CHAPTER 1

Point Defects

What is a point defect?
What is the "effective charge" on a defect?
What is an antisite defect?

1.1 INTRODUCTION

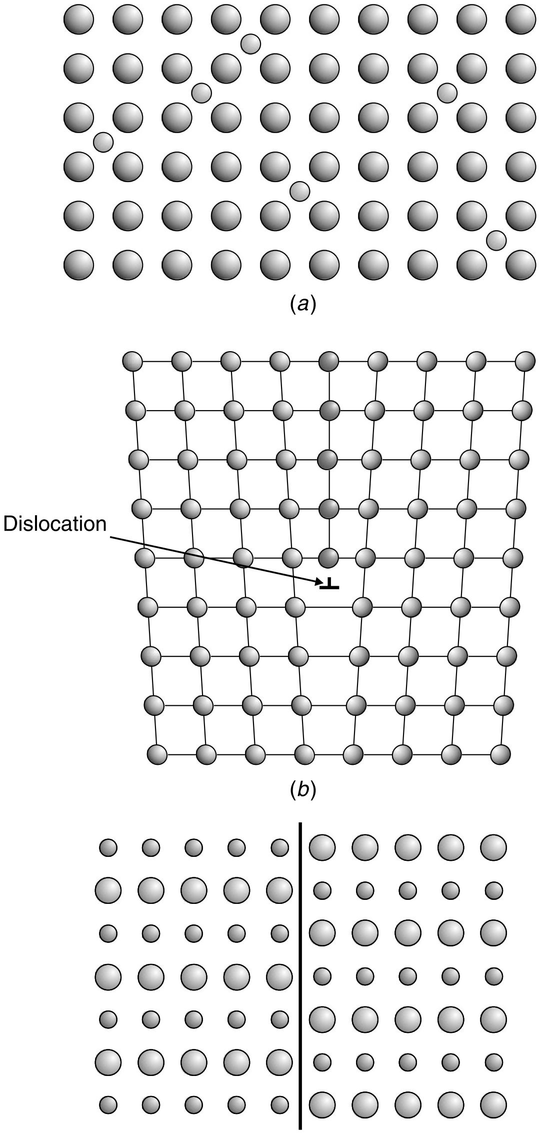
Defects play an important part in both the chemical and physical behavior of solids, and much of modern science and technology centers upon the exploitation or suppression of the properties that defects confer upon a solid. Batteries, fuel cells, displays, data storage, and computer memories all directly utilize, or have evolved from, an understanding and manipulation of defects in inorganic materials. This technology has been developed over some 80–100 years and started with the simplest concepts. However, as the effect of defects upon the properties of the solid gradually became appreciated, the concept of a defect has undergone considerable evolution. The simplest notion of a defect in a solid was the idea of a mistake such as a missing atom or an impurity in place of a normal atom. These structurally simple defects are called point defects. Not long after the recognition of point defects, the concept of more complex structural defects, such as linear defects termed dislocations, was invoked to explain the mechanical properties of metals. In the same period it became apparent that planar defects, including surfaces and grain boundaries, and volume defects such as rods, tubes, or precipitates, have important roles to play in influencing the physical and chemical properties of a solid.

Defects can thus be arranged in a dimensional hierarchy (Fig. 1.1*a*–1.1*d*):

1. Zero-dimensional defects—point defects
2. One-dimensional (linear) defects—dislocations
3. Two-dimensional (planar) defects—external and internal surfaces
4. Three-dimensional (volume) defects—point defect clusters, voids, precipitates.

This and the following chapter are concerned with point defects.

Defects in Solids, by Richard J. D. Tilley



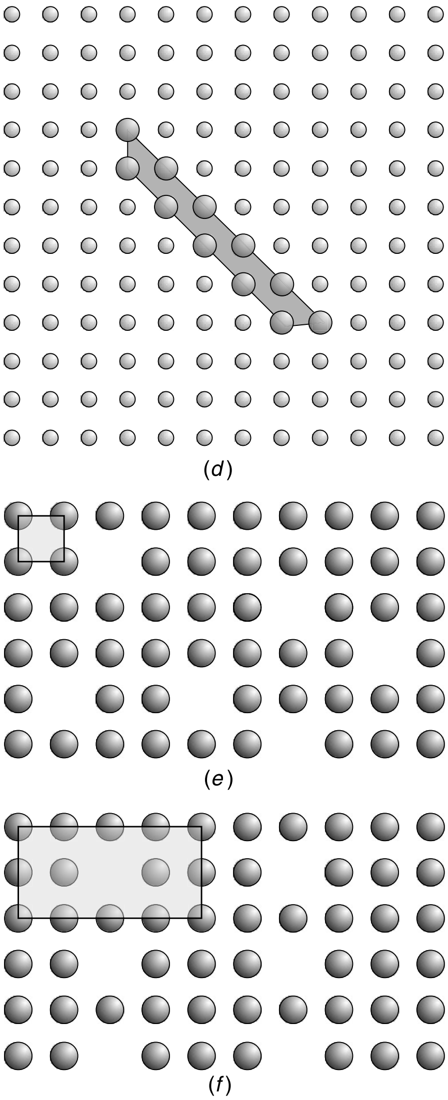
(*a*)

(*b*)

(*c*)

Figure 1.1 Defects in crystalline solids: (*a*) point defects (interstitials); (*b*) a linear defect (edge dislocation); (*c*) a planar defect (antiphase boundary); (*d*) a volume defect (precipitate); (*e*) unit cell (filled) of a structure containing point defects (vacancies); and (*f*) unit cell (filled) of a defect-free structure containing "ordered vacancies."

In addition to the defects listed above, which may be termed structural defects, there are also electronic defects. The first of these are electrons that are in excess of those required for chemical bonding and that, in certain circumstances, constitute charged defects that can carry current. In addition, current in some materials is carried by particles



(d)

(e)

(f)

Figure 1.1 (*Continued*).

that behave rather like positive electrons, and these form the second type of electronic defect.[1] They are called electron holes, positive holes, or more often just holes.

Besides the multiplicity of defects that can be envisaged, there is a wide range of host solid phases within which such defects can reside. The differences between an alloy, a metallic sulfide, a crystalline fluoride, or a silicate glass are significant from

---

[1] These are not the positive equivalent of an electron, a positron, because such a particle would be eliminated instantaneously by combination with an ordinary electron, but are virtual particles equivalent to the absence of an electron. They can be considered to be analogs of a vacancy, which is the absence of an atom.

both a chemical or physical viewpoint. By default, defects have come to be associated with crystalline solids because a crystal has a regular repetition of atoms throughout its volume.[2] A disturbance of this regularity then constitutes the defect (Fig. 1.1*a*–1.1*d*). In this sense, the nature of a defect in, for example, a glass or amorphous polymer, is more difficult to picture.

When circumstances permit, defects can order. This may happen at low temperatures, for example, when defect interactions overcome the disordering effect due to temperature. From a diffraction point of view the ordered sample no longer contains defects, as the characteristic signs disappear. For example, in X-ray analysis broadened diffraction spots and diffuse scattering, both characteristic of disordered crystals, disappear, to be replaced by additional sharp reflections. In such cases the defects are incorporated or assimilated into the crystal structure and so effectively vanish. The original unit cell of the defect-containing phase is replaced by a new unit cell in which the original "defects" are now integral components of the structure (Fig. 1.1*e* and 1.1*f*).

1.2 POINT AND ELECTRONIC DEFECTS IN CRYSTALLINE SOLIDS

The simplest localized defect in a crystal is a mistake at a single atom site in a pure monatomic crystal, such as silicon or iron. Such a defect is called a point defect. Two different types of simple point defect can occur in a pure crystal of an element, M. An atom can be imagined to be absent from a normally occupied position, to leave a vacancy, given the symbol V_M (Fig. 1.2). This may occur, for instance, during crystal growth. A second defect can also be envisaged; namely an extra atom incorporated into the structure, again, say, during crystallization. This extra atom is forced to take up a position in the crystal that is not a normally occupied site: it is called an interstitial site, and the atom is called an interstitial atom (or more often simply an interstitial), given the symbol M_i, (Fig. 1.2). If it is necessary to stress that the interstitial atom is the same as the normal atoms in the structure, it is called a self-interstitial atom.

The various point defects present in a crystal as grown, for example, vacancies, interstitials, excess electrons, holes, and other arrangements, are called *native defects*. These native defects can arise in several ways. They can be introduced during crystal growth, as mentioned above, especially if this is rapid and crystallization is far removed from equilibrium. Defects can also form after the crystal is developed if the crystal is subjected to irradiation by high-energy particles or other forms of radiation. Defects that result from such processes are called *induced defects*. When a pure crystal is heated at a moderate temperature for a long period, a process called annealing, the number of native defects will gradually change. However, no matter how long the sample is annealed, a population of point defects will always remain, even in the purest crystal. These point defects are in thermodynamic equilibrium

---

[2] An introduction to crystal structures and nomenclature is given in the Supplementary Material Section S1.

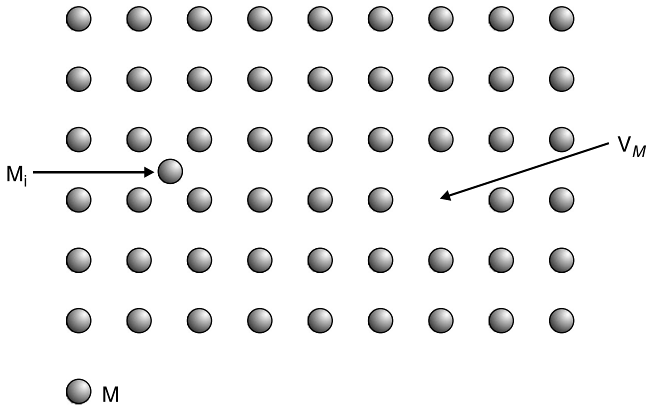


Figure 1.2 Point defects in a pure monatomic crystal of an element M, a vacancy, V_M, and a self-interstitial, M_i.

(Chapter 2) and cannot be eliminated from the solid. They are called *intrinsic* point defects. This residual population is also temperature dependent, and, as treated later (Chapter 2), heating at progressively higher temperature increases the number of defects present.

As well as these intrinsic structural defect populations, electronic defects (excess electrons and holes) will always be found. These are also intrinsic defects and are present even in the purest material. When the equilibria among defects are considered, it is necessary to include both structural and electronic defects.

Turning to pure *compounds*, such as CaO, $MgAl_2O_4$, or FeS, the same intrinsic defects as described above can occur, but in these cases there is more than one set of atoms that can be affected. For example, in a crystal of formula MX, vacancies might occur on metal atom positions, written V_M, or on nonmetal atom positions, given the symbol V_X, or both. Similarly, it is possible to imagine that interstitial metal atoms, written M_i, or nonmetal atoms, written X_i, might occur (Fig. 1.3). The different sets of atom types are frequently called a sublattice, so that one might speak of vacancies on the metal sublattice or on the nonmetal sublattice.

No material is completely pure, and some foreign atoms will invariably be present. If these are undesirable or accidental, they are termed impurities, but if they have been added deliberately, to change the properties of the material on purpose, they are called dopant atoms. Impurities can form point defects when present in low concentrations, the simplest of which are analogs of vacancies and interstitials. For example, an impurity atom A in a crystal of a metal M can occupy atom sites normally occupied by the parent atoms, to form substitutional point defects, written A_M, or can occupy interstitial sites, to form interstitial point defects, written A_i (Fig. 1.4). The doping of aluminum into silicon creates substitutional point defects as the aluminum atoms occupy sites normally filled by silicon atoms. In compounds, the impurities can affect one or all sublattices. For instance, natural sodium chloride often contains

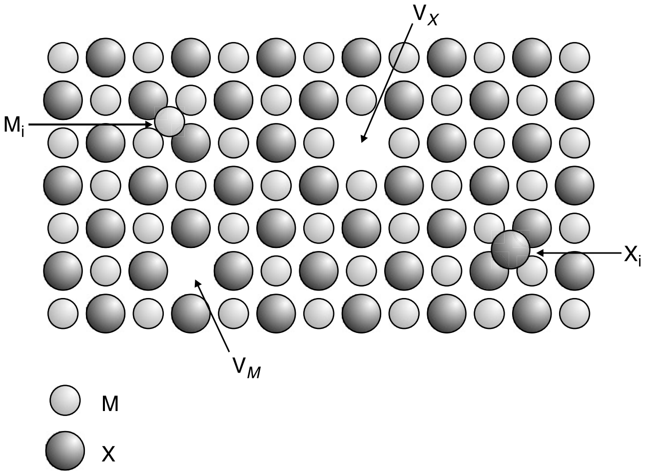


Figure 1.3 Point defects in a crystal of a pure compound, MX, V_M, a metal vacancy; V_X a nonmetal vacancy; M_i, a metal (self-)interstitial; and X_i a nonmetal (self-)interstitial.

potassium impurities as substitutional defects on sites normally occupied by sodium ions, written K_{Na}, that is, the impurities are associated with the metal sublattice. Impurities are called extrinsic defects. In principle, extrinsic defects can be removed by careful processing, but in practice this is very difficult to achieve completely.

Impurities can carry a charge relative to the host structure, as, for example, with a Ca^{2+} ion substituted on a Na^+ site in NaCl or F^- substituted for O^{2-} in CaO. In essence, this means that the impurity carries a different chemical valence, that is,

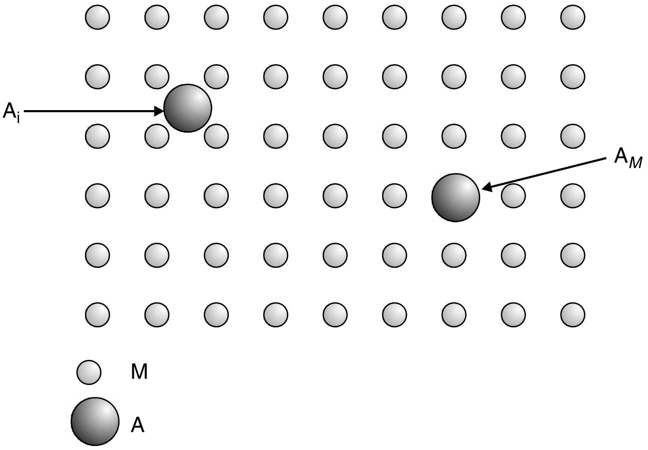


Figure 1.4 Impurity or dopant (A) point defects in a crystal of material M, substitutional, A_M; interstitial, A_i.

it comes from a different group of the periodic table than the host atoms. Deliberate introduction of such impurities is called aliovalent or altervalent doping. The introduction of charged impurities will upset the charge neutrality of the solid, and this must be balanced in some way so as to restore the electronic neutrality of the crystal. One way in which this compensation can be achieved is by the incorporation of other vacancies or interstitial atoms that carry a balancing charge. The substitution of a Ca^{2+} ion for Na^{+} in NaCl can be balanced, for example, by the introduction at the same time, of a sodium ion vacancy, V_{Na}, at the same time. This mechanism is called ionic compensation, structural compensation, or less often self-compensation. However, compensation can also take place electronically by the introduction of appropriate numbers of electrons or holes. For example, the substitution of F^{-} for O^{2-} in CaO could be balanced by the introduction of an equivalent number of electrons, one per added F^{-}. Impurities that are compensated by excess electrons in this way are called donors or donor impurities, while those that are compensated by addition of holes are called acceptors or acceptor impurities. One consequence of these alternatives is that the electrons and holes present in the solid must be included in the overall accounting system used for assessing the defects present in a solid.

The importance of point defects in a crystal cannot be overstated. They can change the physical properties of a solid significantly. To introduce the range of changes possible, Sections 1.3–1.6 outline some of the physical properties that are influenced in this way.

1.3 ELECTRONIC PROPERTIES: DOPED SILICON AND GERMANIUM AS EXAMPLES

Silicon, Si, lies at the heart of most current electronic devices. Both silicon and the similar semiconductor element germanium, Ge, crystallize with the diamond structure (Fig. 1.5). In this structure each atom is surrounded by four others arranged at the corners of a tetrahedron. Each atom has four outer electrons available for chemical bonding, and these are completely taken up by creating a network of tetrahedrally oriented sp^3-hybrid bonds. Although pure silicon and germanium are intrinsic semiconductors, these properties are inadequate for the creation of sophisticated electronic materials for which selected impurity doping is necessary. The resulting materials are called extrinsic semiconductors because of the extrinsic nature of the defects that give rise to the important conductivity changes.

Impurity atoms will upset the orderly arrangement of bonding electrons and this changes the electronic properties. Doping with a very small amount of an impurity from the next *higher* neighboring group of the periodic table, phosphorus (P), arsenic (As), or antimony (Sb), results in the formation of substitutional point defects in which the impurities occupy normal sites, for example, a phosphorus atom on a site normally occupied by a silicon atom, P_{Si} (Fig. 1.6*a*). Each of these atoms has five valence electrons available for bonding, and after using four to form the four sp^3-hybrid bonds, one electron per impurity atom is left over. These electrons are easily liberated from the impurity atoms by thermal energy and are

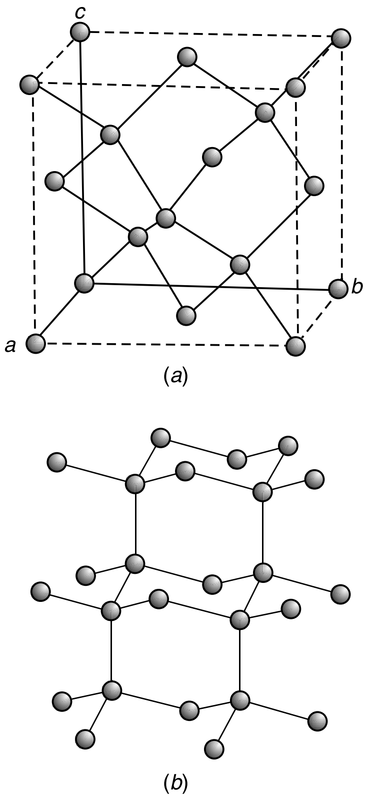


Figure 1.5 Diamond structure: (*a*) unit cell and (*b*) viewed with [111] vertical.

then (more or less) free to move through the crystal under the influence of an external electric field. Each dopant atom contributes one extra electron to the crystal. They are defects, and are given the symbol e' (the superscript indicates a single negative charge relative to the surroundings) to differentiate them from ordinary electrons. The doped material conducts mainly using these electrons, and, as they are negatively charged, the solid is called is called an *n*-type semiconductor. The atoms P, As, or Sb in silicon or germanium are called donors as they donate extra electrons to the crystal. In terms of band theory they are said to occupy states in the conduction band (Supplementary Material S2), the donors themselves being represented by localized energy levels just below the conduction band in the crystal (Fig. 1.6*b*).

An analogous situation arises on doping silicon or germanium with elements from the next *lower* neighboring periodic table group, aluminum (Al), gallium (Ga), and indium (In). The impurity atoms again form substitutional defects, such as $\mathrm{Al_{Si}}$ (Fig. 1.6*c*). In this case the impurities have only three outer bonding electrons available, which are not sufficient to complete four bonds to the surrounding atoms. One bond is an electron short. It simplifies understanding if the missing electron is regarded as a hole, represented by the symbol $h^{\bullet}$. The superscript indicates that the hole carries a

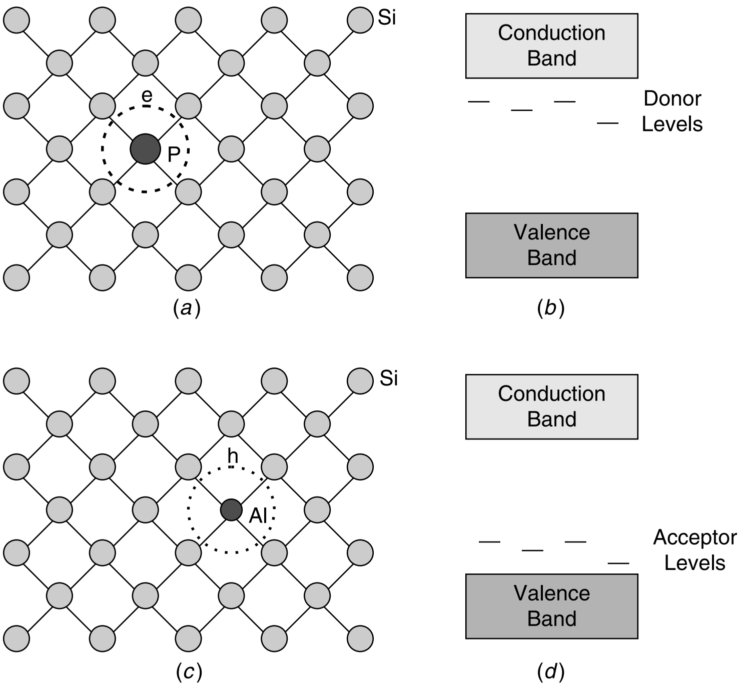


Figure 1.6 (*a*) Donor impurity (P_{Si}) in a silicon crystal. (*b*) Donor energy levels below the conduction band. (*c*) Acceptor impurity (Al_{Si}) in a silicon crystal. (*d*) Acceptor energy levels above the valence band.

positive charge relative to the surroundings. Each impurity atom introduces one positive hole into the array of bonds within the crystal. Thermal energy is sufficient to allow the holes to leave the impurity atom, and these can also move quite freely throughout the crystal. The doped material conducts mainly using these holes, and, as they are positively charged, the solid is called a *p*-type semiconductor. The impurities are termed acceptors because they can be thought of as accepting electrons from the otherwise full valence band. On an energy band diagram the acceptors are represented by normally vacant energy levels just above the top of the valence band (Fig. 1.6*d*), which become occupied by electrons from that band on thermal excitation. The holes thereby created in the valence band provide the means for conductivity.

The idea of a hole is widely used throughout electronics, but its exact definition varies with circumstances. It is frequently convenient to think of a hole as a real particle able to move throughout the crystal, a sort of positive electron. It is equally possible to think of a hole as an electron missing from a localized site such as an atom or a covalent bond. Thus a metal *cation* M^{2+} could be regarded as a metal *atom* plus two localized holes ($M + 2\,h^{\bullet}$). In band theory, a hole is generally regarded as an electron missing from the top of the valence band. All of these designations are more or less equivalent, and the most convenient of them will be adopted in any particular case.

The unique electronic properties of semiconductor devices arise at the regions where p-type and n-type materials are in close proximity, as in p–n junctions. Typical impurity levels are about 0.0001 at %, and their inclusion and distribution need to be very strictly controlled during preparation. Without these deliberately introduced point defects, semiconductor devices of the type now commonly available would not be possible.

1.4 OPTICAL PROPERTIES: F CENTERS AND RUBY AS EXAMPLES

Point defects can have a profound effect upon the optical properties of solids. The most important of these in everyday life is color,[3] and the transformation of transparent ionic solids into richly colored materials by F centers, described below, provided one of the first demonstrations of the existence of point defects in solids.

Defects can impart color to an otherwise transparent solid if they interact with white light. In general the interaction between a defect and the incident light is via electrons or holes. These may pick up some specific frequency of the incident illumination and in so doing are excited from the low-energy ground state, E_0, to one or more higher energy excited states E_1, E_2, E_3 (Fig. 1.7*a*). The light that leaves the crystal is thus depleted in some frequencies and appears a different color to the incident light. The reverse also happens. When electrons drop from the excited states to the ground state E_0, they release this energy and the same light frequencies will be emitted (Fig. 1.7*b*). The relationship between the energy gained or lost, ΔE_n, and the frequency, ν, or the wavelength, λ, of the light absorbed or emitted, is

$$E_n - E_0 = \Delta E_n = h\nu_n = \frac{hc}{\lambda_n}$$

where E_n is the energy of the higher energy level ($n = 1, 2, 3, \ldots$), h is Planck's constant, and c is the speed of light.

The first experiments that connected color with defects were carried out in the 1920s and 1930s by Pohl, who studied synthetic alkali halide crystals. A number of ways were discovered by which the colorless starting materials could be made to display intense colors. These included irradiation by X rays, electrolysis (with color moving into the crystal from the cathode), or heating the crystals at high temperatures in the vapor of an alkali metal. Pohl was a strict empiricist who did not openly speculate upon the mechanics of color formation, which he simply attributed to the presence of Farbzentren (lit. color centers), later abbreviated to F centers.

Leading theoreticians were, however, attracted to the phenomenon and soon suggested models for F centers. In 1930 Frenkel suggested that an F center was an electron trapped in a distorted region of crystal structure, an idea that was incorrect in this instance but led directly to development of the concepts of excitons and

---

[3] Color is the name given to the perception of radiation in the electromagnetic spectrum with a wavelength of between 400 and 700 nm, for an average eye. If all wavelengths in this range are present in a distribution similar to that of radiation from the sun, the light is called white light.

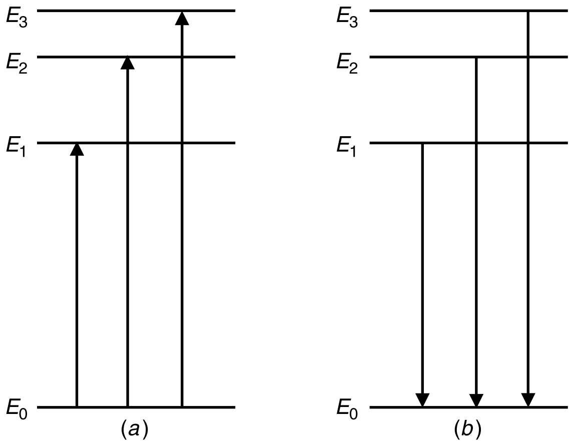


Figure 1.7 (*a*) Light is absorbed when a defect is excited from the ground state, E_0, to a higher energy state, E_1, E_2, or E_3. (*b*) Light is emitted when a defect drops in energy from a higher energy state E_1, E_2, or E_3 to the ground state, E_0.

polarons (Section 7.2). In 1934 Schottky suggested that an F center was an anion vacancy together with a trapped electron, and this model was put onto a firm quantum mechanical footing by Gurney and Mott in 1937 and later by Seitz, who extended the work to other types of color centers.

The origin of the color is as follows. The electron trapped at an anion vacancy in an alkali halide crystal is an analog of a hydrogen atom. The electron can occupy one of a number of orbitals, and transitions between some of these levels absorb light and hence endow the solid with a characteristic color. F centers and related defects are discussed further in Chapter 9.

Color can also be induced into colorless crystals by the incorporation of impurity atoms. The mineral corundum, α-Al_2O_3, is a colorless solid. Rubies are crystals of Al_2O_3 containing atomically dispersed traces of Cr_2O_3 impurity. The formula of the crystal can be written $(Cr_xAl_{1-x})_2O_3$. In the solid the Al^{3+} and Cr^{3+} cations randomly occupy sites between the oxygen ions, so that the Cr^{3+} cations are impurity substitutional, Cr_{Al}, defects. When x takes very small values close to 0.005, the crystal is colored a rich “ruby” red.

The Cr_{Al} defects are responsible for the remarkable change. The color production is mediated by the three 3*d* electrons located on the Cr^{3+} defects. When a Cr^{3+} ion is in a vacuum, these all have the same energy. When the ion is inserted into a corundum crystal, the *d* electrons interact with the surrounding oxygen ions. This causes the *d*-electron orbitals to split into two groups, one at a slightly greater energy than the other, a feature called crystal field or ligand field splitting. An important consequence of this is that three *new* energy levels,[4] ^2E, $^4T_2$, and $^4T_1$, are introduced above the ground state, $^4A_2$, which are not present in pure Al_2O_3.

---

[4] Each separate energy level is given a label called a (spectroscopic) term symbol. For the moment it is sufficient to regard these merely as labels. See Further Reading at the end of this chapter for more information.

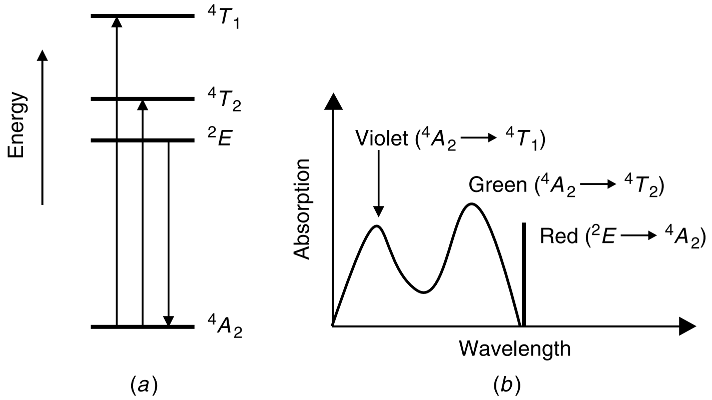


Figure 1.8 Light absorption and emission in ruby: (*a*) energy levels of Cr^{3+} ions and (*b*) absorption spectrum of ruby.

When white light falls onto a crystal of ruby the Cr^{3+} ions selectively absorb some of the radiation and are excited from the ground state $^4A_2$ to energy levels $^4T_2$ and $^4T_1$ (Fig. 1.8*a*). The resulting *absorption spectrum* (Fig. 1.8*b*), consists of two overlapping bell-shaped curves due to these transitions. The absorption curves show that the wavelengths corresponding to violet and green–yellow are strongly absorbed. This means that the color transmitted by the ruby will be red with something of a blue–purple undertone.

At the same time as the excitation is occurring, many of the higher energy Cr^{3+} ions return to the ground state by emitting exactly the same amount of energy as was absorbed, so as to drop back to the ground state from either $^4T_2$ or $^4T_1$. Some ions, however, lose energy to the crystal lattice, warming it slightly, dropping back only to the ^2E energy level. (For quantum mechanical reasons the probability that an ion will pass directly from the ground state to the ^2E state by absorbing energy is low, and so ^2E only becomes filled by this roundabout process.) The ions that return to the ground state from the energy level ^2E emit red light (Fig. 1.8*b*). The color of the best rubies is enhanced by this extra red component. At compositions close to $Cr_{0.005}Al_{0.995}O_3$ this emission can be made to dominate light emission, and the result is laser action (Section 9.9).

The color of rubies (as well as the action of ruby lasers) is thus totally dependent upon substitutional defects.

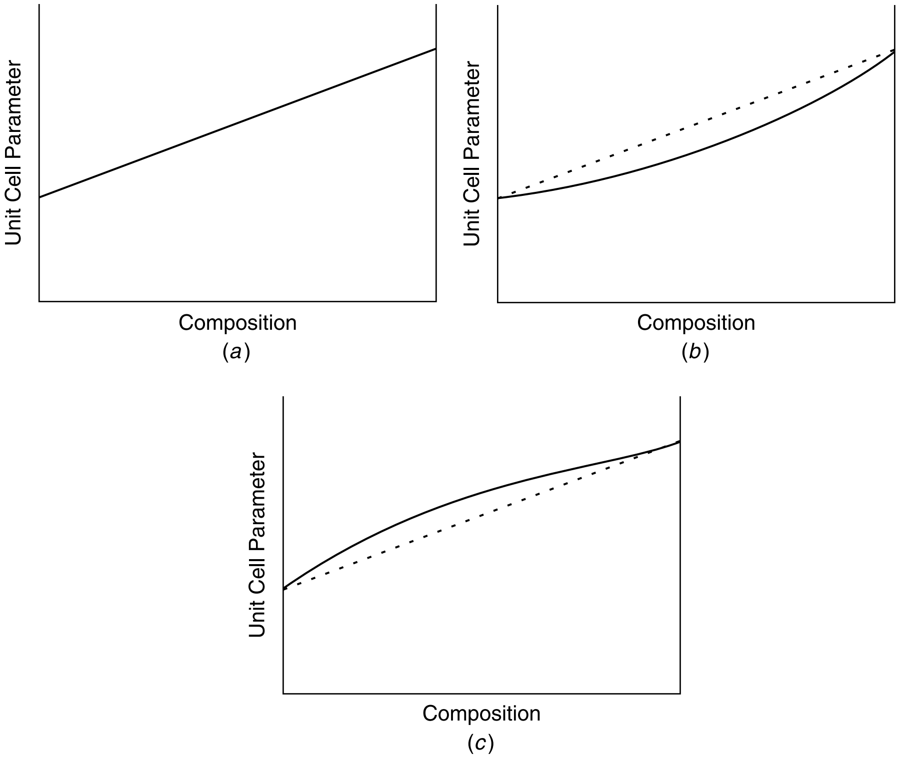
1.5 BULK PROPERTIES

The mechanical consequences of defect populations are less frequently considered than optical or electronic aspects, but they are of importance in many ways, especially when thin films or nanoparticles are considered.

1.5.1 Unit Cell Dimensions

Defects change the dimensions of the unit cell of a crystal. This can be illustrated by reference to ruby, described in the previous section. Rubies represent just a small part of the system spanning the composition range from pure Al_2O_3 to pure Cr_2O_3. The composition of these mixed crystals, $(Cr_xAl_{1-x})_2O_3$, can cover the range $0 < x < 1$, The unit cell parameters of the solid will vary as the composition changes. Vegard's law, first stated in (1921), is an empirical correlation to the effect that the lattice parameters of members of a solid solution formed between two isostructural phases, such as NaCl and KCl, will fall on a straight line joining the lattice parameters of these two parent structures (Fig. 1.9*a*):

$$x = \frac{a_{ss} - a_1}{a_2 - a_1}$$



(*a*) (*b*)

(*c*)

Figure 1.9 Vegard's law relating unit cell parameters to composition for solid solutions and alloys: (*a*) ideal Vegard's law behavior; (*b*) negative deviation from Vegard's law; and (*c*) positive deviation from Vegard's law.

that is,

$$a_{ss} = a_1 + x(a_2 - a_1)$$

where a_1 and a_2 are the lattice parameters of the parent phases, a_{ss} is the lattice parameter of the solid solution, and x is the mole fraction of the parent phase with lattice parameter a_2. (The relationship holds for all unit cell parameters, a, b, and c and any interaxial angles.) This "law" is simply an expression of the idea that the cell parameters are a direct consequence of the sizes of the component atoms in the solid solution. Vegard's law, in its ideal form, is almost never obeyed exactly. A plot of cell parameters that lies below the ideal line (Fig. 1.9*b*) is said to show a *negative* deviation from Vegard's law, and a plot that lies above the ideal line (Fig. 1.9*c*) is said to show a *positive* deviation from Vegard's law. In these cases, atomic interactions, which modify the size effects, are responsible for the deviations. In all cases, a plot of composition versus cell parameters can be used to determine the formulas of intermediate structures in a solid solution.

1.5.2 Density

X-ray diffraction allows the dimensions of the unit cell to be accurately measured. If the structure type of the material is known, the ideal cell contents are also known. Thus, the unit cell of a crystal of composition M_2O_3 that adopts the corundum structure contains 12 M atoms and 18 O atoms (Supplementary Material, S1). This readily allows the theoretical density of a solid to be calculated. The weights of all of the atoms in the cell are added, and this is divided by the cell volume.

It was realized at an early stage that a comparison of the theoretical and measured density of a solid can be used to determine the notional species of point defect present. The general procedure is:

1. Measure the composition of the solid.
2. Measure the density.
3. Measure the unit cell parameters.
4. Calculate the theoretical density for alternative point defect populations.
5. Compare the theoretical and experimental densities to see which point defect model best fits the data.

The method can be illustrated by reference to a classical 1933 study of the defects present in wüstite, iron monoxide. Wüstite adopts the sodium chloride (NaCl) structure, and the unit cell should contain 4 Fe and 4 O atoms in the unit cell, with an ideal composition $FeO_{1.0}$, but in reality the composition is oxygen rich and the unit cell dimensions also vary with composition (Table 1.1). Because there is more oxygen present than iron, the real composition can be obtained by assuming either that there are extra oxygen atoms in the unit cell (interstitial defects) to give a composition FeO_{1+x}, or that there are iron vacancies present, to give a formula $Fe_{1-x}O$. It is

TABLE 1.1 Experimental Data for Wüstite, FeO_x

O/Fe Ratio	Fe/O Ratio	Lattice Parameter/nm	Density/kg m^{-3} Observed	Density/kg m^{-3} Calculated for Interstitial Oxygen	Density/kg m^{-3} Calculated for Iron Vacancies
1.058	0.945	0.4301	5728	6076	5740
1.075	0.930	0.4292	5658	6136	5706
1.087	0.920	0.4285	5624	6181	5687
1.099	0.910	0.4282	5613	6210	5652

Source: Adapted from E. R. Jette and F. Foote, *J. Chem. Phys.*, **1**, 29 (1933).

possible to determine which of these suppositions is correct by comparing the real and theoretical density of the material.

For example, consider the sample specified in the top line of Table 1.1, with an oxygen : iron ratio of 1.058, a measured density of 5728 kg m^{-3} and a cubic lattice parameter, *a*, of 0.4301 nm.

1. Assume that the iron atoms in the crystal are in a perfect array, identical to the metal atoms in the sodium chloride structure, and that the 0.058 excess of oxygen is due to interstitial oxygen atoms being present, over and above those on the normal anion positions. The unit cell of the structure now contains 4 Fe and (4 × 1.058) O. The density is calculated to be 6076 kg m^{-3}.
2. Assume that the oxygen array is perfect and identical to the nonmetal atom array in the sodium chloride structure and that the unit cell contains some vacancies on the iron positions. In this case, one unit cell will contain 4 atoms of oxygen and (4/1.058) atoms of iron, that is, 4 $Fe_{0.945}O$. The density is calculated to 5741 kg m^{-3}.

The difference in the two values is surprisingly large and is well within the accuracy of density determinations. The experimental value is in accord with a model that assumes vacancies on the iron positions, as are all results in Table 1.1, indicating that a formula $Fe_{1-x}O$, in which there are vacancies at some of the Fe positions, better reflects the structure.

Although this analysis is correct, it is a macroscopic method that does not give any true crystallographic information. The oxygen vacancies may be arranged in any number of ways. In fact, it is found that the vacancies form clusters that can be regarded as fragments of the next higher oxide, Fe_3O_4, with the spinel structure embedded in a sodium chloride structure matrix (Section 4.4.2).

Because of Vegard's law, it is clear that the density of a solid solution will also be expected to be a linear function of the densities of the parent phases that make up the limits of a solid solution.

1.5.3 Volume

The ambient pressure surrounding a solid can directly influence the point defects present (Chapter 7). This is most often encountered with respect to oxygen pressure, but water vapor and the partial pressure of volatile metals are also of importance in high-temperature applications. Changes in defect populations have a direct bearing upon unit cell dimensions and hence upon the overall dimensions of the solid. These dimensional changes, like changes in density, can be used to infer the type of point defect present, albeit with a low degree of discrimination. For example, uranium dioxide, UO_2, can gain oxygen to form a hyperstoichiometric phase UO_{2+x}. The partial pressure of oxygen that is in equilibrium with the dioxide, $UO_{2.0}$ is about 10^{-9} Pa, depending upon the temperature. When the oxygen partial pressure surrounding the oxide is increased above this value, the solid gains oxygen and expands. This expansion is not uniform but peaks and then decreases (Fig. 1.10). A possible explanation of this change is that interstitial oxygen ions cause initial expansion. When these reach higher concentrations, clustering might cause the volume to fall. X ray or other structural studies are needed to explore such models further (Section 4.4.3).

Changes in density, unit cell dimensions, and macroscopic volume have serious effects. In an environment where point defects (or aggregates of point defects) are generated, such as in the components of nuclear reactors, or in vessels used for the storage of nuclear waste, where point defects are produced as a result of irradiation, dimensional changes can cause components to seize or rupture.

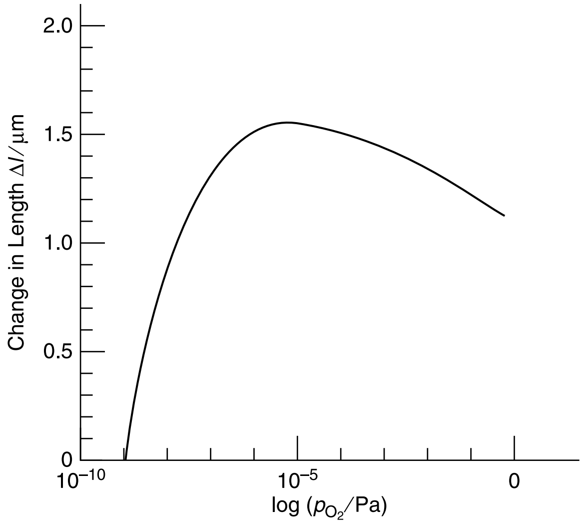


Figure 1.10 Variation of the length of a single crystal of oxygen-rich uranium dioxide, UO_{2+x}, with oxygen partial pressure at 1200°C. [Data redrawn from L. Desgranges, M. Gramond, C. Petot, G. Petot-Ervas, P. Ruello, and B. Saadi, *J. Eur. Ceram. Soc.*, **25**, 2683–2686 (2005).]

Such changes in the defect population can be critical in device manufacture and operation. For example, a thin film of an oxide such as SiO laid down in a vacuum may have a large population of anion vacancy point defects present. Similarly, a film deposited by sputtering in an inert atmosphere may incorporate both vacancies and inert gas interstitial atoms into the structure. When these films are subsequently exposed to different conditions, for example, moist air at high temperatures, changes in the point defect population will result in dimensional changes that can cause the film to buckle or tear.

Ionic conductors, used in electrochemical cells and batteries (Chapter 6), have high point defect populations. Slabs of solid ceramic electrolytes in fuel cells, for instance, often operate under conditions in which one side of the electrolyte is held in oxidizing conditions and the other side in reducing conditions. A significant change in the point defect population over the ceramic can be anticipated in these conditions, which may cause the electrolyte to bow or fracture.

These effects can all be enhanced if the point defects interact to form defect clusters or similar structures, as in $Fe_{1-x}O$ above or UO_{2+x} (Section 4.4). Such clusters can suppress phase changes at low temperatures. Under circumstances in which the clusters dissociate, such as those found in solid oxide fuel cells, the volume change can be considerable, leading to failure of the component.

1.5.4 Young's Modulus (the Elastic Modulus)

The response of a solid to an applied force (a stress, σ) in terms of a change in dimensions (a strain, ε) is given by *Young's modulus, E*:

$$\sigma = E\varepsilon$$

(Hooke's law). [Young's modulus is often called the elastic modulus, but as a number of different (symmetry dependent) elastic moduli are defined for a solid, the generic term, i.e., *the* elastic modulus, is not sufficiently precise.] The value of Young's modulus of a solid containing point defects is noticeably different from a solid that is defect free. This is important is some applications. For example, solid oxide fuel cells (Chapter 6) use oxides such as calcia-stabilized or yttria-stabilized zirconia as the electrolyte. These cells operate at between 700 and 1000°C, and substantial populations of oxygen ion vacancies occur in the electrolyte at these temperatures. Mechanical failure of the cells remains a problem due to the thermal cycling that occurs during operation. It has been found that the Young's modulus of the electrolyte decreases significantly as the number of oxygen ion vacancies increases, exacerbating the mechanical problems encountered.

These examples indicate that it is necessary to keep the possible effect of point defects on bulk and mechanical properties in mind. Although less definitive than electronic and optical properties, they may make the difference in the success or failure of device operation.

1.6 THERMOELECTRIC PROPERTIES: THE SEEBECK COEFFICIENT AS AN EXAMPLE

When the two ends of a material containing mobile charge carriers, holes or electrons, are held at different temperatures, a voltage is produced, a phenomenon called the Seebeck effect (Fig. 1.11). The Seebeck coefficient of a material, α, is defined as the ratio of the electric potential produced when no current flows to the temperature difference present across a material:[5]

$$\alpha = \pm \frac{\phi_H - \phi_C}{T_H - T_C} = \pm \frac{\Delta\phi}{\Delta T}$$

where ϕ_H and ϕ_C are the potentials and T_H and T_C are the temperatures at the hot end and the cold end of the sample, respectively. The main virtue of the Seebeck coefficient in the context of this book is that its sign and magnitude can provide a measure of the concentration of charge carriers, the nature of the charge carriers, and, with some simple assumptions, the number of defects present that give rise to the charge carriers.

In the case of materials that have mobile electrons, that is, *n*-type semiconductors, the colder end of the rod will be *negative* with respect to the hotter end and the sign of the Seebeck coefficient is negative. In the case where the mobile charge carriers are positive holes, that is, *p*-type semiconductors, the colder end of the rod will be *positive* with respect to the hotter end, making the Seebeck coefficient positive. For example, the nonstoichiometric forms of NiO, CoO, and FeO all show positive values for α, indicating that conductivity is by way of holes, whereas nonstoichiometric ZnO has a negative value of α, indicating conductivity by way of electrons. In the case of materials with both types of charge carrier present, the one that is present in greatest numbers dominates the measurement.

The magnitude of the Seebeck coefficient is related to the concentration of mobile charge carriers present and is greatest at low defect concentrations, when other methods of analyzing defect populations give least precision. The relationship between the number of defects and the Seebeck coefficient is obtained by estimating the configurational entropy of the defect-containing material. A number of forms for this estimate are found, each depending upon slightly different approximations (Supplementary Material S3). The most direct is

$$\alpha = \pm \left(\frac{k}{e}\right) \left[\ln\left(\frac{n_0}{n_d}\right)\right]$$

---

[5] The Seebeck coefficient is frequently called the thermoelectric power or thermopower, and labeled Q or S. Neither of these alternatives is a good choice. The units of the Seebeck coefficient are not those of power. The symbol Q is most often used to signify heat transfer in materials. The designation S can easily be confused with the entropy of the mobile charge carriers, which is important because the Seebeck coefficient is equivalent to the entropy per mobile charge carrier (see Supplementary Material S3).

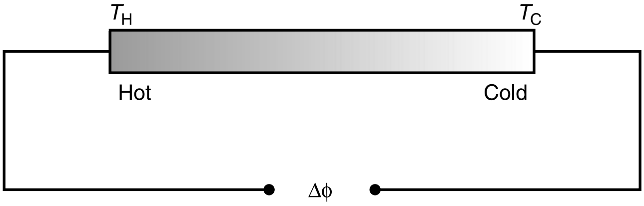


Figure 1.11 Seebeck effect. A sample with one end maintained at a high-temperature T_H and the other at a low-temperature T_C will develop a potential difference $\Delta\phi$.

where n_0 is the number of sites in the sublattice containing defects and n_d is the number of defects giving rise to mobile electrons or holes and $k/e = 86.17\ \mu V\ K^{-1}$. The positive version applies to p-type materials and the negative expression to n-type materials. Note that n_0/n_d increases as the number of defects falls, and so the value of α is expected to be greatest for lowest defect populations.

This equation is formally equivalent to the Heikes equation:

$$\begin{aligned}\alpha &= -\left(\frac{k}{e}\right)\ln\left(\frac{1-c}{c}\right) && \text{for electrons}\\ &= +\left(\frac{k}{e}\right)\ln\left(\frac{1-c}{c}\right) && \text{for holes}^{6}\end{aligned}$$

where c is the fraction of defects (or mobile charge carriers) present (Supplementary Material S3). This form is useful because the value of c is directly related to the composition of the sample. For example, the material $LaNi_xCo_{1-x}O_3$, which has the perovskite structure, can be analyzed in this way. Each Ni^{2+} ion that replaces a Co^{3+} ion in the parent compound $La^{3+}Co^{3+}O_3$ forces one of the other Co^{3+} ions to transform to Co^{4+} to maintain charge neutrality. Each Co^{4+} ion can be considered to be a Co^{3+} ion plus a trapped hole. Electronic conductivity can then be considered to occur by the migration of holes from one Co^{4+} ion to a neighboring Co^{3+} ion:

$$Co^{4+} + Co^{3+} \rightleftharpoons Co^{3+} + Co^{4+}$$

that is,

$$(Co^{3+} + h^{\bullet}) + Co^{3+} \rightleftharpoons Co^{3+} + (Co^{3+} + h^{\bullet})$$

---

[6] Note that the form of the equation for holes is often written

$$\alpha = -\left(\frac{k}{e}\right)\ln\left(\frac{c}{1-c}\right)$$

which is identical to that given as $\ln x = -\ln(1/x)$.

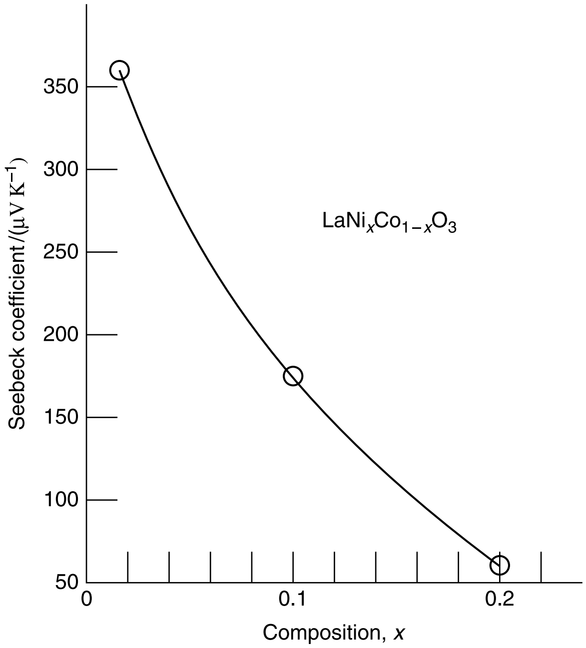


Figure 1.12 Seebeck coefficient of the oxide $LaNi_xCo_{1-x}O_3$ as a function of the composition, x. [Data adapted from R. Robert, L. Becker, M. Trottmann, A. Reller, and A. Weidenkraft, *J. Solid State Chem.*, **179**, 3893–3899, (2006).]

The number of mobile holes is equal to the number of impurity Ni^{2+} ions, and so the fraction c in the Heikes equation is equal to x in $LaNi_xCo_{1-x}O_3$. In accord with the theory, the Seebeck coefficient, α, is positive and greatest at low values of x and decreases as x increase (Fig. 1.12). Substituting a value of $c = 0.02$ into the equation yields a value of $\alpha = +335$ μV K^{-1}, in good agreement with the experimental value of 360 μV K^{-1} (Robert *et al.*, 2006). Note that the above example also shows that an experimentally determined value of the Seebeck coefficient can be used to estimate the concentration of impurity defects in a doped oxide.

1.7 POINT DEFECT NOTATION

Point defect populations profoundly affect both the physical and chemical properties of materials. In order to describe these consequences a simple and self-consistent set of symbols is required. The most widely employed system is *the Kröger–Vink notation*. Using this formalism, it is possible to incorporate defect formation into chemical equations and hence use the powerful methods of chemical thermodynamics to treat defect equilibria.

In the Kröger–Vink notation, empty atom positions, that is, vacancies, are indicated by the symbol V. Acknowledging that V is the chemical symbol for the element vanadium, it is necessary to add that, where confusion may occur, the symbol for a vacancy is written Va. The atom that is absent from a normally occupied site is specified by the normal chemical symbol for the element, written as a subscript. Thus in NiO, for example, the symbol V_O would represent an oxygen atom vacancy and V_{Ni} a nickel atom vacancy.

The position of a defect that has been substituted for another atom in the structure is represented by a subscript that is the chemical symbol of the atom normally found at the site occupied by the defect impurity atom. The impurity is given its normal chemical symbol, and the site occupied is written as a subscript, using the chemical symbol for the atom that normally occupies the site. Thus, an Mg atom on a Ni site in NiO would be written as Mg_{Ni}. The same nomenclature is used if an atom in a crystal occupies the wrong site. For example, antisite defects in GaN would be written as Ga_N and N_{Ga}.

Interstitial positions, positions in a crystal not normally occupied by an atom, are denoted by the subscript i. For example, F_i would represent an interstitial fluorine atom in, say, a crystal of fluorite, CaF_2.

It is possible for one or more lattice defects to associate with one another, that is, to cluster together. These are indicated by enclosing the components of such a cluster in parentheses. As an example, $(V_M V_X)$ would represent a defect in which a vacancy on a metal site and a vacancy on a nonmetal site are associated as a vacancy pair.

1.8 CHARGES ON DEFECTS

One of the most difficult problems when working with defects, especially in ionic crystals, is to decide on the charge on the ions and atoms of importance. The Kröger–Vink notation bypasses the problem of deciding on the *real charges* on defects (z_d), by considering only *effective charges* (q_e). The effective charge on a defect is the charge that the defect has *with respect to the charge that would be present at the same site in a perfect crystal.* Thus, the effective charge is a relative charge. For atomic or ionic species, this is equal to the difference between the real charges on the defect species, z_d, minus the real charge at the site occupied in a perfect crystal, z_s:

$$q_e = z_d - z_s \tag{1.1}$$

To distinguish effective charges from real charges, the superscript $'$ is used for each unit of effective negative charge and the superscript $\bullet$ is used for each unit of effective positive charge. The real charges on a defect are still given the superscript symbols $-$ and $+$.

1.8.1 Electrons and Electron Holes

The charged defects that most readily come to mind are electrons. In a crystal containing defects, some fraction of the electrons may be free to move through the matrix. These are denoted by the symbol e'. The superscript $'$ represents the effective negative charge on the electron (q_e), and it is written in this way to emphasize that it is considered relative to the surroundings rather than as an isolated real point charge. The concentration of electrons that are free to carry current through a crystal is frequently given the symbol n in semiconductor physics.

The counterparts to electrons in semiconducting solids are holes, represented by the symbol $h^{\bullet}$. Each hole will bear an effective positive charge, q_e, of $+1$, which is represented by the superscript $\bullet$ to emphasize that it is considered relative to the surrounding structure. The concentration of holes that are free to carry current through a crystal is often given the symbol p in semiconductor physics.

1.8.2 Atomic and Ionic Defects

Point defects can carry a charge. In ionic crystals, this may be considered to be the normal state of affairs. The effective charge on these defects is the charge that the defect has with respect to the charge present or expected at the same point in the perfect crystal structure. To illustrate this concept, consider the situation in an ionic material such as NaCl, which it is convenient to consider as made up of the charged ions Na^+ and Cl^-. What is the effective charge on a sodium vacancy, V_{Na}, in the NaCl structure? The result is easily obtained using Eq. (1.1). The real charge on the vacancy, z_d, is 0. The real charge at the site in a perfect crystal, z_s, is due to the presence of Na^+, that is, $+1$, hence:

$$q_e = z_d - z_s = (0 - 1) = -1$$

Relative to the normal situation at the site, the vacancy appears to bear an effective negative charge equivalent to -1. Hence, a vacancy at a sodium ion (Na^+) site in NaCl would be written as V'_{Na}.

In general, the absence of a positive ion will leave a vacancy with a negative effective charge relative to the normally occupied site. Multiple effective negative charges can exist and are written using superscript n'. A Ca^{2+} ion vacancy in a crystal of CaO will bear an effective negative charge of $2'$, and the vacancy has the symbol $V_{Ca}^{2'}$ as can be seen from Eq. (1.1):

$$q_e = z_d - z_s = (0 - 2) = -2$$

The same reasoning indicates that the absence of a negative ion will leave a positive effective charge relative to a normal site occupied by a negative ion. A vacancy at a chloride ion (Cl^-) site is positively charged relative to the normal situation prevailing at an anion site in the crystal. Using Eq. (1.1):

$$q_e = z_d - z_s = (0 - (-1)) = +1$$

Hence, the vacancy has an effective charge of $+1$, which would be written $V_{Cl}^{\bullet}$.

In general, the absence of a negative ion will endow a site with a positive effective charge. Multiple effective positive charges can exist and are written using superscript $n\bullet$. An oxide ion (O^{2-}) vacancy in a crystal of CaO will bear an effective positive charge of

$$q_e = z_d - z_s = (0 - (-2)) = +2$$

and the vacancy has the symbol $V_O^{2\bullet}$.

An effective charge relative to the host lattice is possible with any defect. These are added as superscripts to the appropriate symbol: V_M, V_X, M_i, M_X and associated defects such as $(V_M V_X)$.

Interstitial sites, which are normally unoccupied in a crystal, will have no preexisting charge. When an atom or an ion occupies an interstitial site, its real charge is the same as the effective charge. Thus, for a Zn^{2+} ion at an interstitial site, from Eq. (1.1):

$$q_e = z_d - z_s = (2 - 0) = 2$$

The defect is given the symbol $Zn_i^{2\bullet}$.

Substitution of an ion with one valence by another with a different valence, aliovalent substitution, will create a charged defect. For example, a divalent ion such as Ca^{2+} substituted for a monovalent Na^+ on a sodium site in NaCl gives a local electronic charge augmented by one extra positive charge:

$$q_e = z_d - z_s = (2 - 1) = 1$$

The defect has an effective charge of 1 and it is represented by the symbol $Ca_{Na}^{\bullet}$.

Not all defects carry effective charges. Frequently, this need not be noted. For instance, suppose that a sodium ion in NaCl, represented by Na_{Na}, is substituted by a potassium ion, represented by K_{Na}. Clearly, the defect will have no effective charge:

$$q_e = z_d - z_s = (1 - 1) = 0$$

This defect is therefore neutral in terms of effective charge. The same could be said of a neutral lithium atom introduced into an interstitial site in titanium disulfide, TiS_2, which would be written Li_i. However, it is sometimes important to emphasize that the defect is neutral in terms of effective charge. This is made clear by the use of a superscript x. Thus a K^+ ion substituted for a Na^+ ion could be written K_{Na}^{x} when the effective charge situation needs to be specified. Similarly, an interstitial Li atom could be represented as Li_i^{x} to emphasize the lack of an effective charge on the defect when it is essential to do so.

Some defects are termed amphoteric[7] defects. These are defects that can take on either a positive or negative effective charge, depending upon circumstances. For

---

[7] Amphoteric is a chemical term used to describe (mainly) oxides or hydroxides that are able to react both with acids and alkalis. In the present use, it is not related to acid–base properties.

TABLE 1.2 Kröger–Vink Notation for Defects in Crystals[a]

Defect Type	Notation	Defect Type	Notation
Metal vacancy at metal (M) site	V_M	Nonmetal vacancy at nonmetal (Y) site	V_Y
Impurity metal (A) at metal (M) site	A_M	Impurity non-metal (Z) at nonmetal site	Z_Y
Interstitial metal (M)	M_i	Interstitial nonmetal (Y)	Y_i
Neutral metal (M) vacancy	V_M^x	Neutral nonmetal (Y) vacancy	V_Y^x
Metal (M) vacancy with negative effective charge	V_M'	Nonmetal (Y) vacancy with positive effective charge	$V_Y^\bullet$
Interstitial metal (M) with positive effective charge	$M_i^\bullet$	Interstitial nonmetal (X) with negative effective charge	X_i'
Interstitial metal (M) with *n* positive effective charges	$M_i^{n\bullet}$	Interstitial nonmetal (Y) with *n* negative effective charges	$Y_i^{n'}$
Free electron[b]	e'	Free hole[b]	$h^\bullet$
Associated defects (vacancy pair)	$(V_M V_Y)$	Associated defects with positive effective charge	$(V_M V_Y)^\bullet$

[a] The definitive definitions of this nomenclature and further examples are to be found in the *IUPAC Red Book on the Nomenclature of Inorganic Chemistry*, Chapter I.6.

[b] The concentrations of these defects are frequently designated by n and p respectively.

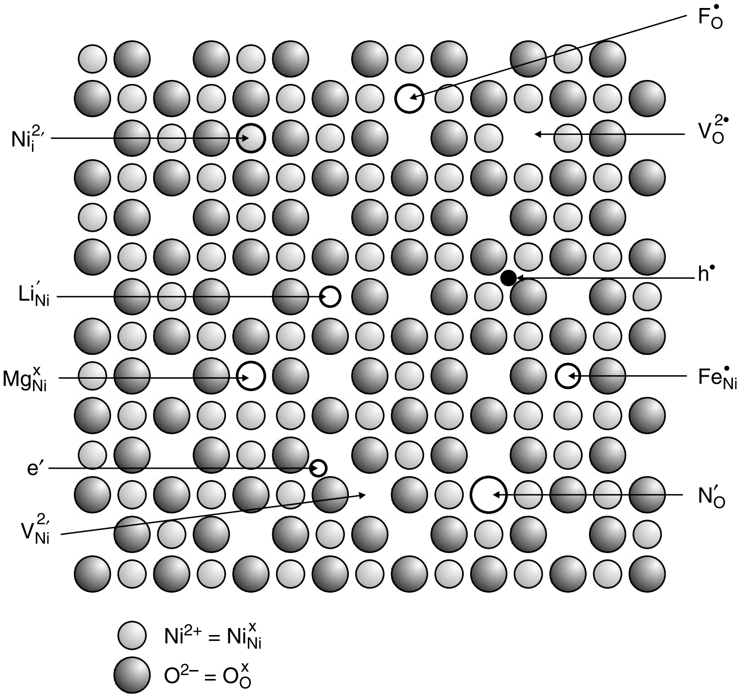


Figure 1.13 Point defects in nickel oxide, NiO (schematic): Ni^{2+} vacancy; Ni^{2+} interstitial; Li^+ on a Ni^{2+} site; Mg^{2+} on a Ni^{2+} site; Fe^{3+} on a Ni^{2+} site; O^{2-} vacancy; N^{3-} on an O^{2-} site; F^- on an O^{2-} site; free electron; free hole.

example, the incorporation of lithium ions, Li^+, into zinc selenide, ZnSe, can lead to lithium interstitials, $Li_i^\bullet$, in which the real charge and the effective charge are identical, as stated above. Alternatively, the ions can form substitution defects in which the Li^+ ions occupy Zn^{2+} sites, in which case the designation is Li'_{Zn}. As the populations of these two alternatives changes, for instance, with lithium concentration, so does the electronic character of the defect population.

The effective charges on an ionic defect can be considered to be linked to the defect by an imaginary bond. If the bond is weak, the effective charge can be liberated, say by thermal energy, so that it becomes free to move in an applied electric field and so contribute to the electronic conductivity of the material. Whether the effective charge on a defect is considered to be strongly associated with the defect or free depends upon the results obtained when the physical properties of the solid are measured.

The main features of the Kröger–Vink notation are summarized in Table 1.2 and are illustrated with respect to point defects in a crystal containing Ni^{2+} and O^{2-} ions in Figure 1.13.

1.9 BALANCED POPULATIONS OF POINT DEFECTS: SCHOTTKY AND FRENKEL DEFECTS

1.9.1 Schottky Defects

Compounds are made up of atoms of more than one chemical element. The point defects that can occur in pure compounds parallel those that occur in monatomic materials, but there is an added complication in this case concerning the composition of the material. In this chapter discussion is confined to the situation in which the composition of the crystal is (virtually) fixed. Such solids are called stoichiometric compounds. (The situations that arise when the composition is allowed to vary are considered in Chapter 4 and throughout much of the rest of this book. This latter type of solid is called a nonstoichiometric compound.) The composition problem can be illustrated with respect to a simple compound such as sodium chloride.

Sodium chloride, also known as rock salt or halite, is composed of equal numbers of sodium (Na) and chlorine (Cl) atoms, has a chemical formula NaCl and a simple structure (Fig. 1.14). It is a good first approximation to regard this material as being composed of ions. The introduction of vacancies on the cation sublattice will upset both the composition and the charge balance. If x such vacancies occur, the formula of the crystal will now be $Na_{1-x}Cl$, and the overall material will have an excess negative charge of $x-$ because the number of chloride ions is greater than the number of sodium ions by this amount. The compound should be written $[Na_{1-x}Cl]^{x-}$. The same will be true for the anion sublattice. If x vacancies are placed on the anion sublattice, the material will take on an overall positive charge because the number of sodium ions now outnumbers the chlorine ions, and the formula becomes $[NaCl_{1-x}]^{x+}$. Ordinary crystals of sodium chloride do not show an overall negative or positive charge or have a formula different to NaCl. Thus, if vacancy defects occur in these crystals, the numbers on the anion and cation

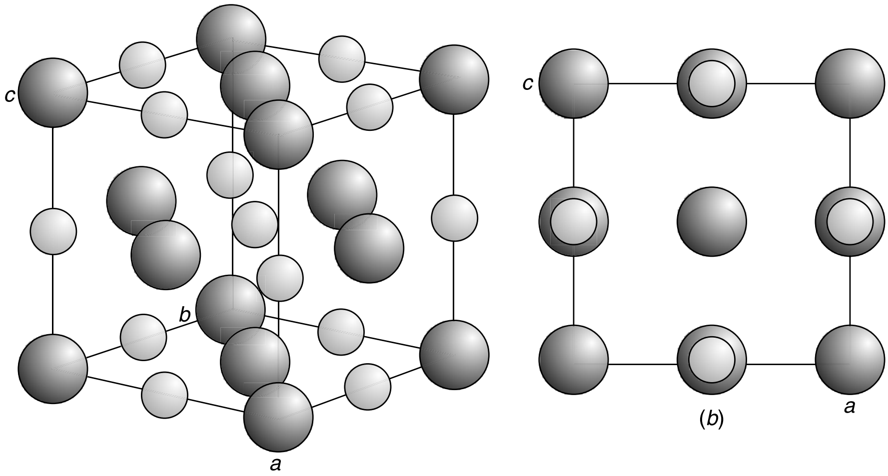


Figure 1.14 Crystal structure of sodium chloride, NaCl: (*a*) a perspective view of one unit cell and (*b*) projection down [010], the *b* axis.

sublattices must be balanced to maintain the correct formula and preserve electrical neutrality. This means that we must introduce equal numbers of vacancies onto both sublattices.

Such a situation was envisaged by Schottky and Wagner, whose ideas were presented in 1931. The defects arising from balanced populations of cation and anion vacancies in any crystal, not just NaCl, are now known as Schottky defects. For example, if the crystal has a formula MX, then the number of cation vacancies will be equal to the number of anion vacancies, in order to maintain the composition and electrical neutrality (or electroneutrality). In such a crystal, one Schottky defect consists of one cation vacancy together with one anion vacancy, although these vacancies are not necessarily imagined to be near to each other in the crystal. It is necessary to remember that the number of Schottky defects in a crystal of formula MX is equal to one-half of *total* the number of vacancies. Schottky defects are frequently represented diagrammatically by a drawing of the sort shown in Figure 1.15*a*. In a real crystal the situation will be more complex, as atoms in the vicinity of the defects will move slightly (relax) due to the changed situation.

In crystals of more complex formula, such as titanium dioxide, TiO_2, a Schottky defect will consist of two anion vacancies and one cation vacancy. This is because it is necessary to counterbalance the loss of one Ti^{4+} ion from the crystal by the absence of two O^{2-} ions in order to maintain composition and electroneutrality. This ratio of two anion vacancies per one cation vacancy will hold in all ionic compounds of formula MX_2. In crystals like Al_2O_3, two Al^{3+} vacancies must be balanced by three O^{2-} vacancies. Thus, in crystals with a formula M_2X_3, a Schottky defect will consist of two vacancies on the cation sublattice and three vacancies on the anion sublattice. These vacancies are not considered to be clustered together but are distributed

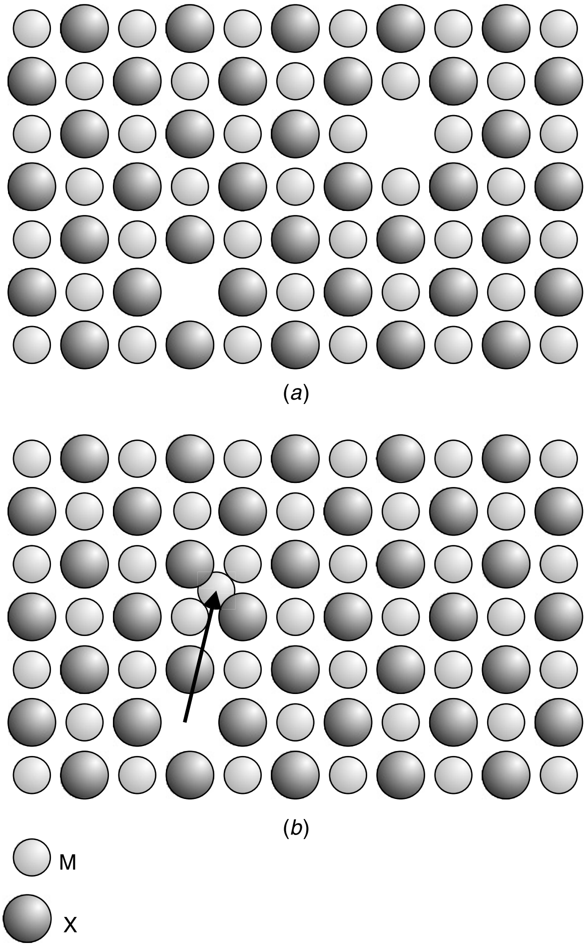


Figure 1.15 Balanced populations of point defects in an ionic crystal of formula MX (schematic): (*a*) Schottky defects and (*b*) Frenkel defects.

at random throughout the crystal structure, in the relative numbers needed to keep the crystals electrically neutral.

The formation energy of Schottky defects is described further in Chapter 2.

1.9.2 Frenkel Defects

As already illustrated, it is also possible to imagine a defect related to ions in interstices, that is, interstitials. Such defects were first suggested as being of importance by Frenkel and are known as *Frenkel defects*. In this case, an atom or ion from one sublattice moves to a normally empty site (an interstitial site) in the crystal, leaving a vacancy behind (Fig. 1.15*b*). A Frenkel defect may involve either the cation or the anion sublattice. Thus, in any crystal of formula MX, a Frenkel defect consists of one interstitial ion plus one vacant site in the sublattice where that ion would normally be found. Because ions are being moved internally, no problem of electronic charge balance occurs. This means that the number of interstitials and vacancies forming a Frenkel defect population is not connected to the formula of the compound. For example, if we have Frenkel defects on the anion sublattice in CaF_2, we can think of just one F^- ion being displaced; it is not necessary to displace two F^- ions to form the Frenkel defect.

The formation energy of Frenkel defects is described further in Chapter 2.

1.10 ANTISITE DEFECTS

An antisite defect is an atom on a site normally occupied by a different chemical species that exists in the compound. Antisite defects are a feature of a number of important materials, especially weakly ionic or covalently bonded ones. In a compound of formula AB the antisite defects that can occur are an A atom on a site normally occupied by a B atom (Fig. 1.16*a*), or a B atom on a site normally occupied by an A atom (Fig. 1.16*b*).

Antisite defects are not important in binary ionic compounds, as the resultant increase in electrostatic energy is generally prohibitive. For example, if an Na^+ ion were to replace a Cl^- ion in NaCl (Fig. 1.16*c*), it would be surrounded by an octahedron of Na^+ cations, leading to strong electrostatic repulsion. Similarly, a Cl^- anion on an Na^+ site would be surrounded by an octahedron of Cl^- anions (Fig. 1.16*d*) again, leading to strong electrostatic repulsion. This is not true when cations of different elements are present together in an ionic compound, as antistite defects can then be accommodated without the electrostatic penalty. An example is the oxide spinel, $MgAl_2O_4$. Notionally all of the Mg^{2+} ions occupy tetrahedrally coordinated sites in the structure, written (Mg), and Al^{3+} ions occupy octahedrally coordinated sites, written $[Al_2]$, so that the structure can be represented as $(Mg)[Al_2]O_4$. In most such crystals it is found that a small number of Mg^{2+} and Al^{3+} cations exchange sites to form antisite defects $(Mg_{1-x}Al_x)[Al_{2-x}Mg_x]O_4$:

$$\mathrm{Mg_{Mg}} + \mathrm{Al_{Al}} \longrightarrow \mathrm{Mg'_{Al}} + \mathrm{Al^{\bullet}_{Mg}}$$

where the effective charges are derived assuming the compound is ionic. This exchange of cations between octahedral and tetrahedral sites is generally designated in terms of normal and inverse spinel structures (Supplementary Material S1). Similarly, in $Er_2Ti_2O_7$, which has the pyrochlore structure, antisite defects occur between the two metal atoms. In this structure, in the absence of defects, the larger atom (Er) is coordinated by eight oxygen atoms in an approximately cubic arrangement, while the smaller atom (Ti) is in distorted octahedral coordination. The antisite defects can be represented as forming by the reaction:

$$\mathrm{Er_{Er}} + \mathrm{Ti_{Ti}} \longrightarrow \mathrm{Er'_{Ti}} + \mathrm{Ti^{\bullet}_{Er}}$$

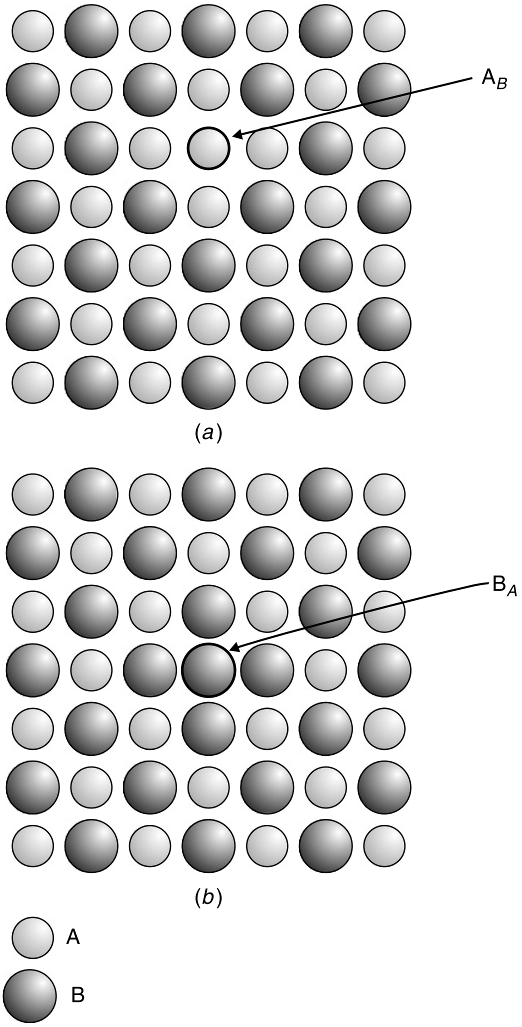


Figure 1.16 Antisite point defects in an ionic crystal of formula MX (schematic): (*a*) A on B sites, A_B; (*b*) B on A sites, B_A; (*c*) Na^+ on a Cl^- site in sodium chloride, $Na_{Cl}^{2\bullet}$; and (*d*) Cl^- on an Na^+ site in sodium chloride, $Cl_{Na}^{2\prime}$.

where the effective charges indicate an ionic model. About 2.9 at % of these defects are found in normal preparations (see also Section 4.4).

In metallic and many semiconducting crystals, the valence electrons are delocalized throughout the solid, so that antisite defects are not accompanied by prohibitive energy costs and are rather common. For example, an important defect in the semiconducting material GaAs, which has the zinc blend structure (Supplementary Material S1), is the antisite defect formed when an As atom occupies a Ga site.

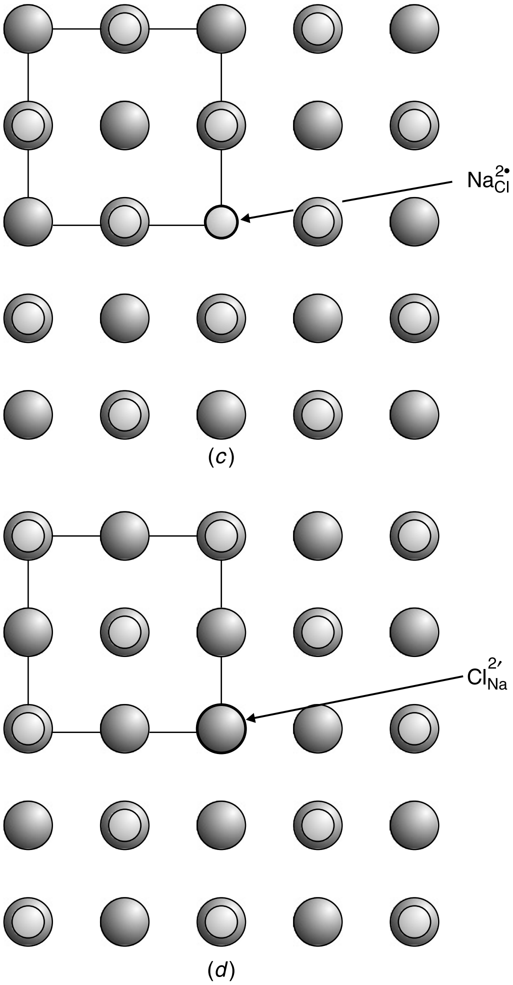


Figure 1.16 (*Continued*).

In cases where the antisite defects are balanced, such as a Ga atom on an As site balanced by an As atom on a Ga site, the composition of the compound is unaltered. In cases where this is not so, the composition of the material will drift away from the stoichiometric formula unless a population of compensating defects is also present. For example, the alloy FeAl contains antisite defects consisting of iron atoms on aluminum sites without a balancing population of aluminum atoms on iron sites. The composition will be iron rich unless compensating defects such as Al interstitials or Fe vacancies are also present in numbers sufficient to restore the stoichiometry. Experiments show that iron vacancies (V_{Fe}) are the compensating defects when the composition is maintained at FeAl.

The creation of antisite defects can occur during crystal growth, when atoms are misplaced on the surface of the growing crystal. Alternatively, they can be created by internal mechanisms once the crystal is formed, provided that sufficient energy is applied to allow for atom movement.

1.11 DEFECT FORMATION AND REACTION EQUATIONS

Defects are often deliberately introduced into a solid in order to modify physical or chemical properties. However, defects do not occur in the balance of reactants expressed in traditional chemical equations, and so these important components are lost to the chemical accounting system that the equations represent. Fortunately, traditional chemical equations can be easily modified so as to include defect formation. The incorporation of defects into normal chemical equations allows a strict account of these important entities to be kept and at the same time facilitates the application of chemical thermodynamics to the system. In this sense it is possible to build up a *defect chemistry* in which the defects play a role analogous to that of the chemical atoms themselves. The Kröger–Vink notation allows this to be done provided the normal rules that apply to balanced chemical equations are preserved.

1.11.1 Addition and Subtraction of Atoms

When writing defect formation equations, the strategy involved is always to add or subtract elements to or from a crystal via *electrically neutral atoms*. When ionic crystals are involved, this requires that electrons are considered separately. Thus, if one considers NiO to be ionic, formation of a V_{Ni} would imply the removal of a neutral Ni atom, that is, removal of a Ni^{2+} ion together with two electrons. Similarly, formation of a V_O would imply removal of a neutral oxygen atom, that is, removal of an O^{2-} ion, followed by the addition of two electrons to the crystal. An alternative way to express this is to say the removal of an O^{2-} ion together with $2h^{\bullet}$. Similarly, only neutral atoms are added to interstitial positions. If ions are considered to be present, the requisite number of electrons must be added or subtracted as well. Thus, the formation of an interstitial Zn^{2+} defect would involve the addition of a neutral Zn atom and the removal of two electrons.

1.11.2 Equation Formalism

The formation of defects can be considered as the *reaction* of a nominally perfect crystal with dopant. The rules for writing equations including defects are similar to those of elementary chemistry, but as the matrix is a crystal structure, quantities must be specified with respect to crystallographic sites rather than molecules or moles.

1. *The number of metal atom sites must always be in the correct proportion to the number of nonmetal atom sites in the crystal.* This is in essence a crystal structure

constraint. In a crystal structure there are a certain number of atoms that occupy specified positions within the space group of the solid. Provided that the crystal structure is (even only loosely) specified, the number of atom sites must remain fixed. Thus, in CaS, which adopts the sodium chloride structure, formula MX, there must always be equal numbers of metal and nonmetal atom positions in the equation. For an oxide such as TiO_2, with the rutile structure, there must always be twice as many anion sites as cation sites, or, in general, for a compound M_aX_b, there must be an a metal atom sites for every b nonmetal atom sites. As long as this proportion is maintained, the *total* number of sites can vary, as this simply corresponds to more or less substance present. If the crystal contains vacancies, these must be counted as part of the total number of sites, as each vacancy can be considered to occupy a site just as legally as an atom. Interstitial atoms do not occupy normal sites and so do not count when this rule is being applied.

2. *The total number of atoms on one side of the equation must balance the total number of atoms on the other side.* This rule is simply an expression of the well-known chemical fact that atoms are neither created nor destroyed during a chemical reaction. Remember that subscripts and superscripts are labels describing charges and sites and are not counted in evaluating the atom balance.

3. *The crystal must always be electrically neutral.* This means not only that the total charge on one side of the equation must be equal to the total charge on the other side, but also that the sum of the charges on each side of the equation must equal zero. In this assessment, both effective and real charges must be counted if both sorts are present.

Recall that only neutral atoms are involved in reactions. After reaction, neutral atoms can dissociate into charged species if this is thought to represent the real situation in the crystal, provided that electroneutrality, as described, is maintained.

Reactions involving the creation, destruction, and elimination of defects can appear mysterious. In such cases it is useful to break the reaction down into hypothetical steps that can be represented by partial equations, rather akin to the half-reactions used to simplify redox reactions in chemistry. The complete defect formation equation is found by adding the partial equations together. The rules described above can be interpreted more flexibly in these partial equations but must be rigorously obeyed in the final equation. Finally, it is necessary to mention that a defect formation equation can often be written in terms of just structural (i.e., ionic) defects such as interstitials and vacancies or in terms of just electronic defects, electrons, and holes. Which of these alternatives is preferred will depend upon the physical properties of the solid. An insulator such as MgO is likely to utilize structural defects to compensate for the changes taking place, whereas a semiconducting transition-metal oxide with several easily accessible valence states is likely to prefer electronic compensation.

To illustrate exactly how these rules work, a number of examples follow. In the first, the formation of antisite defects, a simple example that does not involve changes in atom numbers or charges on defects, is described. Secondly, two reactions involving oxides, nickel oxide and cadmium oxide, both of which are nonstoichiometric, but for opposite reasons, indicate how to deal with a solid–gas interaction in terms of electronic compensation. Thirdly, the formation of calcia-stabilized zirconia, $ZrO_2 \cdot xCaO$, where x is of the order of 10–15 mol % outlines the method for a reaction between two solids utilizing structural compensation. Finally, defect formation in more complex oxides containing two cation species is outlined. None of these reactions can be represented by conventional chemical equations.

1.11.3 Formation of Antisite Defects

The creation of a complementary pair of antisite defects consisting of an A atom on a B atom site, A_B, and a B atom on an A atom site, B_A, can be written in terms of a chemical equation:

$$A_A + B_B \longrightarrow A_B + B_A$$

Antisite defects can be created via the intermediate formation of a Frenkel defect by the following rules:

1. Frenkel defect formation on the A sublattice:

$$A_A \longrightarrow A_i + V_A$$

2. Exchange of the interstitial A_i atom with a B atom on its correct site:

$$A_i + B_B \longrightarrow A_B + B_i$$

3. The interstitial B_i atom can eliminate the vacancy to form a B atom antisite defect:

$$B_i + V_A \longrightarrow B_A$$

4. If these three equations are added, the result is

$$A_A + B_B \longrightarrow A_B + B_A$$

The same equations can also be written in terms of interstitial B atom formation and exchange of the interstitial B_i with an A_A atom, and so forth.

1.11.4 Nickel Oxide

Nickel oxide, NiO, which adopts the sodium chloride structure (Fig. 1.14), can readily be made slightly oxygen rich, and, because the solid then contains more oxygen than nickel, the crystal must also contain a population of point defects. This situation can formally be considered as a reaction of oxygen gas with stoichiometric NiO, and the simplest assumption is to suppose that the extra oxygen extends the crystal by adding extra oxygen sites. Atoms are added as neutral atoms, and initially it is reasonable to suppose that NiO is built of neutral atoms. Because of rule 1 above, each oxygen added must be balanced by the creation of a corresponding nickel site, so that the reaction equation is

$$\tfrac{x}{2}\mathrm{O_2(NiO)} \longrightarrow x\,\mathrm{O_O} + x\,\mathrm{V_{Ni}} \tag{1.2}$$

The reactant, oxygen gas, is written as in a normal equation. Strictly speaking the entry (NiO) to specify the host crystal is superfluous, but its inclusion avoids confusion (see later examples). The host crystal is written in parentheses to emphasize that it is not really part of the equation.[8] It might be considered to be more realistic if NiO is supposed to be ionic, containing Ni^{2+} and O^{2-} ions. In an ionic crystal each Ni vacancy must carry two effective negative charges, written $V_{Ni}^{2\prime}$, because a normal Ni^{2+} ion has been removed. The neutral oxygen atom introduced onto the oxygen ion site will be missing two electrons compared to O^{2-}, so that it carries two effective positive charges, written $O_O^{2\bullet}$. This transformation can be designated by extending Eq. (1.2) in the following way:

$$x\,\mathrm{O_O} + x\,\mathrm{V_{Ni}} \longrightarrow x\,\mathrm{O_O^{2\bullet}} + x\,\mathrm{V_{Ni}^{2\prime}} \tag{1.3}$$

The charges balance, in accordance with rule 3 above. Finally, it may be regarded as reasonable to assume that the added oxygen ends up as O^{2-}. In this case each added atom of oxygen must gain two electrons. These will be taken from another source in the crystal, thereby generating two positive holes, $h^{\bullet}$:

$$x\,\mathrm{O_O^{2\bullet}} \longrightarrow x\,\mathrm{O_O} + 2x\,\mathrm{h^{\bullet}} \tag{1.4}$$

The complete equation is derived by adding Eqs. (1.2), (1.3), and (1.4):

$$\tfrac{x}{2}\mathrm{O_2(NiO)} \longrightarrow x\,\mathrm{O_O} + x\,\mathrm{V_{Ni}^{2\prime}} + 2x\,\mathrm{h^{\bullet}} \tag{1.5}$$

The creation of each vacancy is accompanied by the creation of a hole. If the ionic assumption is correct, the solid would therefore be expected to behave as a *p*-type semiconductor. This is, in fact, the case for NiO. However, if for some analogous case this is not confirmed experimentally, the equation is not valid.

Chemically, it might be preferable to specify exactly the source of the electrons donated to form O^{2-} ions in NiO. The physical properties of the solid suggest that the two electrons actually come from separate Ni^{2+} ions, converting each of them to the ion Ni^{3+}. The defect is then a Ni^{3+} ion located on a Ni^{2+} site. This defect has an effective positive charge of one unit compared to the Ni^{2+} ion, so it would be written $Ni_{Ni}^{\bullet}$. Taking this into account, the reaction Eq. (1.4) needs to be

---

[8] Frequently, in defect formation equations the host crystal is written above the reaction arrows. Throughout this book the convention of placing the host crystal in parentheses on the left side of the equation will be followed.

written as

$$x\,\mathrm{O_O^{2\bullet}} + 2x\,\mathrm{Ni_{Ni}} \longrightarrow x\,\mathrm{O_O} + 2x\,\mathrm{Ni_{Ni}^{\bullet}} \tag{1.6}$$

Adding Eqs. (1.2), (1.3), and (1.6) gives

$$\tfrac{x}{2}\mathrm{O_2(NiO)} + 2x\,\mathrm{Ni_{Ni}} \longrightarrow x\,\mathrm{O_O} + x\,\mathrm{V_{Ni}^{2\prime}} + 2x\,\mathrm{Ni_{Ni}^{\bullet}} \tag{1.7}$$

This reveals that two alternative defect structures can be imagined, one with free holes and one with $\mathrm{Ni^{3+}}$ defects. A further possibility is that the hole may be lightly bound to an $\mathrm{Ni^{2+}}$ ion to give a defect complex that could be written ($\mathrm{Ni_{Ni}} + \mathrm{h^{\bullet}}$). All of these descriptions are valid. The one adopted would be the one most consistent with the measured properties of the solid.

1.11.5 Cadmium Oxide

Cadmium oxide, CdO, like nickel oxide, also adopts the sodium chloride structure (Fig. 1.14). However, unlike nickel oxide, this compound can be made to contain more metal than oxygen. The defects that cause this metal excess are usually considered to be interstitial Cd atoms or ions. In this case the reaction is one in which the solid formally loses oxygen. Because of the rules of equation writing, this must involve the removal of neutral oxygen atoms. Each oxygen lost results in the loss of a nonmetal site. In order to keep the site ratio correct, a metal site must also be lost, forcing the metal into interstitial sites:

$$x\,\mathrm{CdO} \longrightarrow \tfrac{x}{2}\mathrm{O_2} + x\,\mathrm{Cd_i}$$

If CdO is considered to be ionic, and built of $\mathrm{Cd^{2+}}$ and $\mathrm{O^{2-}}$ ions, the cadmium interstitial will show two effective positive charges, written $\mathrm{Cd_i^{2\bullet}}$. At the same time, removal of an oxygen atom will require that the two electrons that are part of the $\mathrm{O^{2-}}$ ion are left in the crystal:

$$x\,\mathrm{CdO} \longrightarrow \tfrac{x}{2}\mathrm{O_2} + x\mathrm{Cd_i^{2\bullet}} + 2x\,\mathrm{e'}$$

The presence of free electrons means that metal-rich CdO should be an *n*-type semiconductor. It is possible to imagine that the divalent cadmium interstitials can take up an electron to form monovalent interstitials:

$$x\,\mathrm{CdO} \longrightarrow \tfrac{x}{2}\mathrm{O_2} + x\,\mathrm{Cd_i^{\bullet}} + x\,\mathrm{e'}$$

As in the case of nickel oxide, the equation that best represents the real situation can only be determined experimentally. In this case one possible method would be to measure the way in which the electronic conductivity varies with the partial pressure of the surrounding oxygen atmosphere (Chapter 7).

1.11.6 Calcia-stabilized Zirconia

The same principles apply when two solids react, as can be illustrated by the reaction of a crystal of zirconia, ZrO_2, with a small amount of calcia, CaO, to produce a crystal of calcia-stabilized zirconia. There are two principal ways that this reaction can be imagined to occur: either the Ca atoms occupy Zr sites or they occupy interstitial sites.

Suppose that on reaction Ca atoms are located on normal metal sites in ZrO_2. In order to comply with rule 1 in Section 1.11.2, it is necessary to create two anion sites per Ca atom introduced. Initially, these are vacant. However, as an oxygen atom from the CaO must also be accommodated in the crystal, it is reasonable to place it in one of these sites. The other site remains vacant. If we take the ZrO_2 to be ionic and composed of Zr^{4+} and O^{2-} ions, the added (neutral) atoms will acquire effective charges. The neutral Ca atom at the Zr^{4+} displays an effective charge of 4 with respect to the charge experienced when a normal Zr^{4+} ion is present, so it is written as $Ca_{Zr}^{4\prime}$. Similarly, for ionic $(Zr^{4+}O_2^{2-})$ a neutral oxygen atom at an occupied oxygen site acquires an effective charge of 2•, and is hence written $O_O^{2\bullet}$. The same is true of the oxygen vacancy. Compared to a site occupied by an O^{2-} ion, the site has an effective positive charge of 2•, so the vacancy is written $V_O^{2\bullet}$. Formally, one can therefore write

$$CaO(ZrO_2) \longrightarrow Ca_{Zr}^{4\prime} + V_O^{2\bullet} + O_O^{2\bullet} \tag{1.8}$$

The host crystal, ZrO_2, is written in parentheses (ZrO_2). Note that the equation conserves mass balance, electrical charge balance, and site numbers in accordance with the rules given in Section 1.11.2. Now CaO is normally regarded as an ionic compound, as implied by the foregoing allocation of charges, so that *ions* should occupy the sites, not neutral atoms. To achieve this, two electrons are transferred from the neutral Ca atom to create Ca^{2+} and lodged upon the neutral O atom to create O^{2-}. The effective charge on the Ca ion will now be just two negative units instead of four, and there will be no effective charge on the oxygen ion as all normal oxygen sites contain O^{2-}. To stress this electroneutrality, it is given the symbol O_O^{x}. There is no change in the status of the oxygen vacancy. The process is summarized by

$$Ca_{Zr}^{4\prime} + V_O^{2\bullet} + O_O^{2\bullet} \longrightarrow Ca_{Zr}^{2\prime} + V_O^{2\bullet} + O_O^{x} \tag{1.9}$$

The final outcome for the reaction between the two *ionic* crystals is obtained by adding Eqs. (1.8) and (1.9) to give

$$CaO(ZrO_2) \longrightarrow Ca_{Zr}^{2\prime} + V_O^{2\bullet} + O_O^{x}$$

For chemical reasons it might be argued that the Ca^{2+} ions do not occupy Zr^{4+} sites but prefer interstitial positions, while the oxygen atoms occupy newly formed sites. The Ca interstitial atoms do not affect site numbers, but the oxygen atoms must maintain the site ratio of the ZrO_2 matrix, so that one Zr vacancy must be created for each pair of oxygen atoms added. Because neutral atoms are added, the vacancy and oxygen atoms will carry effective charges as above. The reaction is

$$2CaO(ZrO_2) \longrightarrow 2Ca_i + 2O_O^{2\bullet} + V_{Zr}^{4\prime} \tag{1.10}$$

If the added species are ions (Ca^{2+} and O^{2-}) the Ca^{2+} ions will have an effective charge of 2•, written $Ca_i^{2\bullet}$. Each oxide ion has no effective charge compared to a normal oxide anion, written O_O^x. There is no change as far as the Zr vacancy is concerned. The transformation can formally be expressed as

$$2Ca_i + 2O_O^{2\bullet} + V_{Zr}^{4\prime} \longrightarrow 2Ca_i^{2\bullet} + 2O_O^x + V_{Zr}^{4\prime} \tag{1.11}$$

The reaction of ionic CaO with ionic ZrO_2 can be obtained by addition of Eqs. (1.10) and (1.11):

$$2CaO(ZrO_2) \longrightarrow 2Ca_i^{2\bullet} + 2O_O^x + V_{Zr}^{4\prime}$$

Experimental evidence is needed to decide which (if any) of these formal equations represents the true situation in the material (see Section 4.4.5).

1.11.7 Ternary Oxides

The principles described above apply equally well to oxides with more complex formulas. In these materials, however, there are generally a number of different cations or anions present. Generally, only one of the ionic species will be affected by the defect forming reaction while (ideally) others will remain unaltered. The reactant, on the other hand, can be introduced into any of the suitable ion sites. This leads to a certain amount of complexity in writing the defect equations that apply. The simplest way to bypass this difficulty is to "decompose" the complex oxide into its major components and treat these separately. Two examples, using the perovskite structure, can illustrate this.

The perovskite structure, ABO_3 (where A represents a large cation and B a medium-size cation) is adopted by many solids and solid solutions between them can readily be prepared. Vacancy-containing systems with the perovskite structure are of interest as electrolytes in solid-state batteries and fuel cells. Typical representatives of this type of material can be made by introducing a higher valence cation into the A sites or a lower valance cation into the B sites.

The first example is the substitution of La^{3+} for the alkaline earth Ca^{2+} cation in $CaTiO_3$. In this reaction, the Ti sites are unaffected, and it can be written simply in terms of the reaction between the impurity La_2O_3 and a CaO matrix. The site maintenance rule means that three oxygen anions added (from La_2O_3) must be balanced by the creation of three cation sites. Assuming ionic materials:

$$La_2O_3(CaO) \longrightarrow 2La_{Ca}^{\bullet} + V_{Ca}^{2\prime} + 3O_O$$

One vacancy is generated for every two La^{3+} substituents. If one wishes to include the TiO_2, which is a "sleeping partner," it is only necessary to maintain the correct stoichiometry. The creation of three new A cation sites ($2La_{Ca}^{\bullet} + V_{Ca}^{2\prime}$) requires the inclusion of three Ti sites and the corresponding number of anion sites:

$$La_2O_3(CaTiO_3) \longrightarrow 2La_{Ca}^{\bullet} + V_{Ca}^{2\prime} + 3Ti_{Ti} + 9O_O$$

Exactly the same approach can be employed for substitution on the B sites, and this can provide our second example. Assuming ionic compounds, consider the substitution of Cr^{3+} for Ti^{4+} in $CaTiO_3$. This time the CaO component is the sleeping partner and can be ignored so the reaction can be considered to be between the Cr_2O_3 impurity and the TiO_2 matrix. The site maintenance rule requires that inclusion of two Cr^{3+} ions on Ti^{4+} sites requires the creation of four O^{2-} sites:

$$Cr_2O_3(TiO_2) \longrightarrow 2Cr_{Ti}^{\prime} + 3O_O + V_O^{2\bullet}$$

One anion vacancy is generated for every two Cr^{3+} substituents. As before, if it is helpful, include the CaO, simply maintain the stoichiometry. The creation of two new B cation sites ($2Cr_{Ti}^{\prime}$) requires the inclusion of two Ca sites and two anion sites:

$$Cr_2O_3(CaTiO_3) \longrightarrow 2Ca_{Ca} + 2Cr_{Ti}^{\prime} + 5O_O + V_O^{2\bullet}$$

1.12 COMBINATIONS OF POINT DEFECTS IN PURE MATERIALS

There is no obvious reason why only one defect type should occur in a crystal, and several different species would be expected to be present. However, the formation energy of each defect type is different, and it is often a reasonable approximation to assume that only one or a small number of defect types will dominate the chemical and physical properties of the solid.

For example, the formation of an intrinsic interstitial defect requires the simultaneous creation of a vacancy. These may not remain close together in the crystal, and it is legitimate to consider that the two defects occur in equal numbers. Thus, in silicon it is possible to write the formation equation for silicon self-interstitials, Si_i, as

$$Si_{Si} \longrightarrow Si_i + V_{Si}$$

where Si_{Si} represents the perfect crystal before defect formation, Si_i represents the self-interstitial, and V_{Si} represents a vacancy on a normally occupied silicon site. Similarly, in the alloy FeAl, antisite defects consisting of Fe atoms on Al sites,

Fe_{Al}, coexist with vacancies on Fe sites, V_{Fe} (Section 1.10). The iron sites, therefore, contain populations of two defects, vacancies and aluminum atoms.

Calculations of defect formation energies (Section 2.10) will allow estimates of the numbers of each kind of defect present to be made using formulas of the type:

$$n_d \propto \exp\left(\frac{-E_d}{RT}\right)$$

where n_d is the number of defects present, E_d the molar formation energy of the defect, R the gas constant, and T the temperature (K).

1.13 STRUCTURAL CONSEQUENCES OF POINT DEFECT POPULATIONS

The various types of point defect found in pure or almost pure stoichiometric solids are summarized in Figure 1.17. It is not easy to imagine the three-dimensional consequences of the presence of any of these defects from two-dimensional diagrams, but it is important to remember that the real structure of the crystal surrounding a defect can be important. If it is at all possible, try to consult or build crystal models. This will reveal that it is easier to create vacancies at some atom sites than others, and that it is easier to introduce interstitials into the more open parts of the structure.

Moreover, point defects will undoubtedly cause significant crystallographic distortions in the region of the affected site, which endow the defects with a

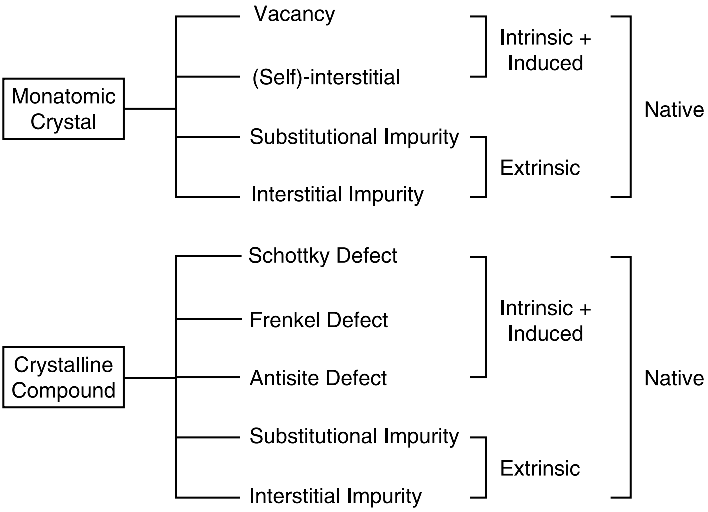


Figure 1.17 Point defects in almost pure stoichiometric crystals.

spatial extension greater than that implied in the term point defect. The resulting strain field plays an important part in the aggregation of point defects into clusters and other extended defects. The fraction of point defects that aggregate will depend upon the relative energy gain when a cluster forms, as well as the mobility of the defects in the crystal. The electronic and other interactions between defects can be calculated using approaches analogous to the Debye–Hückel treatment used for electrolytes or the Lorentz theory of the polarizability of insulating solids. In fact, simulations and quantum mechanical calculations (Chapter 2) suggest that defect association is the rule rather than the exception.

1.14 ANSWERS TO INTRODUCTORY QUESTIONS

What is a point defect?

A point defect is a localized defect that consists of a mistake at a single atom site in a solid. The simplest point defects that can occur in pure crystals are missing atoms, called vacancies, or atoms displaced from the correct site into positions not normally occupied in the crystal, called self-interstitials. Additionally atoms of an impurity can occupy a normal atom site to form substitutional defects or can occupy a normally vacant position in the crystal structure to form an interstitial. Other point defects can be characterized in pure compounds that contain more than one atom. The best known of these are Frenkel defects, Schottky defects, and antisite defects.

What is the effective charge on a defect?

The effective charge on a defect is the charge with respect to that normally occurring at the same site in a perfect crystal. Interstitial sites are normally unoccupied, so an atom or ion at an interstitial site will have an effective charge equal to its real charge. For example, interstitial Ti^{4+} ions will have an effective charge of $4\bullet$. Any neutral atom, Zn, for example, will have no effective charge at an interstitial site. When an impurity atom or an impurity ion occupies a site that normally contains an ion, the effective charge will be the difference between the real charge on the impurity atom or ion and the real charge on the normal ion. Thus, a Ca^{2+} ion located on a Na^{+} site in a crystal will carry an effective charge of $2 - 1 = 1$, that is, it will be written $Ca_{Na}^{\bullet}$. Similarly, a neutral atom such as Li located at the same site would carry an effective charge of $0 - 1 = -1$, that is, it will be written $Li_{Na}^{\prime}$.

What is an antisite defect?

An antisite defect is an atom on an inappropriate site in a crystal, that is, a site normally occupied by a different chemical species. In a compound of formula AB the antisite defects that can occur are an A atom on a site normally occupied by a B atom, or a B atom on a site normally occupied by an A atom. Antisite defects are not very important in binary ionic compounds, as the misplacement of an ion is energetically costly, and so unfavorable. In ternary ionic compounds, however, such as spinels, AB_2O_4, the transfer of A ions to B sites and vice versa, is not expensive on energy and can occur. In metallic and many semiconducting crystals, the valence electrons are delocalized throughout the solid, so that also in these cases antisite defects are not accompanied by prohibitive energy costs. For example, most important defect in the semiconducting material GaAs is the antisite defect formed when an As atom occupies a Ga site.

PROBLEMS AND EXERCISES

Quick Quiz

1. Point defects that form in a crystal exposed to radioactivity are called:
   (a) Intrinsic defects
   (b) Induced defects
   (c) Native defects

2. Intrinsic defects in crystals:
   (a) Are always present
   (b) Only form if two or more different atom species occur
   (c) Only form in the presence of impurities

3. Impurity atoms are:
   (a) Intrinsic defects
   (b) Extrinsic defects
   (c) Native defects

4. An impurity P atom on a Si site in a crystal of silicon is a:
   (a) Interstitial defect
   (b) Antisite defect
   (c) Substitutional defect

5. The number of Schottky defects in a crystal of formula MX is equal to:
   (a) Half the total number of vacancies present
   (b) The total number of vacancies present
   (c) Double the total number of vacancies present

6. A single anion Frenkel defect in an ionic crystal of formula MX_2 needs to be balanced by:
   (a) A cation Frenkel defect
   (b) Two cation Frenkel defects
   (c) Nothing

7. A metal atom N on an M site in a crystal MX would be written:
   (a) N_M
   (b) M_N
   (c) N_{MX}

8. An anion vacancy in an ionic oxide has the symbol:

- **(a)** V_O^{x}
- **(b)** $V_O^{2\prime}$
- **(c)** $V_O^{2\bullet}$

9. A Ca^{2+} ion on a Na^{+} site in a crystal is written:

- **(a)** Ca_{Ca}^{x}
- **(b)** $Ca_{Na}^{\bullet}$
- **(c)** $Ca_{Na}^{\prime}$

10. The equation representing doping of La_2O_3 into CaO is:

- **(a)** $La_2O_3\ (CaO) \longrightarrow 2La_{Ca}^{\bullet} + V_{Ca}^{2\prime} + 3O_O$
- **(b)** $La_2O_3\ (CaO) \longrightarrow 2La_{Ca}^{3\bullet} + V_{Ca}^{2\prime} + 3O_O$
- **(c)** $La_2O_3\ (CaO) \longrightarrow 2La_{Ca}^{3\prime} + V_{Ca}^{2\prime} + 3O_O$

Calculations and Questions

1. **(a)** A ruby crystal has a composition $(Al_{0.99}Cr_{0.01})_2O_3$. How many Cr^{3+} ions are there in a ruby of dimensions 1 cm^3? (The unit cell dimensions are $a =$ 0.4763 nm, $c = 1.3009$ nm; Z (the unit cell contents) $= 6Al_2O_3$; the volume of the unit cell can be taken as $0.966a^2c$.)

(b) A solid solution between the spinel structure phases $MgGa_2O_4$ ($a =$ 0.82780 nm) and $MgGaMnO_4$ ($a = 0.83645$ nm) obeys Vegard's law. What is the composition of the phase with a lattice parameter of 0.83000 nm?

2. **(a)** The unit cell of zirconium sulfide, ZrS, is sodium chloride type, $a = 0.514$ nm. The measured density is 4800 kg m^{-3}. Suggest a possible defect structure and formula for the solid. There are four ZrS units per unit cell, relative molar mass: Zr, 91.22 g mol^{-1}; S, 32.07 g mol^{-1}.

(b) A sample of a cubic calcia-stabilized zirconia ceramic has a density of 5720 kg m^{-3} and a lattice parameter of 0.5130 nm. Assuming $V_O^{2\bullet}$ are the dominant point defects, suggest a composition for the solid. There are four MO_2 units in the unit cell. Molar mass: Zr, 91.22 g mol^{-1}; Ca, 40.08 g mol^{-1}; O, 16.00 g mol^{-1}.

3. The Seebeck coefficient for pure $LaCoO_3$ is $+600\ \mu V\ K^{-1}$. (**a**) What are the mobile charge carriers? (**b**) Suppose these occur because the crystal contains a trace of an impurity, Co^{4+}, calculate the defect concentration and the formula of the material (data from Robert *et al.*, 2006).

4. Estimate the Seebeck coefficient for the reduced oxide $TiO_{1.94}$, assuming that the defects are Ti^{3+} ions and the parent phase is TiO_2.

5. Plot a graph of the expected Seebeck coefficient of $La_{1-x}Sr_xMnO_3$ for $x = 0$–0.2.

6. A complex Hg-Sr-Ca-Co containing oxide has a Seebeck coefficient of 115 μV K^{-1}, due to the presence of both Co^{3+} and Co^{4+}. If the defects are the Co^{4+} ions, what is the average valence of the Co ions (data from Pelloquin *et al.*, 2003).

7. Write defect equations for the following reactions:
   - **(a)** MgO doped with Li_2O
   - **(b)** Slight reduction of PrO_2
   - **(c)** TiO_2 doped with Ga_2O_3
   - **(d)** $LaCoO_3$ doped with CaO
   - **(e)** $MgAl_2O_4$ doped with V_2O_3 to form $Mg(Al_{2-x}V_x)_2O_4$

8. Nickel oxide, NiO, is doped with lithium oxide, Li_2O, to form $Li_xNi_{1-x}O$ with the sodium chloride structure. (**a**) Derive the form of the Heikes equation for the variation of Seebeck coefficient, α, with the degree of doping, x. The following table gives values of α versus $\log[(1-x)/x]$ for this material. (**b**) Are the current carriers holes or electrons? (**c**) Estimate the value of the constant term k/e.

α/mV K^{-1}	log $[(1-x)/x]$
0.12	0.2
0.18	0.5
0.28	1.0
0.39	1.5
0.48	2.0
0.59	2.5
0.69	3.0

Data adapted from E. Antolini, *Mater. Chem. Phys.*, **82**, 937–948 (2003).

9. The oxides with general formulas A_2O, AO, A_2O_3, AO_2, and A_2O_5 are doped into MgO so that the cation substitutes for Mg. Write general defect equations for the reactions, assuming cation vacancies rather than electronic compensation occurs.

10. The crystallographic data for Al_2O_3 are $a = 0.47628$ nm, $c = 13.0032$ nm, and for V_2O_3 are $a = 0.4955$ nm, $c = 14.003$ nm. The structure of both oxides is the corundum type, the volume of the hexagonal unit cell can be taken as $0.966a^2c$ and Z (the unit cell contents) = 6 formula units. (**a**) Calculate the density of Al_2O_3. (**b**) Calculate the density of V_2O_3. (**c**) Provided Vegard's law is obeyed, estimate the density of $V_{0.6}Al_{1.4}O_3$. (**d**) The measured density is 3620 kg m^{-3}. What does this suggest about Vegard's law dependence?

REFERENCES

D. Pelloguin, A. Maignan, S. Hébert, C. Michel, and B. Raveau, *J. Solid State Chem.*, **170**, 374–381 (2003).

R. Robert, L. Becker, M. Trottman, A. Reller, and A. Weidenkraft, *J. Solid State Chem.*, **179**, 3893–3899 (2006).

FURTHER READING

Much of the material in this chapter is introductory. Background information on solids is given in:

R. J. D. Tilley, *Understanding Solids*, Wiley, Chichester, 2004.

R. J. D. Tilley, *Crystals and Crystal Structures*, Wiley, Chichester, 2006.

A. F. Wells, *Structural Inorganic Chemistry*, 5th ed., Oxford University Press, Oxford, United Kingdom, 1984.

A. R. West, *Solid State Chemistry*, Wiley, Chichester, 1984.

Expansion and further explanation with respect to defects and defect chemistry and physics will be found in:

F. Agullo-Lopez, C. R. A. Catlow, and P. D. Townsend *Point Defects in Materials*, Academic, New York, 1988.

D. M. Smyth *The Defect Chemistry of Metal Oxides*, Oxford University Press, Oxford, United Kingdom, 2000.

A. M. Stoneham, *The Theory of Defects in Solids*, Oxford University Press, Oxford, United Kingdom, 1985.

Various authors *Mat. Res. Soc. Bull.*, **XVI**, November (1991) and December (1991).

Many aspects of the study and importance of defects in solids presented from a historical and materials perspective are to be found in:

R. W. Cahn, *The Coming of Materials Science*, Pergamon/Elsevier, Oxford, United Kingdom, 2001.

Details about the Heikes formula is found in:

R. R. Heikes and R. W. Ure, Jr., *Thermoelectricity; Science and Engineering*, Wiley/Interscience, New York, 1961, p. 40.