Ewald) interference conditions and to the structure factor.

11.6.1 Elastic scattering from molecules

Equation (11.87) can be developed by separating the terms $m = n$ from the terms $m \neq n$:

$$I_{\text{e.u.}}(\vec{K}) \;=\; \sum_{m} f_m^2(K) \;+\; \sum_{m}\sum_{n\neq m} f_m(K)\, f_n(K)\, e^{i\vec{K}\cdot\vec{R}_{mn}}\,. \tag{11.89}$$

(We consider here spherically symmetric atoms with real scattering factors, which only depend on the magnitude K).

The sums in (11.89) are over the atoms of a single molecule.

1. The first sum on the right hand side of (11.89) is the independent contribution of the N atoms of the molecule. For a molecule containing only one atomic species, at $K = 0$ the first sum is NZ^2. This contribution decreases monotonically when K increases, according to the behaviour of the squared scattering factor $f_m^2(K)$.
2. The second sum over the $N(N-1)$ pairs of atoms gives rise to interference patterns, that appear as oscillations around the smooth independent contribution. The interference patterns contain the structural information (basically, information on inter-atomic distances R_{mn}).

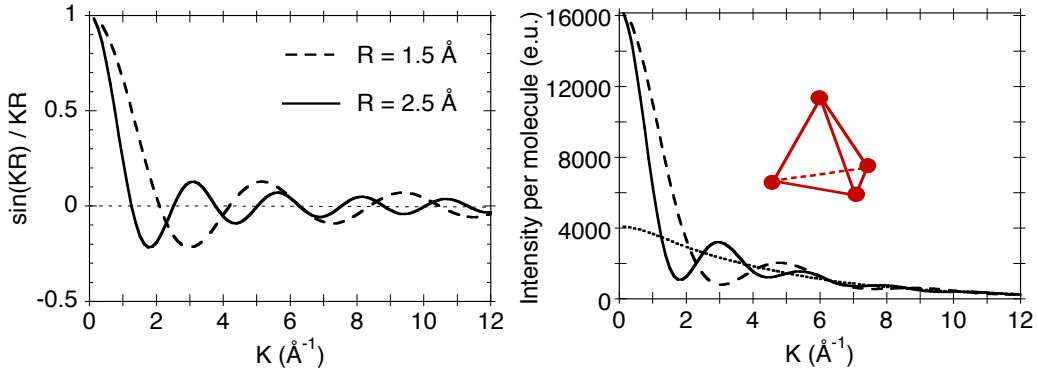

Figure 11.10: Left: function $\sin(KR)/KR$ for two different values of R. Right: elastic scattering intensity calculated according to (11.93) for an hypothetical tetrahedral molecule of Ge atoms. The continuous line is for an interatomic distance R=2.5 Å, the dashed line is for R=1.5 Å. The dotted line gives the independent terms contribution, equal to $4|f(\vec{K})|^2$.

When dealing with a randomly oriented molecule (or a nano-cluster or a crystallite), the expression (11.89) can be simplified by averaging over all possible orientations of the interatomic distances $\vec{R}_{mn}$ and their projections on the scattering vector:

$$I_{\text{e.u.}}(K) \;=\; \sum_m f_m^2(K) \;+\; \sum_m \sum_{n\neq m} f_m(K)\, f_n(K) \left\langle e^{\mathrm{i}\vec{K}\cdot\vec{R}_{mn}} \right\rangle . \tag{11.90}$$

If one assumes spherical symmetry of the orientation of the vectors $\vec{R}_{mn}$, one gets (§ 11.10.6)

$$\left\langle e^{\mathrm{i}\vec{K}\cdot\vec{R}_{mn}} \right\rangle \;=\; \frac{\sin(KR_{mn})}{KR_{mn}} , \tag{11.91}$$

so that one can arrive to the *Debye formula* for one molecule

$$I_{\text{e.u.}}(K) \;=\; \sum_m f_m^2(K) \;+\; \sum_m \sum_{n\neq m} f_m(K)\, f_n(K)\, \frac{\sin(KR_{mn})}{KR_{mn}} . \tag{11.92}$$

The Debye formula can be applied to any randomly oriented system.

Eq. (11.92) gives the scattering intensity once the structure is known. The actual interest is in the inverse process, say in the determination of the structure from the intensity of scattering patterns $I_{\text{e.u.}}(K)$. To that purpose, one can guess a structural model, calculate the expected intensities through (11.92), compare with the experimental intensities and progressively refine the model until the best agreement is obtained. For large systems, such as a crystallite in a crystalline powder, such a procedure is still quite prohibitive in terms of computer time, because of the huge number of terms to be summed up.

Note that the Debye formula only gives one-dimensional information (on the magnitude, not on the direction, of distances) and directly connects the structural information to the scattering intensity, not to the amplitude. No three-dimensional information can be obtained.

Problem: Plot the function $\sin(KR)/KR$ against KR and verify its damped oscillating behaviour. The function is zero for $KR = n\pi$. Plot $\sin(KR)/KR$ against K for different values of R and verify that the frequency of the damped oscillations increases when R increases (see Fig. 11.10, left).

Debye formula for a monatomic molecule

If the N atoms of the molecule are equal, the Debye formula (11.92) becomes

$$I_{\text{e.u.}}(K) \;=\; f^2(K) \left[N + \sum_m \sum_{n \neq m} \frac{\sin(K R_{mn})}{K R_{mn}} \right] . \tag{11.93}$$

The right panel of Fig. 11.10 refers to the hypothetical case of a randomly oriented tetrahedral Ge_4 molecule, where only one value of interatomic distance is present. The calculated contributions to elastic scattering are shown: the frequency of the damped oscillations is proportional to the interatomic distance.

Further considerations

1. The cross section for elastic scattering form one molecule is $r_e^2\, I_{\text{e.u.}}(K)$, where $r_e = 1.6 \times 10^{-5}$ Å^2. In an experimental spectrum, the inelastic Compton contribution is added to the elastic contributions; the Compton contribution increases with the scattering vector K. The relative importance of Compton scattering is stronger for lighter elements (see Fig. 11.8). The calculated Compton contribution has to be subtracted from the experimental signal before analysing it in terms of elastic scattering.

2. The total measured intensity is the sum of the intensities scattered by the single molecules (scattering by different molecules is incoherent). The intensity for one molecule, given by (11.93), is multiplied by the number $\mathcal{N}$ of molecules encountered by the X-ray beam.

3. For most molecules, a number of different distances are present and the interference function is a superposition of oscillating functions of different frequencies.

4. The accuracy of structural determinations is increased by increasing the extent of experimental information, e.g. by increasing the range of K values. The maximum value of the scattering vector $K_{\max} = 4\pi/\lambda$ can be increased only by reducing the wavelength. Synchrotron Radiation sources, thanks to high intensity and the continuum spectrum, allow the choice of shorter wavelengths with respect to laboratory sources. When K increases, however, the scattering length $f_{aT}(K)$ decreases. In many cases it can be more convenient to resort to thermal neutrons, whose scattering length only depends on K through the thermal factor.

11.6.2 Elastic scattering from non-crystalline systems

The application of the Debye formula to liquids or amorphous solids is impracticable: the structure of non-crystalline condensed phases cannot be specified in terms of all the atomic positions. As it was shown in § 2.6, the structure of non-crystalline condensed phases is statistically characterised in terms of a radial distribution function $R(r)$ or, equivalently, of a pair distribution function $g(r)$ (Fig. 2.16)

$$R(r) \;=\; 4\pi r^2\, \rho_0\, g(r)\,, \tag{11.94}$$

where ρ_0 is the average atomic number density. The information contained in $R(r)$ and $g(r)$ is purely unidimensional.

In order to establish a viable connection between the results of X-ray scattering experiments and the structure described by the pair distribution function, a suitable formalism has been developed. For simplicity, we consider here only monatomic systems and spherically symmetric atoms.

Equation (11.89) for molecules is modified by considering only one atomic species and by substituting the last sum $\sum_{m \neq n}$ with an integral over the distribution $\rho_m(\vec{r}_{mn})$ of the interatomic

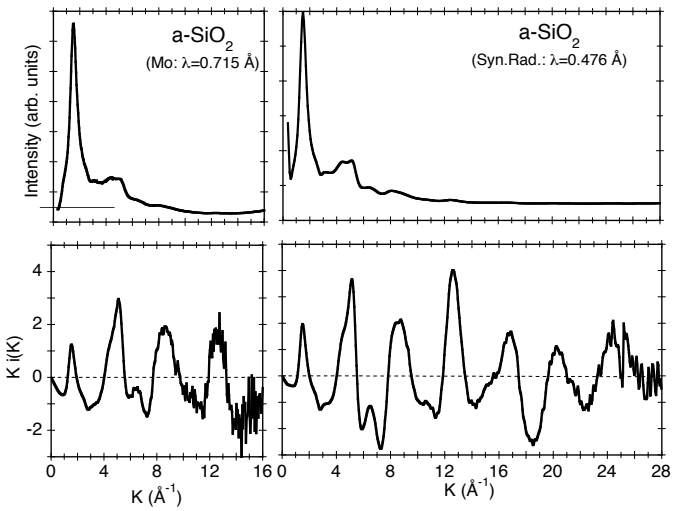

Figure 11.11: X-ray scattering from non-crystalline silica (a-SiO_2). Top panels: experimental spectra; bottom panels: interference functions $K\,i(K)$. The left and right panels refer to measurements performed with different wavelengths and different resolutions (laboratory Mo tube and Synchrotron Radiation, respectively).

distances from a given central atom m to all other atoms:

$$I_{\text{e.u.}}(\vec{K}) \;=\; N\,f^2(K) \;+\; f^2(K)\sum_m \int_V \rho_m(\vec{r}_{mn})\,e^{i\vec{K}\cdot\vec{r}_{mn}}\,dV_n\,. \tag{11.95}$$

The sought structural information is represented by the oscillations of the distribution $\rho_m(\vec{r}_{mn})$ around the average value ρ_0, so it is convenient to evidence these oscillations by splitting the integral in (11.95) as the sum of two integrals

$$I_{\text{e.u.}}(\vec{K}) \;=\; N\,f^2(K) \;+\; f^2(K)\sum_m \int_V [\rho_m(\vec{r}_{mn})-\rho_0]\,e^{i\vec{K}\cdot\vec{r}_{mn}}\,dV_n \tag{11.96}$$

$$+\quad f^2(K)\sum_m \int_V \rho_0\,e^{i\vec{K}\cdot\vec{r}_{mn}}\,dV_n\,. \tag{11.97}$$

a) Only the first two terms on the right-hand side of (11.97) are considered when studying the structure of non-crystalline systems at the atomic level. The first term is the contribution of independent scattering from the N atoms of the system. The second term contains the Fourier transform of the oscillations of $\rho_m(\vec{r}_{mn})$ around the average value ρ_0, say the Fourier transform of the sought structural information. We expect that the oscillations of $\rho_m(\vec{r}_{mn})$ damp out rapidly when r_{mn} increases.

b) The third term is the Fourier transform of the average density ρ_0, which is a constant quantity, at least at the length scale of the fluctuations of $\rho_m(\vec{r}_{mn})$. The third term gives rise to a narrow peak near the origin, typically for $K \le 0.02\text{Å}^{-1}$. The analysis of the possible structures superimposed on this peak gives information on the long-range fluctuations of the average density ρ_0, typically over lengths of the order of 100 Å (the related technique is called SAXS, small angle X-ray scattering technique).

Let us focus on the first two terms on the right-hand side of (11.97). After averaging over the central atom m and considering spherical symmetry, one gets:

$$I_{\text{e.u.}}(K) \;=\; N\,f^2(K) \;+\; N\,f^2(K)\int_0^\infty [\rho(r)-\rho_0]\,4\pi r^2\,\frac{\sin(Kr)}{Kr}\,dr\,. \tag{11.98}$$

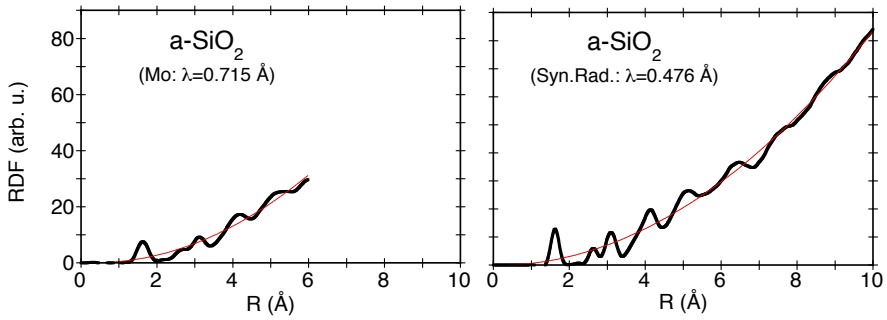

Figure 11.12: Radial distribution functions of a-SiO_2 reconstructed from measurements with laboratory Mo tube (left) and Synchrotron Radiation (right). The longer K range of Synchrotron Radiation allows a better resolution of the structures in the RDF. The better resolution and better signal to noise ratio of Synchrotron Radiation allows a longer r range in the RDF.

Experimental signals measured for SiO_2 with a laboratory source and with a Synchrotron Radiation source are shown in the top panels of Fig. 11.11.
From the experimental signal, once the Compton contribution has been subtracted, one can directly obtain the static scattering function (or static structure function) $S(K)$ defined in (11.8):

$$S(K) \;=\; \frac{I_{\text{e.u.}}(K)}{N\,f^2(K)} \;=\; 1 \;+\; \int_0^\infty [\rho(r)-\rho_0]\,4\pi r^2\,\frac{\sin(Kr)}{Kr}\,dr\,. \tag{11.99}$$

Alternatively, one considers the interference function

$$i(K) \;=\; \frac{I_{\text{e.u.}}/N - f^2(K)}{f^2(K)} \;=\; \int_0^\infty [\rho(r)-\rho_0]\,4\pi r^2\,\frac{\sin(Kr)}{Kr}\,dr\,. \tag{11.100}$$

Examples of interference functions are given in the bottom panels of Fig. 11.11.
Since $\rho(r) = \rho_0 g(r)$, the interference function $i(K)$ and the pair distribution function $g(r)$ are connected by the Fourier sin-transforms:

$$K\,i(K) \;=\; 4\pi\rho_0 \int_0^\infty [g(r)-1]\,r\,\sin(Kr)\,dr\,, \tag{11.101}$$

$$g(r)-1 \;=\; \frac{1}{2\pi^2\,r\,\rho_0}\int_0^\infty i(K)\,\sin(Kr)\,K\,dK\,. \tag{11.102}$$

Once the $g(r)$ has been obtained, one can thus reconstruct the radial distribution functions $R(r) = 4\pi r^2\rho_0 g(r)$. Examples of radial distribution functions for SiO_2 are given in Fig. 11.12.

11.6.3 Elastic scattering from crystals

A crystal structure is the convolution of a lattice function and a basis function (§ 2.4):

$$S(\vec{r}) \;=\; L(\vec{r}) \,*\, F(\vec{r}) \;=\; \sum\nolimits_{\vec{T}} F(\vec{r}-\vec{T})\,. \tag{11.103}$$

The lattice function is

$$L(\vec{r}) \;=\; \sum\nolimits_{\vec{T}} \delta(\vec{r}-\vec{T})\,, \qquad \vec{T} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3\,. \tag{11.104}$$

For the interpretation of elastic scattering patterns it is convenient to refer to the *conventional* unit cells, which can contain more than one Bravais lattice point, and generally exhibit a larger number of symmetry elements than the primitive cells. The translation vector $\vec{T}$ in (11.104) labels the conventional unit cells.

The basis function can be formally expressed, for a *monatomic system*, as

$$F(\vec{r}) \;=\; \rho_T(\vec{r}) \;*\; \sum\nolimits_{\alpha} \delta(\vec{r} - \vec{R}_\alpha) \tag{11.105}$$

where the sum is over all the atoms forming the basis and $\rho_T(\vec{r}) = \rho(\vec{r}) * w(\vec{r})$ is the electron density for one atom defined in (11.81), which includes the effect of atomic vibrations.

Scattering amplitude

The amplitude of scattering (11.86) can be written as the Fourier transform of the function $S(\vec{r})$:

$$A_{\text{e.u.}}(\vec{K}) \;=\; \mathcal{F}\,[S(\vec{r})] \;=\; \int S(\vec{r})\, e^{i\vec{K}\cdot\vec{r}}\, d\vec{r} \tag{11.106}$$

and, according to the convolution theorem,

$$\mathcal{F}\,[S(\vec{r})] \;=\; \mathcal{F}\,[L(\vec{r})] \;\times\; \mathcal{F}\,[F(\vec{r})]\;. \tag{11.107}$$

One can easily verify that the intensity $I_{\text{e.u.}}(\vec{K}) = |A_{\text{e.u.}}(\vec{K})|^2$ is the Fourier transform of the autocorrelation function of $S(\vec{r})$.

Let us now separately consider first the Fourier transform of the lattice function, then the Fourier transform of the basis function.

Fourier transform of the lattice function

The scattering amplitude factor due to the lattice is the Fourier transform of the lattice function. For a finite crystal (see § 3.1 and 3.2) the scattering amplitude is

$$A_{\text{e.u.}}(\vec{K}) \;=\; \mathcal{F}\,[L(\vec{r})] \;=\; \sum\nolimits_{\vec{T}} e^{i\vec{K}\cdot\vec{T}} \tag{11.108}$$

and the scattering intensity is (see § 11.10.7 for mathematical details)

$$\begin{aligned} I_{\text{e.u.}}(\vec{K}) \;&=\; |\mathcal{F}\,[L(\vec{r})]|^2 \;=\; \left|\sum\nolimits_{\vec{T}} e^{i\vec{K}\cdot\vec{T}}\right|^2 \\ &=\; \frac{\sin^2(N_1\vec{K}\cdot\vec{a}_1/2)}{\sin^2(\vec{K}\cdot\vec{a}_1/2)} \times \frac{\sin^2(N_2\vec{K}\cdot\vec{a}_2/2)}{\sin^2(\vec{K}\cdot\vec{a}_2/2)} \times \frac{\sin^2(N_3\vec{K}\cdot\vec{a}_3/2)}{\sin^2(\vec{K}\cdot\vec{a}_3/2)} \end{aligned} \tag{11.109}$$

where N_i is the number of cells along the i-th direction and $\vec{a}_i$ the corresponding lattice parameter. The righ-hand side of (11.109) is the *Laue interference function*. One can easily verify that the function

$$\frac{\sin^2(N\vec{K}\cdot\vec{a}_i/2)}{\sin^2(\vec{K}\cdot\vec{a}_i/2)} \tag{11.110}$$

has the value N^2 for $\vec{K}\cdot\vec{a}_i/2 = n_i\pi$, where n_i is an integer number (Fig. 11.13). When N is large, the function is nearly zero for $\vec{K}\cdot\vec{a}_i/2 \neq n_i\pi$.

Conditions for diffraction

For an infinite crystal, the Fourier transform of the lattice function (§ 3.2) is

$$\mathcal{F}\,[L(\vec{r})] \;=\; \frac{1}{V}\sum\nolimits_{\vec{G}} \delta(\vec{K} - \vec{G}), \tag{11.111}$$

where $\vec{G}$ is a vector of the reciprocal space.

For an infinite crystal the intensity of the scattered beam is different from zero only if the following condition (the *Ewald condition*) is satisfied:

$$\boxed{\vec{K} = \vec{G}}, \tag{11.112}$$

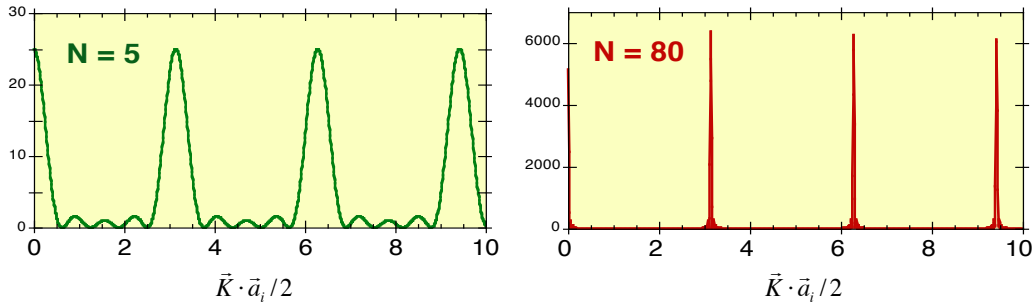

Figure 11.13: Laue function for N=5 and N=80.

say when the scattering vector $\vec{K} = \vec{k}_{\rm out} - \vec{k}_{\rm in}$ corresponds to a vector $\vec{G}$ of the reciprocal lattice of the crystal.

Expressing the reciprocal lattice vector in terms of the primitive vectors of the reciprocal space, $\vec{G} = m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3$ and remembering the relations between primitive vectors of the real and reciprocal spaces, it is easy to verify that the Ewald condition (11.112) corresponds to the so-called *Laue condition*

$$\left.\begin{array}{rl} \vec{K}\cdot\vec{a}_1 = & 2\pi\, h \\ \vec{K}\cdot\vec{a}_2 = & 2\pi\, k \\ \vec{K}\cdot\vec{a}_3 = & 2\pi\, \ell \end{array}\right\} \quad (h, k, \ell \text{ integer}) \tag{11.113}$$

which defines the values of $\vec{K}$ corresponding to the maxima of the Laue interference function.

For a perfectly monochromatic and collimated beam impinging along a fixed direction on a single crystal with a given orientation, the Ewald or Laue conditions are generally not fulfilled.
To overcome this difficulty and obtain a number of diffracted beams suitably large to extract structural information, one can use different experimental approaches, e.g.

- white beam instead of monochromatic beam (Laue method)
- rotation of the single crystal during the experiment (single crystal method)
- polycrystalline sample (powder method)

For any reciprocal lattice vector $\vec{G}$, there is a family of lattice planes normal to $\vec{G}$ and separated by a distance d, where $2\pi/d$ is the length of the shortest reciprocal lattice vector parallel to $\vec{G}$.
Taking into account the expression for the magnitudes of $\vec{K}$ and $\vec{G}$, the Ewald condition (11.112) can be expressed as

$$K \;=\; 4\pi\frac{\sin\theta_B}{\lambda} \;=\; G \;=\; n\,\frac{2\pi}{d}\,, \tag{11.114}$$

which leads to the well known *Bragg condition* for diffraction

$$\boxed{2d\,\sin\theta_B = n\lambda} \tag{11.115}$$

where θ is the angle of incidence (between the incoming beam and the plane) and n is an integer.

Example 1: Suppose that the Bragg condition is satisfied for $n = 1$ by a family of planes $(hkl) = (311)$. This situation corresponds to a reciprocal lattice vector $\vec{G}_1$ perpendicular to the planes (311) and of magnitude $G_1 = 2\pi/d_{311}$. Physically, the Bragg condition with $n = 1$ corresponds to a path difference between the beams scattered by two adjacent planes of one wavelength.

Example 2: Suppose now that the Bragg condition is satisfied by the same family of planes, but with $n = 2$. This corresponds to a reciprocal lattice vector $\vec{G}_2 = 2\vec{G}_1$, again perpendicular to planes (311), but with magnitude $G_2 = 2\pi/d_{622}$. This situation corresponds formally to a family of planes at a distance $d_{311}/2$. Physically, the Bragg condition with $n = 2$ corresponds to a path difference between the beams scattered by two adjacent (311) planes of two wavelengths.

Reciprocal space and Brillouin condition

The Bragg condition is completely cast into the direct space. The Laue conditon can be considered from the viewpoint solely of the real space. The Ewald condiction establishes a connection between direct and reciprocal space.
In Solid State Physics, when considering the scattering of the valence electrons by the ions of the crystal lattice, it is useful to consider a further condition, the so called *Brillouin condition*, which only refers to the reciprocal space.
Let us consider the incoming wavevector $\vec{k}_{\text{in}}$ as a vector of the reciprocal space, with the origin at the origin of the reciprocal space. Let us rewrite the Ewald condition as

$$\vec{k}_{\text{out}} = \vec{k}_{\text{in}} - \vec{G}\,. \tag{11.116}$$

Notice that both $\vec{k}_{\text{in}}$ and $\vec{k}_{\text{out}}$ are vectors of the reciprocal space, but not vectors of the reciprocal lattice. $\vec{G}$ is instead a vector of the reciprocal lattice.
Squaring (11.116) and taking into account that $k_{\text{in}} = k_{\text{out}} = k$, one obtains

$$k^2 = k^2 - 2\vec{k}_{\text{in}} \cdot \vec{G} + G^2\,, \tag{11.117}$$

say

$$\boxed{\vec{k}_{\text{in}} \cdot \hat{G} = G/2} \tag{11.118}$$

where $\hat{G}$ is the unit vector in the direction of $\vec{G}$.
Diffraction of valence electrons is thus possible if the tip of the wavevector $\vec{k}_{\text{in}}$ lays on a plane of the reciprocal space that bisects a reciprocal lattice vector. Such planes (of the reciprocal space!) are called "Bragg planes". The Bragg planes divide the reciprocal space into regions. The innermost region is the Wigner-Seitz cell around the origin of the reciprocal space, which is called First Brillouin zone.

The structure factor

Let us now come back to (11.107). The scattering amplitude factor due to the basis is the Fourier transform of the basis function $F(\vec{r})$ and is called structure factor:

$$F(\vec{K}) \;=\; \sum_{\alpha} f_\alpha(\vec{K})\, e^{i\vec{K}\cdot\vec{R}_\alpha}\,, \tag{11.119}$$

where f_α is the scattering factor of atom α, including the thermal factor.
The sought information in diffraction experiments on crystals is generally the structure within the unit cells (primitive or conventional). This information is contained in the basis functions $F(\vec{r})$ or, equivalently, in the structure factor $F(\vec{K})$.

If the position vectors and the scattering vector are decomposed as

$$\vec{R}_\alpha \;=\; x_{1\alpha}\vec{a}_1 + x_{2\alpha}\vec{a}_2 + x_{3\alpha}\vec{a}_3\,, \qquad \vec{K} \;=\; \vec{G} \;=\; h\vec{b}_1 + k\vec{b}_2 + \ell\vec{b}_3\,, \tag{11.120}$$

the structure factor is expressed as

$$F(\vec{K}) \;=\; \sum_{\alpha} f_\alpha(\vec{K})\, \exp[2\pi i(x_{1\alpha}h + x_{2\alpha}k + x_{3\alpha}\ell)]\,. \tag{11.121}$$

Global scattering intensity

The complete expression of the scattering intensity from a finite crystal is

$$I_{\text{e.u.}} \;=\; \prod_{s=1}^{3} \frac{\sin^2(N_s\vec{K}\cdot\vec{a}_s/2)}{\sin^2(\vec{K}\cdot\vec{a}_s/2)} \;\times\; |F(\vec{K})|^2\,, \tag{11.122}$$

where the first factor depends on the geometric lattice, the second factor depends on the content of the unit cell.
The structure factor $F(\vec{K})$ modulates the relative intensities of the diffracted beams. A diffraction experiment samples the structure factor $F(\vec{K})$ only at the reciprocal lattice points $\vec{K} = \vec{G}$ for which the conditions for lattice diffraction are fulfilled. For these $\vec{K}$ values, the intensity is enhanced by a factor proportional to the square of the number of lattice points.
The interference between waves scattered by different atoms of the same cell can give rise to selection rules. For example, it is easy to verify that, for an fcc cell containing one atom per Bravais lattice point (it is the case of Cu, Ag, Au):

F = 4f when h, k, ℓ are all even or all odd,
F = 0 when h, k, ℓ are mixed.

At last, notice again that we measure an intensity $|F(\vec{K})|^2$, so that information of the phase of $F(\vec{K})$ is lost.

11.7 Temperature effects on diffraction peaks

Atomic vibrations have a remarkable effect on the intensity of diffraction peaks. In order to maintain the notation as simple as possible, we eliminate the distinct label for cells and atoms inside the cell. Atomic positions are labelled by one index, $\vec{R}_m$.
If atoms were fixed at their equilibrium positions, the amplitude of X-rays scattering would be

$$A_{\text{eu}}(\vec{K}) \;=\; \sum_{m=1}^{N} f_m(\vec{K})\, e^{i\vec{K}\cdot\vec{R}_m} \tag{11.123}$$

where $\vec{K}$ is the scattering vector and $f_m(\vec{K})$ is the atomic scattering factor.
To take into account the effect of atomic vibrations, one substitutes $\vec{R}_m \to \vec{R}_m + \vec{u}_m$, where $\vec{u}_m$ is the instantaneous displacement of atom m, so that (11.123) becomes, for an instantaneous configuration of atomic positions,

$$A_{\text{eu}}(\vec{K}) \;=\; \sum_{m=1}^{N} f_m(\vec{K})\, e^{i\vec{K}\cdot(\vec{R}_m+\vec{u}_m)}\,. \tag{11.124}$$

The connection between atomic displacements $\vec{u}_m$ and normal modes has been explored in Sections 7.2 and 7.4.
The measured intensity is the squared modulus of the amplitude (11.123) averaged over the atomic displacements:

$$I_{\text{eu}}(\vec{K}) \;=\; \sum_{mn} f_m(\vec{K})\, f_n^*(\vec{K})\, e^{i\vec{K}\cdot\vec{R}_{mn}} \left\langle e^{i\vec{K}\cdot(\vec{u}_m-\vec{u}_n)} \right\rangle . \tag{11.125}$$

Let us now find a more interesting expression for the average $\langle \ldots \rangle$ in (11.125). One can demonstrate (see § 11.10.8) that for a gaussian distribution, such as the distribution of positions of a quantum harmonic oscillator,

$$\left\langle e^{ix} \right\rangle \;=\; e^{-\langle x^2\rangle/2}\,. \tag{11.126}$$

As a consequence, the average of (11.125) can be rewritten as the product of three factors:

$$\begin{aligned} \left\langle e^{i\vec{K}\cdot(\vec{u}_m-\vec{u}_n)} \right\rangle \;&=\; e^{-\frac{1}{2}\langle[\vec{K}\cdot(\vec{u}_m-\vec{u}_n)]^2\rangle} \\ &=\; e^{-\frac{1}{2}\langle[\vec{K}\cdot\vec{u}_m]^2\rangle}\; e^{-\frac{1}{2}\langle[\vec{K}\cdot\vec{u}_n]^2\rangle}\; e^{\langle(\vec{K}\cdot\vec{u}_m)(\vec{K}\cdot\vec{u}_n)\rangle} \end{aligned} \tag{11.127}$$

It is convenient to extract the magnitude of the scattering vector from the average, and rewrite (11.127) as

$$\left\langle e^{i\vec{K}\cdot(\vec{u}_m-\vec{u}_n)} \right\rangle \;= e^{-\frac{1}{2}K^2\langle[\hat{K}\cdot\vec{u}_m]^2\rangle}\; e^{-\frac{1}{2}K^2\langle[\hat{K}\cdot\vec{u}_n]^2\rangle}\; e^{K^2\langle(\hat{K}\cdot\vec{u}_m)(\hat{K}\cdot\vec{u}_n)\rangle}\,. \tag{11.128}$$

1. The first two factors in (11.128) depend on the independent atomic oscillations along the direction of the scattering vector $\hat{K}$.
   It is easy to recognise that the average values correspond to the mean square displacements MSD considered in § 7.4: $\langle[\hat{K}\cdot\vec{u}_n]^2\rangle = \langle u_\alpha^2(\kappa)\rangle$, where the direction α corresponds to the direction $\hat{K}$ of the scattering vector.
   The temperature dependence of the MSD has been studied in § 7.4.
2. The last factor in (11.128) depends on the correlation of motion of atom m and atom n. The correlation rapidly reduces when the distance increases, so the contribution of this term is quite weak when summed over all atomic positions.
   It is convenient to expand the last factor in (11.128) as

$$e^{K^2\langle(\hat{K}\cdot\vec{u}_m)(\hat{K}\cdot\vec{u}_n)\rangle} \simeq 1 + K^2\langle(\hat{K}\cdot\vec{u}_m)(\hat{K}\cdot\vec{u}_n)\rangle + \dots . \tag{11.129}$$

Making use of the expansion (11.129), the intensity (11.125) can be rewritten as a sum of terms of decreasing importance.

$$\begin{aligned} I_{\rm eu}(\vec{K}) &= \sum_{mn} f_m(\vec{K}) e^{-\frac{1}{2}K^2\langle(\hat{K}\cdot\vec{u}_m)^2\rangle}\ f_n^*(\vec{K}) e^{-\frac{1}{2}K^2\langle(\hat{K}\cdot\vec{u}_n)^2\rangle}\ e^{i\vec{K}\cdot\vec{R}_{mn}} \\ &+ \sum_{mn} f_m(\vec{K}) e^{-\frac{1}{2}K^2\langle(\hat{K}\cdot\vec{u}_m)^2\rangle}\ f_n^*(\vec{K}) e^{-\frac{1}{2}K^2\langle(\hat{K}\cdot\vec{u}_n)^2\rangle}\ e^{i\vec{K}\cdot\vec{R}_{mn}} K^2\langle(\hat{K}\cdot\vec{u}_m)(\hat{K}\cdot\vec{u}_n)\rangle \\ &+ \cdots . \end{aligned} \tag{11.130}$$

1. The first term in (11.130) is the most important. The sum gives the Laue interference function and is responsible for the diffraction peaks (Chapter 11). The effect of temperature is accounted for by multiplying each atomic scattering factor by a thermal factor

$$f_m(\vec{K}) \quad\to\quad f_m(\vec{K})\ e^{-\frac{1}{2}K^2\langle(\hat{K}\cdot\vec{u}_m)^2\rangle} \tag{11.131}$$

   that corresponds to the thermal factor (11.82) introduced in Section 11.5.
   The thermal effect induces a reduction of the intensity of Bragg peaks; the reduction is larger for higher temperatures and for larger magnitudes of the scattering vector $\vec{K}$.
   For a monatomic Bravais crystal, the atomic scattering factors and the MSDs are equal for all atoms, $f_m = f_n$ and $\langle(\hat{K}\cdot\vec{u}_m)^2\rangle = \langle(\hat{K}\cdot\vec{u}_n)^2\rangle$, so that each term of the sum contains the their product

$$|f_m(\vec{K})|^2\, e^{-K^2\langle(\hat{K}\cdot\vec{u}_m)^2\rangle} . \tag{11.132}$$

   The factor $\exp[-K^2\langle(\hat{K}\cdot\vec{u}_m]^2\rangle)$ is called Debye-Waller factor.
2. In the second term of (11.130), the correlation factor $K^2\langle(\hat{K}\cdot\vec{u}_m)(\hat{K}\cdot\vec{u}_n)\rangle$ is important only for neighbouring atoms, and progressively decreases when the inter-atomic distance increases. The second term in (11.130) gives rise to the so called diffuse scattering, a low intensity background of the diffraction spectra, slightly structured in correspondence of the diffraction peaks generated by the first term.

As for the specific heat, also for the Debye-Waller factor the harmonic theory is sufficient to account for the main properties. In general, anharmonicity effects are much less important.

Note 1: For non-monatomic crystals, thermal effects are different for different atoms. The difference is accounted for by the magnitude of eigenvectors of the dynamical matrix.

Note 2: Thermal diffuse scattering contains information on the correlation of atomic vibrations. For non monatomic crystals, it depends on the phase relationships bewteen eigenvectors of the dynamical matrix.

Note 3: The knowledge of the lattice dynamical properties (eigenfrequencies and eigenvectors) allows one to calculate the thermal factors of diffraction. Viceversa, one cannot recover the full information on lattice dynamics from the thermal factors.

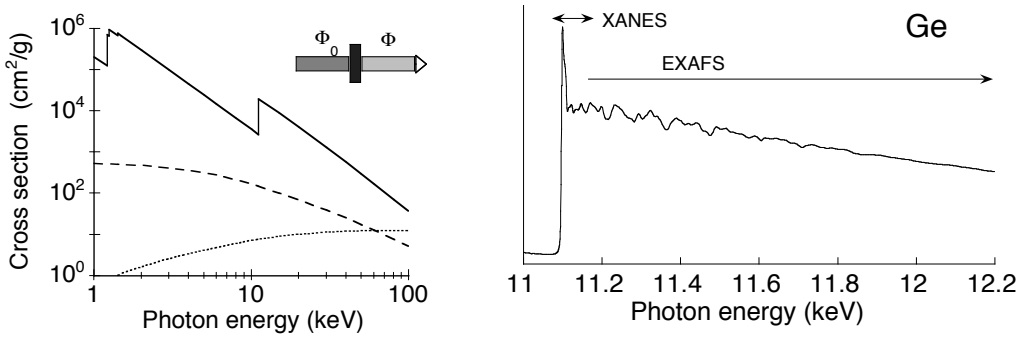

Figure 11.14: Left: Cross section for photoelectric absorption (continuous line), elastic scattering (dashed line) and Compton scattering (dotted line) for Germanium. Right: Fine structures at the K absorption edge of Germanium

11.8 Extended X-ray Absorption Fine Structure (EXAFS)

Let us now introduce a structural probe that is based on the joint effect of photoelectric absorption of an X-ray photon by an atom and scattering of the emitted photoelectron by other surrounding atoms. EXAFS gives information on the short-range structure around selected atomic species.

11.8.1 Introduction

Photo-absorption versus scattering

The intensity Φ_0 of an X-ray beam traversing a sample of thickness x is reduced according to the exponential law

$$\Phi = \Phi_0 \exp[-\mu(\omega)x]. \tag{11.133}$$

where the linear attenuation coefficient $\mu(\omega)$ depends on the energy of X-ray photons and on the sample composition and density. In the energy range between about 1 and 40 keV, two different basic mechanisms contribute to the X-ray beam attenuation (fig. 11.14, right):

1. The dominant interaction is photo-electric absorption: one photon is absorbed from the beam and an atom is ionized or excited.
2. The other interaction is scattering, which has been considered in previous Sections.

In the energy range from 1 to 40 keV, photo-electric absorption is dominant.

The X-ray absorption coefficient

Let us now consider only photo-electric absorption. When the energy $\hbar\omega$ of X-ray photons increases, the absorption coefficient $\mu(\omega)$ progressively decreases. This smooth behavior is interrupted by sharp discontinuities, the *absorption edges* (fig. 11.14), which originate when the photon energy becomes high enough to extract an electron from a deeper level. The highest-energy absorption edges, the K edges, correspond to the extraction of an electron from the deepest level (1s level). Since the binding energies increase monotonically with the atomic number Z, an edge energy corresponds to a well defined atomic species.
If the energy of the absorbed photon is larger than the binding energy, the electron (*photo-electron*) is ejected from the atom. Above the edge, the absorption coefficient exhibits an oscillating structure. The structure extending from 30÷50 eV up to typically one thousand eV above the edge (fig. 11.14, right, is called *EXAFS* (Extended X-ray Absorption Fine Structure), and carries information on the local geometric structure surrounding a given atomic species,

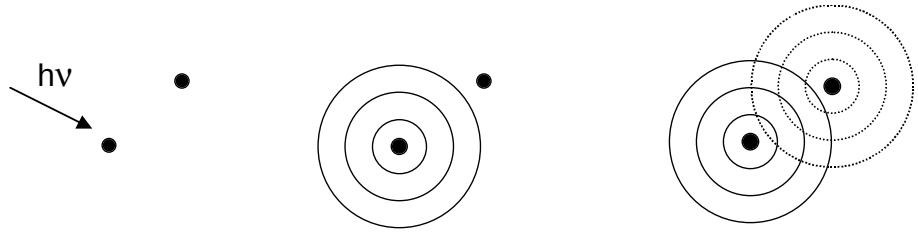

Figure 11.15: Schematic picture of the EXAFS phenomenon: the incoming X-ray photon (left), the outgoing photo-electron spherical wave (centre), the final state superposition of outgoing and backscattered spherical waves (right).

The EXAFS mechanism

The kinetic energy of the photo-electron is the difference between the photon energy $\hbar\omega$ and the core binding energy E_b. The outgoing photo-electron is described as a spherical wave, whose wavelength λ decreases when the photon energy $\hbar\omega$ increases (Fig. 11.15). If the absorber atom is not isolated, the photo-electron can be back-scattered by neighboring atoms, giving rise to an incoming spherical wave. The absorption coefficient $\mu(\omega)$ depends, to first order, on the dipole matrix element between the initial electron core state $|\psi_i\rangle$ and the photo-electron final state $|\psi_f\rangle$, which in turn is a superposition of the outgoing and the incoming spherical waves. The phase relationship between outgoing and incoming waves depends on the photo-electron wavelength and on the inter-atomic distance R. The variation of the phase relationship as a function of photon energy $\hbar\omega$ influences the final state amplitude at the core site, giving rise to a modulation of the absorption coefficient. The frequency of the EXAFS oscillations depends on the distance between absorber and back-scatterer atoms. Their amplitude is proportional to the number of back-scatterer atoms.
Two main peculiarities characterize EXAFS:

1. the selectivity of atomic species, which is achieved by tuning the X-ray energy to the corresponding absorption edge;
2. the insensitivity to long-range order, due to the short mean free path of the photo-electron, typically limited to about 10 Å.

11.8.2 Formal treatment

Photoelectric absorption

The absorption coefficient $\mu(\omega)$ depends on the atomic density n of the sample and on the probabilities of transition per unit time w_{fi} from the stationary initial ground state $|\Psi_i\rangle$ of the atom to all the possible stationary final states $|\Psi_f\rangle$:

$$\mu(\omega) \;=\; \frac{2\hbar}{\epsilon_0 \omega A_0^2 c}\; n \sum\nolimits_f w_{fi}\,, \tag{11.134}$$

where A_0 is the magnitude of the vector potential of the electromagnetic field,

$$\vec{A}(\mathbf{r}) \;=\; A_0\, \hat{\eta}\, e^{i\vec{k}\cdot\vec{r}} \;+\; c.c. \tag{11.135}$$

Within the framework of the time dependent perturbation theory, the density of transition probability per unit time and energy interval is given by the golden rule

$$w_{fi} \;=\; \frac{2\pi}{\hbar}\; \left|\langle\, \Psi_f|\hat{H}_I|\Psi_i\,\rangle\right|^2\, \rho(E_f)\, \delta(E_f - E_i - \hbar\omega)\,, \tag{11.136}$$

where $\hat{H}_I$ is the interaction hamiltonian for photoelectric absorption

$$\hat{H}_I \;=\; \frac{e}{m} \sum\nolimits_j \mathbf{p_j} \cdot \vec{A}(\vec{r}_j)\,, \tag{11.137}$$

$\rho(E_f)$ is the density of final continuum states and the sum is over all the electrons inside the atom.
The problem is simplified by a set of approximations.

1. One-electron approximation: only one core electron changes its state and the remaining $N-1$ passive electrons simply relax their orbitals around the core hole.
2. Electric dipole approximation: The exponential in (11.135) is approximated $\exp(i\vec{k}\cdot\vec{r}) \simeq 1$.
3. Sudden approximation: in the EXAFS region, the photo-electron energy is high enough that its interaction with the passive electrons can be neglected.

Within this set of approximations, the one-electron absorption coefficient is

$$\mu_{\rm el}(\omega) \;\propto\; |\,\langle\,\psi_i|\;\hat{\eta}\cdot\vec{r}\;|\psi_f\,\rangle\,|^2\;\; S_0^2\;\rho(\epsilon_f) \tag{11.138}$$

where ψ_i and ψ_f are one-electron wavefunctions and S_0^2 is the superposition integral of the passive electrons wavefunctions, whose value is generally included between 0.7 and 0.9.
It is the final state ψ_f of eq. (11.138) that carries the sought structural information.

The EXAFS function

It is convenient to introduce the photo-electron wavevector $k \;=\; (2\pi/\lambda)$:

$$k \;=\; \sqrt{(2m/\hbar^2)\;\epsilon_f} \;=\; \sqrt{(2m/\hbar^2)\;(\hbar\omega - E_b)} \tag{11.139}$$

where ϵ_f is the photoelectron energy and E_b is the core electron binding energy.
If the absorber atom is isolated (like in monatomic gases), the final state $|\psi_f^0\rangle$ is represented by an outgoing wave.
In molecular gases and condensed systems, the photo-electron can interact with the surrounding atoms and undergo scattering. The interaction causes a weak perturbation to the final state $|\psi_f\,\rangle \;=\; |\psi_f^0 \,+\, \delta\psi_f\,\rangle$ and the absorption coefficient becomes

$$\mu(\omega) \;\propto\; \left|\,\langle\,\psi_i|\;\hat{\eta}\cdot\mathbf{r}\;|\psi_f^0 \,+\, \delta\psi_f\,\rangle\,\right|^2 . \tag{11.140}$$

The EXAFS function is defined as the difference between the actual absorption coefficient μ and the atomic absorption coefficient μ_0, normalized to μ_0:

$$\chi(k) \;=\; (\mu - \mu_0)/\mu_0\,. \tag{11.141}$$

Typical EXAFS functions are shown in Fig. 11.16. The EXAFS function is similar to the interference function $i(K)$ encountered in § 11.6.2 when considering X-ray scattering from liquids and non-crystalline solids. Due to the insensitivity of EXAFS to long range order, the EXAFS function is very similar for crystalline and amorphous samples. Besides, the scattering vector of diffraction experiments corresponds to twice the EXAFS photoelectron, $K = 2k$.

EXAFS for a two-atomic system

The simplest system consists of two atoms, an absorber one A and a back-scatterer one B; let R be the distance between the two atoms and neglect any vibrations.
One can show that the EXAFS oscillations for an isotropic sample can be interpreted as

$$\chi(k) \;=\; \frac{1}{kR^2}\;\mathrm{Im}\,\left\{f_B(k,\pi)\;e^{2i\delta_1}\;e^{2ikR}\right\}\,, \tag{11.142}$$

where $f_B(k,\pi)$ is the complex amplitude of backscattering from atom B and δ_1 is the phase-shift due to the potential of atom A. By separating magnitude and phase of the complex backscattering amplitude and grouping the phase terms,

$$f_B(k,\pi)\;e^{2i\delta_1} \;=\; |f_B(k,\pi)|\;e^{i\phi}, \tag{11.143}$$

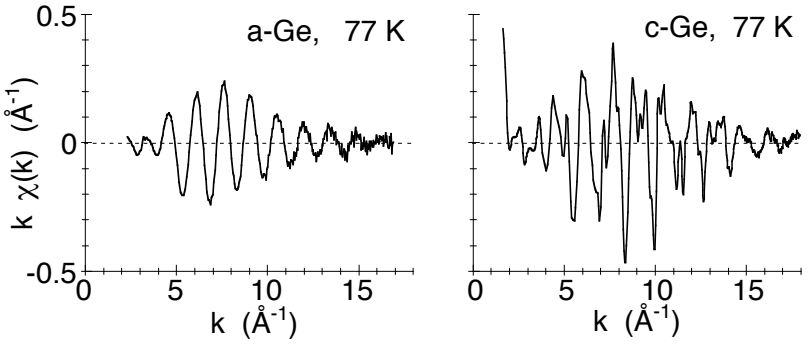

Figure 11.16: EXAFS oscillations in amorphous (rleft) and crystalline (right) Germanium, measured at 77 K.

one can write eq. (11.142) in the purely real form:

$$\chi(k) \;=\; \frac{1}{kR^2}\;|f_B(k,\pi)|\;\sin\left[2kR+\phi(k)\right]. \tag{11.144}$$

Basically, the EXAFS signal has a sinusoidal behaviour in the k space, with frequency $2R$ proportional to the inter-atomic distance. The phase of the sine function is perturbed by the phase-shift $\phi(k)$, while the amplitude is modulated by $|f_B(k,\pi)|$. The k dependence of backscattering amplitudes and phaseshifts is different for different atomic species.

EXAFS for many-atomic systems

For a system composed of any number of atoms, the EXAFS function can be built up as the sum of many two-atomic contributions, with different interatomic distances R_i from the absorber atom, which can be grouped into coordination shells, each one containing N_s atoms of the same species at the same distance R_s from the absorber atom:

$$\boxed{\chi(k) \;=\; (S_0^2/k)\sum\nolimits_j (e^{-2R_j/\lambda}/R_j^2)\;\mathrm{Im}\;\left\{f_j(k,\pi)\;e^{2i\delta_1}\;e^{2ikR_j}\right\}} \tag{11.145}$$

where two further factors have been introduce, to account for inelastic effects:

- *Intrinsic inelastic effects*, say multiple excitations within the absorber atom, give rise to a reduction of the coherent EXAFS signal with respect to that expected for purely elastic excitations, and are taken into account by the factor S_0^2, which typically amounts to $0.7 \div 0.9$.
- *Extrinsic inelastic effects* are due to the inelastic interactions with other electrons outside the absorber atom and are taken into account by a phenomenological mean free path factor $\exp(-2R_j\lambda)$, $\lambda \simeq 5 \div 15$ Å, which contributes to the short-range sensitivity of EXAFS.

11.8.3 Disorder effects on EXAFS

The period of atomic vibrations ($\simeq 10^{-12}$ s) is much larger than the photo-electron time of flight ($10^{-16} \div 10^{-15}$ s). An EXAFS spectrum samples a very large set of instantaneous atomic configurations, corresponding to a distribution of instantaneous interatomic distances. The distribution of interatomic distances can be further enlarged and modified by the presence of structural disorder.

EXAFS formula including disorder

Due to disorder, the distance between absorber and back-scatterer atoms varies according to a probability distribution $\rho(r)$, which can be considered as a partial radial distribution function, say the radial distribution function relative to a given absorption species

The EXAFS equation for one coordination shell becomes:

$$\boxed{\chi_s(k) \;=\; (S_o^2/k)\; N_s \mathrm{Im}\left\{ f_s(k,\pi)\; e^{2i\delta_1} \int_0^\infty \rho(r)\; (e^{-2r/\lambda}/r^2)\; e^{2ikr}\; dr \right\}} \tag{11.146}$$

The EXAFS function is the imaginary part of the Fourier transform of the distance distribution, weighted by the factor $\exp(-2r/\lambda)/r^2$.

Note: The probability distribution $\rho(r)$ of (11.146) corresponds to the radial distribution function $R(r)$ introduced in § 2.6, and shouldn't be confused with the density distribution of § 2.6 and § 11.6.2. The mismatch has historical origin.

The fundamental problem of EXAFS analysis is to recover the distribution $\rho(r)$ from the experimental spectrum $\chi(k)$. No exact solution can be given to this problem, because every experimental spectrum has a finite extension, within the values $k_{\min}$ and $k_{\max}$.
An approximate solution to the inversion problem from $\chi(k)$ to $\rho(r)$ consists in hypothesizing physically sound structural models and in fitting the parameters of their distributions $\rho(r)$ to the experimental EXAFS spectrum.

Parametrization of EXAFS formula

In the great majority of applications, the extent of disorder is sufficiently small to allow the expression of the EXAFS formula in terms of a few standard parameters, which characterise the position, width and shape of the distance distribution.
The EXAFS formula (11.146) can be expressed as

$$\chi_s(k) = \frac{S_o^2}{k}\; N_s\; |f_s(k,\pi)|\; \frac{e^{-2C_1/\lambda}}{C_1^2}\; e^{-2k^2C_2+2k^4C_4/3\cdots}\; \sin\left[2kC_1 - \frac{4k^3C_3}{3} \ldots + \phi(k)\right], \tag{11.147}$$

where the parameters C_i are called cumulants of the distribution. The cumulant C_1 corresponds, to within a correction factor, to the average value $\langle r \rangle$ of the distribution, $C_2 = \sigma^2$ is the variance and C_3 is a measure of the distribution asymmetry.

Summary

The following parameters can be obtained for each coordination shell:

- Coordination number, say number of atoms within the coordination shell.
- Average inter-atomic distance $\langle r \rangle$.
- Variance σ^2 of the distribution of distances, which carries information on thermal and structural disorder. The exponential $\exp[-2k^2\sigma^2]$ is referred to as the EXAFS Debye-Waller factor.
- Third cumulant, which measures the asymmetry of the distribution of distances.

This information can be obtained from EXAFS spectra, provided the "physical quantities" $|f(k,\pi)|$, ϕ, S_0^2 and λ are known. Two different approaches are possible: *a)* these quantities are experimentally obtained from the EXAFS of a reference sample of known structure, *b)* they are calculated ab initio.

11.8.4 Interpretation of EXAFS parameters

Let us now discuss the physical meaning of the parameters obtained from EXAFS. To this purpose, it is convenient to focus the attention on the first coordination shell of a crystal and compare the effect of thermal motion on the EXAFS signal and on the Bragg diffraction patterns.

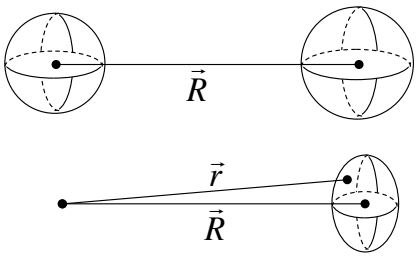

Figure 11.17: Effect of atomic vibrations. Atomic thermal ellipsoids measured by Bragg diffraction (top) and relative thermal ellipsoid sampled by EXAFS (bottom); $\vec{R}$ is the distance between average atomic positions, and $\vec{r}$ is the instantaneous distance.

Effects of atomic thermal displacements

Let $\vec{R}$ be the equilibrium distance and $\vec{u}_a$ and $\vec{u}_b$ the instantaneous thermal displacements of the absorber and back-scatterer atoms. The instantaneous distance is

$$\vec{r} = \vec{R} + \Delta\vec{u}, \tag{11.148}$$

where $\Delta\vec{u} = \vec{u}_b - \vec{u}_a$ is the relative displacement of the two atoms. It is convenient to introduce the projections of the relative displacement $\Delta\vec{u}$ parallel and perpendicular to the inter-atomic bond:

$$\Delta u_{\parallel} = \hat{\vec{R}} \cdot \Delta\vec{u}\,, \qquad \Delta u_{\perp}^2 = (\Delta u)^2 - \Delta u_{\parallel}^2\,. \tag{11.149}$$

Using eqs. (11.148) and (11.149), the scalar distance r is expressed as:

$$r \simeq R + \Delta u_{\parallel} + \Delta u_{\perp}^2/2R - [\Delta u_{\parallel}\Delta u_{\perp}^2]/2R^2. \tag{11.150}$$

Average interatomic distance

In the harmonic approximation one can see that, to first order,

$$\langle r\rangle \simeq R + \langle\Delta u_{\perp}^2\rangle/2R\,. \tag{11.151}$$

Eq. (11.151) relates the average distance measured by EXAFS, $\langle r\rangle = \langle|\vec{r}_b - \vec{r}_a|\rangle$ (average bond length) to the distance between average atomic positions $R = |\langle\vec{r}_b\rangle - \langle\vec{r}_a\rangle|$ measured by Bragg diffraction. The last term in eq. (11.151) is proportional to the Mean Square Relative Displacement in direction normal to the bond direction $\langle\,\Delta u_{\perp}^2\,\rangle$, which is always positive.
The average distance $\langle r\rangle$ is thus always larger than the crystallographic distance R; EXAFS and diffraction give different and complementary information on the inter-atomic distance in crystals. (In non-crystalline systems, EXAFS and diffraction are instead both sensitive to the average distance $\langle r\rangle$).

Parallel MSRD

The variance$\sigma^2 = \langle(r - \langle r\rangle)^2\rangle$ of the distribution $\rho(r)$ corresponds to the parallel Mean Square Relative Displacement (MSRD) and can be calculated from eq. (11.150) as

$$\sigma^2 \simeq \langle\Delta u_{\parallel}^2\rangle = \left\langle[\hat{R}\cdot(\vec{u}_b - \vec{u}_a)]^2\right\rangle \tag{11.152}$$

Expanding the square in eq. (11.152), the parallel MSRD can be rewritten as

$$\langle\Delta u_{\parallel}^2\rangle = \langle(\hat{R}\cdot\vec{u}_b)^2\rangle + \langle(\hat{R}\cdot\vec{u}_a)^2\rangle - 2\,\langle(\hat{R}\cdot\vec{u}_b)(\hat{R}\cdot\vec{u}_a)\rangle. \tag{11.153}$$

The first two terms in eq. (11.153) are the independent Mean Square Displacements (MSD) of absorber and back-scatterer atoms along the bond direction. They can be obtained from the refinement of Bragg diffraction patterns. The third term in eq. (11.153), the Displacement Correlation Function (DCF), depends on the correlation of atomic motions. The stronger the correlation,

the larger is the DCF and the smaller is the parallel MSRD $\langle \Delta u_{\parallel}^2 \rangle$, say the second cumulant σ^2. The correlation term DCF decreases with increasing distance and vanishes for very large distances. The sensitivity to correlation in crystals is peculiar to EXAFS: in diffraction experiments, as a consequence of long-range averaging, the effect of short-range correlations is dispersed into thermal diffuse scattering, and the Debye-Waller factor measures only the uncorrelated MSD. (In non-crystalline systems, both EXAFS and diffraction are equally sensitive to correlation, and measure the MSRD).

Perpendicular MSRD

The perpendicular MSRD $\langle \Delta u_{\perp}^2 \rangle$ cannot be directly obtained from experiment. It can however be calculated by inverting eq. (11.151) if R is known from diffraction experiments.
For isotropic relative atomic displacements, one expects that $\langle \Delta u_{\perp}^2 \rangle / 2 = \langle \Delta u_{\parallel}^2 \rangle$. Actually, one finds that the ratio $\langle \Delta u_{\perp}^2 \rangle / 2 \langle \Delta u_{\parallel}^2 \rangle$ is larger than one: the relative displacements of nearest-neighbouring atoms are anisotropic, and correspond to disc-shaped relative thermal ellipsoids.

Bond thermal expansion

The temperature dependence of the average distance $\langle r \rangle$ corresponds to is thermal expansion. For the first shell, it is the bond thermal expansion. Since the perpendicular MSRD $\langle \Delta u_{\perp}^2 \rangle$ increases when the temperature increases, the thermal expansion of the average distance $\langle r \rangle$ measured by EXAFS is larger than the thermal expansion of the crystallographic distance R measured by Bragg diffraction.

11.9 Inelastic scattering

The dispersion curves can be measured by inelastic scattering of suitable probes, such as neutrons or photons. Let us remember here that the differential scattering cross section can be expressed as

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_0 S(\vec{K}) \tag{11.154}$$

for elastic scattering, and

$$\frac{\partial^2 \sigma}{\partial \Omega \, \partial \omega} = \left(\frac{d\sigma}{d\Omega} \right)_0 S(\vec{K}, \omega) \tag{11.155}$$

for inelastic scattering, where the first factor on the right in both equations, $(d\sigma/d\Omega)_0$, is the intrinsic cross section, different for photons and for neutrons.

As for the second factors on the right,

- the static scattering function $S(\vec{K})$ contains information on the static structure;
- the dynamic scattering function $S(\vec{K}, \omega)$ contains information on the collective excitations of the crystal, such as normal modes of vibration.

11.9.1 Phonon scattering

Basic equations of scattering

Let

$\vec{k}_{\text{in}}$ and ω_{in} be the wavevector and frequency of the incoming particles (neutron or photons)
$\vec{k}_{\text{out}}$ and ω_{out} be the wavevector and frequency of the outgoing particles (neutron or photons)

Linear momentum and energy are given by the well known relations $\vec{p} = \hbar \vec{k}$ and $E = \hbar \omega$.

Elastic scattering

As previously noticed, in an elastic scattering event the energy conservation implies that

$$\hbar\omega_{\text{out}} = \hbar\omega_{\text{in}} \tag{11.156}$$

and the momentum conservation

$$\hbar\vec{k}_{\text{out}} = \hbar\vec{k}_{\text{in}} + \hbar\vec{G}\,, \tag{11.157}$$

where $\vec{G}$ is a reciprocal lattice vector. Equation (11.157) corresponds to the Laue interference condition, where the scattering vector is equal to a reciprocal lattice vector, $\vec{K} = \vec{G}$. Elastic scattering from a crystal lattice can take place only if the impulse released to the lattice corresponds to the in-phase movement of all the lattice points.

Inelastic scattering, general expression

Inelastic scattering takes place with exchange of energy and momentum between the probe (neutrons or photons) and the vibrational normal modes of the crystal.
The energy conservation equation is now

$$\hbar\omega_{\text{out}} = \hbar\omega_{\text{in}} - \sum_{\vec{q}s} \hbar\omega(\vec{q}s)\,\Delta n(\vec{q}s)\,, \tag{11.158}$$

where $\Delta n(\vec{q}s)$ is the variation of the number of phonons of mode $(\vec{q}s)$.
The momentum conservation law is

$$\hbar\vec{k}_{\text{out}} = \hbar\vec{k}_{\text{in}} + \hbar\vec{G} - \sum_{\vec{q}s} \hbar\vec{q}\,\Delta n(\vec{q}s)\,, \tag{11.159}$$

where $\hbar\vec{q}$ is the so-called *crystal momentum* (or quasi-momentum) of the phonons of mode $(\vec{q}s)$.

One-phonon inelastic scattering

Let us now consider the inelastic processes where only one phonon is annihilated or created, and the corresponding energy and momentum are exchanged with the radiation particles (neutrons or photons).

a) One phonon of wavevector $\vec{q}$ and frequency ω is annihilated. The corresponding energy and momentum are transferred from the crystal to particles of the probe:

   - energy conservation

$$\hbar\omega_{\text{out}} = \hbar\omega_{\text{in}} + \hbar\omega(\vec{q}s)\,, \tag{11.160}$$

   - momentum conservation

$$\hbar\vec{k}_{\text{out}} = \hbar\vec{k}_{\text{in}} + \hbar\vec{G} + \hbar\vec{q}\,. \tag{11.161}$$

b) One phonon of wavevector $\vec{q}$ and frequency ω is created. The corresponding energy and momentum are transferred from the probe particles to the crystal:

   - energy conservation

$$\hbar\omega_{\text{out}} = \hbar\omega_{\text{in}} - \hbar\omega(\vec{q}s)\,, \tag{11.162}$$

   - momentum conservation

$$\hbar\vec{k}_{\text{out}} = \hbar\vec{k}_{\text{in}} + \hbar\vec{G} - \hbar\vec{q}\,. \tag{11.163}$$

11.9.2 On the phonon crystal momentum

The conservation of total linear momentum of any isolated system is a consequence of the symmetry of the Hamiltonian of the system with respect to any spatial translation.

The conservation of crystal momentum is instead connected to the lattice symmetry of a crystal.

For the sake of concreteness, let us consider the interaction of a neutron with a crystal lattice. Let $\vec{r}$ and $\vec{p}$ be position and momentum of the neutron, $\vec{R}$, $\vec{u}$ and $\vec{P}$ be position, instantaneous displacement and momentum of the generic ion of the lattice.
The total Hamiltonian of the system is

$$H = \sum_{\vec{R}} \frac{P^2(\vec{R})}{2M} + \frac{1}{2} \sum_{\vec{R}\vec{R}'} \Phi \left[\vec{R} + \vec{u}(\vec{R}) - \vec{R}' - \vec{u}(\vec{R}') \right] + \frac{p^2}{2m} + \sum_{\vec{R}} V \left[\vec{r} - \vec{R} - \vec{u}(\vec{R}) \right] . \quad (11.164)$$

Let us now consider two different operations.

A) A generic translation $\vec{r}_0$ in the real space, inducing the following transformations

$$\vec{r} \rightarrow \vec{r} + \vec{r}_0 \quad (11.165)$$

$$\vec{u}(\vec{R}) \rightarrow \vec{u}(\vec{R}) + \vec{r}_0 . \quad (11.166)$$

The invariance of the total Hamiltonian (11.164) with respect to these transformations leads to the conservation of the total linear momentum.

B) A translation $\vec{R}_0$ in the real space, corresponding to a Bravais lattice vector, inducing the following transformations

$$\vec{r} \rightarrow \vec{r} + \vec{R}_0 \quad (11.167)$$

$$\vec{u}(\vec{R}) \rightarrow \vec{u}(\vec{R} - \vec{R}_0) . \quad (11.168)$$

The invariance of the total Hamiltonian (11.164) with respect to these transformations leads to the conservation of the crystal momentum and to (11.159).

[For more details, see Ashcroft-Mermin, Chapter 24 and Appendix M.]

11.9.3 Neutrons and photons scattering

Neutron and photon scattering are the most suited probes to study the atomic structure and the lattice dynamical properties of condensed matter, by means of elastic and inelastic scattering, respectively.

General properties of photons and neutrons

Let us review the different relation between energy and momentum (or equivalently between angular frequency and wavevector) for photons and neutrons. We consider here only thermal neutrons, say neutron thermalised at room temperature.
The relativistic relation energy-momentum is

$$E^2 = (pc)^2 + (mc^2)^2 . \quad (11.169)$$

- For *photons* $m = 0$ and (11.169) becomes

$$E = pc . \quad (11.170)$$

  Numerical relations:

$$k[^{-1}] = E[\text{keV}]/1.97 , \qquad s[] = 12.4/E[\text{keV}] .$$

- For *thermal neutrons* the kinetic energy is $E_k \ll mc^2$ and (11.169) becomes

$$E_k = p^2/2m . \quad (11.171)$$

  Numerical relation:

$$k[^{-1}] = (E[\text{keV}]/2.08)^{1/2} .$$

Orders of magnitude of the quantities to be measured

In structural and lattice dynamical studies of crystals the quantities to be considered and their order of magnitudes are the following

- Structural studies:
  - lattice parameters and inter-atomic distances $a \simeq 1\,\text{Å}$
- Lattice dynamical studies:
  - normal mode wavevectors $q \simeq 1\,\text{Å}^{-1}$
  - normal mode frequencies $\omega \simeq 1\,\text{THz}$
  - phonon energies $E = \hbar\omega \simeq 0.01$ eV

Neutrons

The order of magnitudes of relevant quantities for thermal neutrons are

$$\lambda \simeq 1\,\text{Å}\,, \qquad k \simeq 1\,\text{Å}^{-1}\,, \qquad E \simeq 50\,\text{meV}\,. \tag{11.172}$$

Thermal neutrons are thus suitable for both structural and dynamical studies.

Neutrons for scientific research can be utilised at large national or international facilities. Two kind of sources are available:

- Reactors, where neutrons are produced by the fission process as a continuous flux. The largest facility of this type in Europe is the ILL laboratory in Grenoble.
- Accelerators, where neutrons are produced by the spallation process as a pulsed flux. The largest facility of this type in Europe is the ISIS laboratory near Oxford (UK).

X-ray photons

The order of magnitudes of relevant quantities for X-ray photons are

$$\lambda \simeq 1\,\text{Å}\,, \qquad k \simeq 1\,\text{Å}^{-1}\,, \qquad E \simeq 10\,\text{keV}\,. \tag{11.173}$$

X-rays have the right wavelenghts for structural studies. The X-ray wavevectors are suitable for dynamical studies. However, the X-ray energies are several orders of magnitude larger than the phonon energies. The experimental difficulty of measuring relatively small variations of energy (some meV over some keV) has been recently resolved thanks to the peculiar properties of Synchrotron Radiation (high intensity, tunability and collimation). The largest facility for Synchrotron Radiation in Europe, where inelastic X-ray scattering experiments can be performed, is the ESRF laboratory in Grenoble.

Visible photons

The order of magnitudes of relevant quantities for visible photons are

$$\lambda \simeq 4000 \div 8000\,\text{Å}\,, \qquad k \simeq 10^{-4}\,\text{Å}^{-1}\,, \qquad E \simeq 1\,\text{eV}\,. \tag{11.174}$$

The inelastic scattering of visible photons gives information on the lattice dynamics near the BZ centre, say for small q values, corresponding to long wavelengths modes. The suitable sources are laser beams. One distinguishes

- Brillouins scattering if acoustic modes are studied
- Raman scattering if optic modes are studied

11.10 Complements and demonstrations

11.10.1 Classical electron radius

The Thomson scattering length is often called Òclassical electron radiusÓ. The classical electron radius is defined through a phenomenological argument: the potential energy of a charge e distributed on a spherical shell or on a homogeneous sphere of radius r_e is

$$U = \frac{1}{2}\left(\frac{1}{4\pi\epsilon_0}\frac{e^2}{r_e}\right), \qquad U = \frac{3}{5}\left(\frac{1}{4\pi\epsilon_0}\frac{e^2}{r_e}\right), \tag{11.175}$$

respectively.
Neglecting the numerical pre-factors and equating the potential energy U to the relativistic rest energy

$$\left(\frac{1}{4\pi\epsilon_0}\frac{e^2}{r_e}\right) = m_e c^2 \tag{11.176}$$

one obtains

$$r_e = \frac{e^2}{4\pi\epsilon_0\, c^2 m_e} = 2.8\times10^{-15}\,\mathrm{m}\,. \tag{11.177}$$

11.10.2 Total Thomson cross section

Demonstration of (11.40):

$$\begin{aligned}
\sigma_{\mathrm{Th}} &= \int \sigma(2\theta,\phi)\, d\Omega \\
&= r_e^2 \int_0^{2\pi} d\phi \int_0^{\pi} d(2\theta)\, \sin(2\theta)\, \frac{1+\cos^2(2\theta)}{2} \\
&= 2\pi\, r_e^2 \left[\frac{1}{2}\int_0^{\pi} d(2\theta)\, \sin(2\theta) + \frac{1}{2}\int_0^{\pi} d(2\theta)\, \sin(2\theta)\, \cos^2(2\theta)\right] = \frac{8}{3}\,\pi\, r_e^2
\end{aligned} \tag{11.178}$$

11.10.3 Thomson scattering intensity from Z electrons

Let us consider the amplitudes and intensities in electronic units. From (11.41) and (11.42),

$$\begin{aligned}
I_{\mathrm{e.u.}}(\vec{K}) &= \sum_{i=1}^{Z} e^{i\vec{K}\cdot\vec{r}_i} \times \sum_{j=1}^{Z} e^{-i\vec{K}\cdot\vec{r}_j} = \sum_{i=1}^{Z} e^{i\vec{K}\cdot\vec{r}_i} \times \left[e^{-i\vec{K}\cdot\vec{r}_i} + \sum_{j\neq i} e^{-i\vec{K}\cdot\vec{r}_j}\right] \\
&= Z + \sum_i \sum_{j\neq i} \cos(\vec{K}\cdot\vec{r}_{ij}) + i \sum_i \sum_{j\neq i} \sin(\vec{K}\cdot\vec{r}_{ij})\,,
\end{aligned} \tag{11.179}$$

where $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$. Since $\vec{r}_{ij} = -\vec{r}_{ji}$, the sum over the sinusoidal terms is zero.
For $Z = 2$, say for two point-like electrons at positions $\vec{r}_1$ and $\vec{r}_2$, (11.179) becomes

$$I_{\mathrm{e.u.}}(\vec{K}) = 2 + 2\cos(\vec{K}\cdot\vec{r})\,. \tag{11.180}$$

11.10.4 Scattering factor for spherically symmetric atoms

The atomic scattering factor is defined in (11.46) as

$$f_0(\vec{K}) = \int \rho(\vec{r})\, e^{i\vec{K}\cdot\vec{r}}\, dV\,. \tag{11.181}$$

Let us consider a fixed $\vec{K}$ vector directed along the z axis, and let α be the angle between the vector position $\vec{r}$ and $\vec{K}$.

The volume element can be expressed as the product

$$dV \;=\; (dr)\;(2\pi r\sin\alpha)\;(r\,d\alpha) \;=\; 2\pi r^2\,\sin\alpha\,dr\,d\alpha\,, \tag{11.182}$$

where $0 \le r < \infty$ and $0 \le \alpha < \pi$.
The integral (11.181) becomes

$$\begin{aligned} f_0(\vec{K}) \;&=\; 2\pi\int_0^\infty r^2\rho(r)\,dr\int_0^\pi e^{iKr\cos\alpha}\sin\alpha\,d\alpha \\ &=\; 2\pi\int_0^\infty r^2\,\rho(r)\,dr\left[\int_0^\pi \cos(Kr\cos\alpha)\,\sin\alpha\,d\alpha + i\int_0^\pi \sin(Kr\cos\alpha)\,\sin\alpha\,d\alpha\right] \\ &=\; 4\pi\int_0^\infty r^2\,\rho(r)\,\frac{\sin(Kr)}{Kr}\,dr \;=\; f_0(K)\,, \end{aligned}$$

because the second integral in square parentheses is zero.

11.10.5 Potential for electron scattering from atoms

For a spherically symmetric charge distribution, the Poisson equation is conveniently written in spherical coordinates.
The Laplacian in spherical coordinates is

$$\nabla^2 \;=\; \frac{\partial^2}{\partial r^2} \;+\; \frac{2}{r}\frac{\partial}{\partial r} \;+\; \frac{1}{r^2}\frac{\partial^2}{\partial\phi^2} \;+\; \frac{1}{r^2\tan\theta}\,\frac{\partial}{\partial\phi} \;+\; \frac{1}{r^2\sin^2\phi}\frac{\partial^2}{\partial\theta^2} \tag{11.183}$$

For spherical symmetry, the potential Φ doesn't depend on the angles θ and ϕ, so that the Poisson equation is

$$\frac{\partial^2\Phi(r)}{\partial r^2} \;+\; \frac{2}{r}\,\frac{\partial\Phi(r)}{\partial r} \;=\; -\frac{1}{\epsilon_0}\,[\rho_+(r) - \rho_-(r)]\;, \tag{11.184}$$

where $\rho(r)$ is the charge density per unit volume.
We assume the boundary condition

$$\Phi(r) \to 0 \quad \text{for} \quad r \to \infty\,. \tag{11.185}$$

The total potential is the sum of the contributions from the positive and negative charges:

$$\Phi(r) \;=\; \Phi_+(r) + \Phi_-(r) \tag{11.186}$$

and the potential energy of the scattered electron is $E_p(r) = -e\Phi(r)$.

Contribution of the positive charge

The density of positive nuclear charge is

$$\rho_+(r) \;=\; +Ze\,\delta(r)\,. \tag{11.187}$$

The simplest procedure to calculate Φ is to first find the electric field $\mathcal{E}(r)$ (Coulomb law) the integrate over r, taking into account the boundary condition (11.185):

$$\Phi_+(r) \;=\; \frac{1}{4\pi\epsilon_0}\,\frac{Ze}{r}\,. \tag{11.188}$$

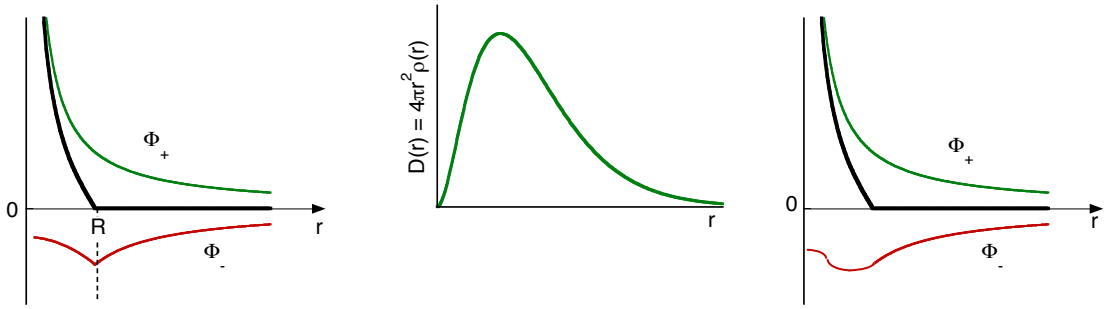

Figure 11.18: Left: potentials generated by a point-like positive charge (green) and by a spherical homogeneous distribution of negative charge or radius R (red); the sum is the black thick line. Centre: radial distribution of negative charge $D(r)$. Right: corresponding realistic potentials. For a scattered electron, the potential energy is $E_p = -e\Phi$.

Ideal case: homogeneous distribution of negative charge

Let us consider a homogeneous negative distribution inside a sphere of radius R:

$$\rho_-(r) \;=\; \begin{cases} -\rho & \text{if } r < R \\ 0 & \text{if } r > R \end{cases} \tag{11.189}$$

with the normalisation condition

$$-\,4\pi \int_o^R \rho\, r^2\, dr \;=\; -Ze\,. \tag{11.190}$$

The potential is (Fig. 11.18, left)

$$\Phi_-(r) \;=\; \begin{cases} -(\rho/6\epsilon_0)\, r^2 - \rho R^2/6\epsilon_0 & \text{if } r < R \\ -(1/4\pi\epsilon_0)\,(Ze/r) & \text{if } r > R \end{cases} \tag{11.191}$$

Realistic case of the distribution of negative charge in atoms

Let us introduce the radial distribution of negative charge (Fig. 11.18, centre)

$$D(r) \;=\; 4\pi r^2\, \rho_-(r)\,. \tag{11.192}$$

The corresponding potential is (Fig. 11.18, right)

$$\Phi_-(r) \;=\; \frac{1}{4\pi\epsilon_0} \left[\frac{\int_0^r D(r')\, dr'}{r'} \;+\; \int_0^\infty \frac{D(r')}{r'}\, dr' \right]$$

11.10.6 Derivation of equation (11.91) (Debye formula)

Let us consider a fixed direction of $\vec{K}$, e.g. along the z axis, and let the vector $\vec{R}_{mn}$ be randomly oriented with respect to $\vec{K}$. The angle α between the directions of $\vec{K}$ and $\vec{R}_{mn}$ is distributed within the interval $0 \le \alpha < \pi$ with probability density $P(\alpha)$ given by the ratio between the solid angle subtended by α and the total solid angle 4π:

$$P(\alpha)\, d\alpha \;=\; (1/4\pi)\, 2\pi\, \sin\alpha\, d\alpha \;=\; (1/2)\, \sin\alpha\, d\alpha\,. \tag{11.193}$$

One can calculate the average

$$\left\langle e^{i\vec{K}\cdot\vec{R}} \right\rangle = \int_0^\pi \cos[KR\cos\alpha]\,\frac{1}{2}\,\sin\alpha\,d\alpha + i\int_0^\pi \sin[KR\cos\alpha]\,\frac{1}{2}\,\sin\alpha\,d\alpha\,. \tag{11.194}$$

The second integral is zero, and the first one is

$$\int_0^\pi \cos[KR\cos\alpha]\,\frac{1}{2}\,\sin\alpha\,d\alpha = \frac{\sin(KR)}{KR}\,. \tag{11.195}$$

11.10.7 Derivation of the Laue interference function

The scattering amplitude factor due to the lattice is the Fourier transform of the lattice function. For a finite crystal (see § 3.1 and 3.2)

$$\begin{aligned} A^{\rm lat}_{\rm e.u.}(\vec{K}) &= \mathcal{F}\left[L(\vec{r})\right] = \sum\nolimits_{\vec{T}} e^{i\vec{K}\cdot\vec{T}} \\ &= \sum_{n_1=0}^{N_1-1} e^{in_1\vec{K}\cdot\vec{a}_1} \times \sum_{n_2=0}^{N_2-1} e^{in_2\vec{K}\cdot\vec{a}_2} \times \sum_{n_3=0}^{N_3-1} e^{in_3\vec{K}\cdot\vec{a}_3} \\ &= \frac{e^{iN_1\vec{K}\cdot\vec{a}_1}-1}{e^{i\vec{K}\cdot\vec{a}_1}-1} \times \frac{e^{iN_2\vec{K}\cdot\vec{a}_2}-1}{e^{i\vec{K}\cdot\vec{a}_2}-1} \times \frac{e^{iN_3\vec{K}\cdot\vec{a}_3}-1}{e^{i\vec{K}\cdot\vec{a}_3}-1}\,. \end{aligned} \tag{11.196}$$

The last equality is justified by

$$\sum_{k=0}^{n-1} = \frac{a^n-1}{a-1}\,. \tag{11.197}$$

The corresponding intensity is

$$\begin{aligned} I_{\rm e.u.}(\vec{K}) &= A^{\rm lat}_{\rm e.u.}(\vec{K})\left[A^{\rm lat}_{\rm e.u.}(\vec{K})\right]^* \\ &= \left(\frac{e^{iN_1\vec{K}\cdot\vec{a}_1}-1}{e^{i\vec{K}\cdot\vec{a}_1}-1}\right)\left(\frac{e^{-iN_1\vec{K}\cdot\vec{a}_1}-1}{e^{-i\vec{K}\cdot\vec{a}_1}-1}\right) \times \ldots \times \ldots \\ &= \frac{\sin^2(N_1\vec{K}\cdot\vec{a}_1/2)}{\sin^2(\vec{K}\cdot\vec{a}_1/2)} \times \ldots \times \ldots \end{aligned} \tag{11.198}$$

where only the factors involving the primitive vector $\vec{a}_1$ have been written, the symbol ... staying for the similar factors involving $\vec{a}_2$ and $\vec{a}_3$.
The last equality is justified by

$$(e^{ix}-1)(e^{-ix}) = 2(1-\cos x) = 2\sin^2(x/2)\,. \tag{11.199}$$

11.10.8 Average of $\exp(ix)$

We want to demonstrate that

$$\langle e^{ix}\rangle = e^{-\langle x^2\rangle/2}\,. \tag{11.200}$$

For small values of x

From the power expansion of the exponential, for sufficiently small values of x

$$\begin{aligned} \langle e^{ix}\rangle &= 1 + i\langle x\rangle - \frac{1}{2}\langle x^2\rangle - \frac{1}{6}\langle x^3\rangle + \ldots \\ &\simeq 1 - \frac{1}{2}\langle x^2\rangle \simeq e^{-\langle x^2\rangle/2}\,. \end{aligned} \tag{11.201}$$

The approximation $x \ll 1$ is generally not suitable for the case considered in Section 11.7, where $x = \vec{K}\cdot(\vec{u}_m - \vec{u}_n)$. We have then to rely on the following most general demonstration.

General demonstration for a gaussian distribution

Let us assume that x is normally distributed around the mean value $x = 0$:

$$\rho(x) = \frac{\alpha}{\sqrt{2\pi}}\, e^{-\alpha^2 x^2}\,. \tag{11.202}$$

From the symmetry of the normal distribution and the parity properties of trigonometric functions

$$\langle e^{ix} \rangle \;=\; \langle \cos x \rangle + i\, \langle \sin x \rangle \;=\; \langle \cos x \rangle\,. \tag{11.203}$$

We can thus evaluate $\langle e^{ix} \rangle$ as

$$\langle e^{ix} \rangle \;=\; \int_{-\infty}^{+\infty} \frac{\alpha}{\sqrt{\pi}}\, e^{-\alpha^2 x^2}\, \cos x\, dx \;=\; e^{-1/4\alpha^2}\,. \tag{11.204}$$

On the other hand, it is well known that

$$\langle x^2 \rangle \;=\; \int_{-\infty}^{+\infty} \frac{\alpha}{\sqrt{\pi}}\, e^{-\alpha^2 x^2}\, x^2\, dx \;=\; \frac{1}{2\alpha^2}\,. \tag{11.205}$$

By substituting α from (11.205) into (11.204) one finds (11.200).

11.11 Bibliography of Chapter 11

- N.W. Aschroft and N.D. Mermin: *Solid State Physics* (various editions). Chapters 5 (recirpocal lattice) and 6 (X-ray diffraction).
- C. Kittel: *Introduction to Solid State Physics*, 8th edition, Wiley 2005. Chapter 2 (diffraction and reciprocal lattice).
- B.K. Vainshstein: *Modern Crystallography*, vol. 1, Springer 1981. Chapter (introductory treatment for X-rays, neutrons and electrons scattering).
- Sow-Hsin Chen and M. Kotlarchyk: *Interactions of photons and neutrons with matter*, World Scientific 2007 (a modern textbook).
- B.E. Warren: *X-ray diffraction*, 1969 (a classical textbook on diffraction).
- J. Als-Nielsen and D. McMorrow: *Elements of modern X-ray Physics*, 2nd ed., Wiley 2011. Chapters 4 and 5 (modern presentation of diffraction with recent examples). Chapter 7 (X-ray absorption).
- S. Mobilio, F. Boscherini, C. Meneghini eds.: *Synchrotron Radiation: basics, methods and applications*, Springer 2015. Chapter 4 (Introduction to matter radiation interaction, by S. Mobilio). Chapter 6 (Introduction to X-ray abdorption spectroscopy, by P. Fornasini). Chapter 8 (X-ray diffraction by crystalline materials, by D. Viterbo and G. Zanotti).

Appendix A

Physical constants and conversion factors

Table A.1: Values of some fundamental constants. The values including uncertainty are taken from the 2014 compilation of CODATA [*CODATA Recommended Values of the Fundamental Physical Constants: 2014*, by P. J. Mohr, D.B. Newell and B. N. Taylor, Rev. Mod. Phys. 88, 035009 (2016), Table XXXII].

Constant	*Symbol*	*Value*	*Unit*
Speed of light in vacuum	c	299 792 458	m s^{-1}
Planck constant	h	$6.626\,070\,040(81)\times10^{-34}$	J s
		$4.135\,67\times10^{-15}$	eV s
Planck constant reduced	$\hbar = h/2\pi$	$1.054\,571\,800(13)\times10^{-34}$	J s
		$6.582\,12\times10^{-16}$	eV s
Elementary charge	e	$1.602\,176\,6208(98)\times10^{-19}$	C
Electron mass	m_e	$9.109\,383\,56(11)\times10^{-31}$	kg
		0.510 998/,902(21)	MeV/c^2
Atomic mass unit	u	$1.660\,539\,040(20)\times10^{-27}$	kg
		9.31×10^{8}	eV/c^2
Rydberg constant	R_∞	10 973 731 . 568 508(65)	m^{-1}
Avogadro constant	N_A	$6.022\,140\,857(74)\times10^{23}$	mol^{-1}
Boltzmann constant	k_B	$1.380\,648\,52(79)\times10^{-23}$	J K^{-1}
		$8.617\,342(15)\times10^{-5}$	eV K^{-1}
Newtonian constant of gravitation	G	$6.674\,08(31)\times10^{-11}$	$m^3kg^{-1}s^{-2}$

Table A.2: Conversions.

Energy	electronvolt	1 eV	$= 1.602\,176\,5\times10^{-19}$ J
	Rydberg	1 Ry	$\simeq$ 13.6 eV
	Hartree	1 Hartree	$\simeq$ 27.2 eV
Wavenumber	Kayser (old)	1 cm^{-1}	$= 1.23984\times10^{-4}$ eV
Electric dipole moment	Debye	1 D	$\simeq 3.336\times10^{-30}$ C m