11

Point Defects, Charge, and Diffusion

CHAPTER PREVIEW

Point defects are particularly important in ceramics because of the role they can play in determining the properties of a material. The entire semiconductor industry is possible because of minute concentrations of point defects that are added to Si: the dopants determine if the Si is n-type, p-type, or semi-insulating: they determine the electrical properties. Solid-oxide fuel cells work because of the large concentrations of oxygen vacancies present: the vacancies provide fast ion conduction pathways. CZ is cubic because of the presence of point defects that stabilize the cubic structure.

We will address three principal questions in this chapter and leave the properties associated with point defects to later chapters where we will discuss conduction, light, and color, for example.

- What point defects are possible?
- How many point defects are present?
- How easily can they move?

To determine what defects are possible in a crystal we need to know its structure. Point defects also exist in glass, but there is the extra problem of how to define what is being “defective.” (What is the perfect structure of the glass?) To estimate the concentration of point defects, we should know (1) what types of defects are present and (2) how they form, which, in turn, determines (3) how many of each kind there will be. Answering the third question will give us many of the properties of the materials containing these defects. In some cases we want them to move quickly and in other cases we actually would rather they were immobile. To really understand this subject thoroughly you will need a good knowledge of thermodynamics and kinetics. These topics are so large that there are many texts devoted entirely to them.

11.1 ARE DEFECTS IN CERAMICS DIFFERENT?

Until now we have considered the ideal structures of crystals only when each atom or ion is on a regular site in the crystal. Real crystals contain a variety of imperfections or defects. In crystalline ceramics and glasses, the structure and chemistry of the material will be determined by the kinetics of defect movement. For example, the kinetics of the glass-to-crystal transformation are slow if the temperature is low (typically less that 1000°C) because the transformation occurs by atoms moving—in ceramics, this usually occurs by point defects moving. If point defects move too slowly, the structure with the lowest energy (the equilibrium structure) may never actually be achieved. How fast they move is determined by their structure.

THE OTHER DEFECT CHAPTERS

Dislocations—Chapter 12
Surfaces—Chapter 13
Grain boundaries—Chapter 14
Phase boundaries, particles, and voids—Chapter 15

The really special feature of ceramics is that they can contain charged defects; metals cannot. There are some definite differences when we compare defects in ceramics to similar ones in face-centered cubic (fcc) metal crystals such as Cu.

- The concentration of impurities in ceramics is usually much greater than that of intrinsic defects.
- Dislocations are usually much less important for deformation mechanisms than they are in metals.
- Surfaces and interfaces are even more important for ceramics.
- Pores and voids are often very much more important for ceramics.

The importance of pores and voids is a good illustration of how time changes our view of materials. In the past,

TABLE 11.1 Defect Hierarchy

Dimension	*Name*	*Example*
Zero	Point defect	Vacancy
One	Line defect	Dislocation
Two	Planar defect	Grain boundary
Three	Volume defect	Pore

processing ceramics often meant trying to remove pores and voids during sintering because they weakened the final material. While they still generally do so, we are now also very interested in highly porous ceramics because such materials can have a very large surface area, which may be critical in their application as catalysts, catalyst supports, or filters. In this chapter we concentrate on point defects, but remember that they do relate to all other defects in ceramics and their importance is that they determine many important properties.

11.2 TYPES OF POINT DEFECTS

Defects are often classified in terms of a dimensionality. This is known as the defect hierarchy. The classifications are given in Table 11.1. In spite of this table, remember that all these defects are three-dimensional. We will first summarize the different types of point defect that can occur in crystalline materials. To provide some idea of the importance of point defects, we will consider some specific examples.

Vacancies: If an atom is not present on the site that it should occupy in a perfect crystal then a vacancy is located at that site. A Schottky defect is a set of vacancies created by removing one atom for each atom in the chemical formula. Thus, in a stoichiometric crystal such as MgO, we get a pair of vacancies, one on the Mg sublattice and one on the O sublattice, as shown in Figure 11.1. In spinel, a Schottky defect consists of seven vacancies.

Interstitials: If an atom is present on any site that would be unoccupied in a perfect crystal then that atom is an interstitial. A Frenkel defect is a vacancy + interstitial pair formed by removing an atom from its site in the crystal structure and putting it into an interstice as illustrated in Figure 11.2. Frenkel defects formed in iodine-containing AgBr are essential to the photographic process.

Misplaced atoms: If an atom is present on a crystal site that should be occupied by a different atom, that atom is a misplaced atom and may be called an antisite defect. Antisite defects usually form in covalent ceramics such as AlN and SiC, but can also occur in complex oxides that have several different types of cation, for example, spinels and garnets. (We do not expect to see cations on anion sites and vice versa.)

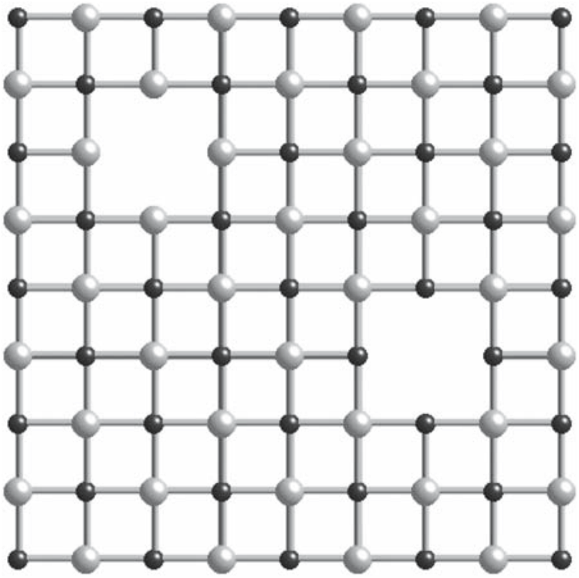

FIGURE 11.1 Schematic of a Schottky defect drawn in 2D.

Associated centers: When two point defects interact so that they can be considered as a single defect they are called an associated center or, if more atoms are involved, a defect cluster or a defect complex. Exposing a material to ionizing radiation such as X-rays and γ-rays can create large numbers of defect clusters.

Solute (substitutional) atoms: In Cu alloys, we can add up to 30 at% Zn before the structure is changed. All the Zn atoms sit on Cu sites so they substitute for the Cu and the crystal is said to be a solid solution of Zn in Cu. The effect is to add electrons to the *d* band of the alloy and it is this change in the electronic structure that determines the solubility limit. We can similarly substitute Ge in Si, but the solubility is limited due to the difference in atomic size. In GaAs we can replace the Ga atom by Al on the group III sublattice

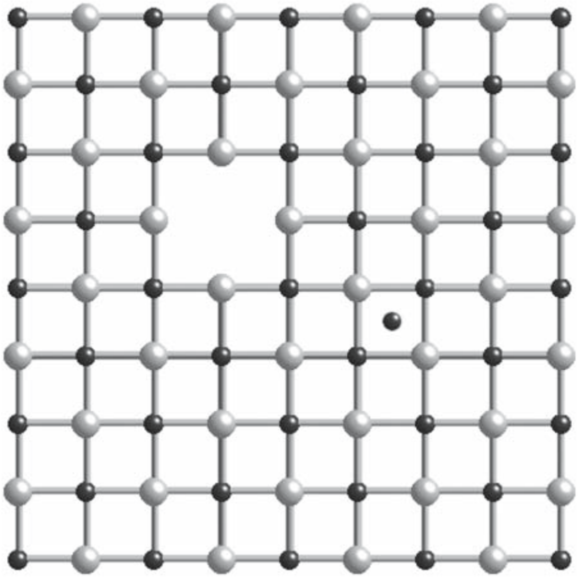

FIGURE 11.2 Schematic of a Frenkel defect drawn in 2D.

TABLE 11.2 The Kröger–Vink Notation for the Binary MX

Defect	*Notation*	*Comments*
Vacancy	V_x	V is also the chemical symbol for the element vanadium; if there is a possibility for confusion, such as expressing a vacancy on a vanadium site, then use "Va" for the vacancy
Interstitial	M_i and X_i	No "site"
Antisite atom	M_X and X_M	Usually in covalent solids
Associated defect	(M_iX_M)	Larger associations called "clusters"
Solute	S_X	Impurity substitution
Electron and hole	e' and h	
Schottky	(V_MV_X)	Special associated defects
Frenkel	(V_MM_i)	

giving a complete solid solution, $Ga_xAl_{1-x}As$, with x running from 0 to 1. GaAs and AlAs have the same structure and similar lattice parameters. We can replace As by P on the group V sublattice to give another continuous solid solution. This type of substitution occurs in both covalent and ionic ceramics as well as in metals.

Electronic defects: Electrons and holes can both exist in ceramics. They may be associated with particular ions, in which case they are simply charged point defects. This leads to the topic of color centers, which color alkali halide crystals, amethyst, diamond, etc.

We use the Kröger–Vink notation to identify these different point defects, which is summarized in Table 11.2. This notation is completely general in that it can apply to any crystalline compound or even to pure crystals. In this notation, structural elements are denoted as S_P^c.

We will make use of Kröger–Vink notation in many sections in this chapter. We can write simple equations to describe the formation of point defects or their interactions. For example, if we remove a molecule, NaCl, from a crystal of rocksalt, we form a vacancy on both anion and cation sublattices, so we have to have a method for distinguishing these defects. The Kröger–Vink notation for the reaction we have just described is

$$V_{NaCl} \rightarrow V'_{Na} + V^{\bullet}_{Cl}$$

NOTATION SUMMARY FOR S_P^c

S: Species
P: Position in the crystal
c: Charge relative to the perfect crystal
Charge notation:
  Negative: dash, ′
  Positive: point, •
  Neutral: *x*

We need such a notation because of one of the most special features about ceramics—the charge. Other notations are sometimes used, but the Kröger–Vink notation is the most widely accepted. (You may see variations in this notation so be careful in translating from one text to another.) The topic of point defects should not be completely new to you. Some of the fields in which you may have encountered point defects before are listed in Table 11.3.

11.3 WHAT IS SPECIAL FOR CERAMICS?

Point defects are special in ceramics when they involve either charge (ionic materials) or dangling bonds (covalent materials).

- Vacancies, interstitials, and substitutional defects can all be charged. The special point defect in ceramics is the charged vacancy. Frenkel and Schottky defects are overall neutral.
- Association of defects is particularly important for ceramics because Coulombic interactions are both strong and long-range.
- Electronic structure is increasingly important. In the past, the electronic structure was important only because the materials have a large band gap. Ionic materials are insulators. This is no longer the case—many old "insulators" (like SiC) are now also wide-band-gap semiconductors!
- Nonstoichiometry is so important because the concentration of point defects can be very large. In ionic materials, point defects are usually charged and they are often numerous.

TABLE 11.3 Examples of Commonly Encountered Point Defects in Materials

Type of point defect	*Application*	*Comments*
Solute atoms	Doping silicon for ICs	Concentration of dopant atoms very small, ~0.0001% can cause a 100 times increase in conductivity
Solute atoms	Solid solution strengthening of metals	A 50 Cu–50 Ni alloy has twice the tensile strength of pure copper
Interstitial	Carburizing of iron and steel	Surface hardening by heating metal in a hydrocarbon gas
Vacancies	Kirkendall effect	Failure mechanism in Au–Al bonds, e.g., Au wire bonds on Al metallizations in ICs
Vacancies and interstitials	Cold working of metals	At $T < 0.4\,T_m$ solid-state diffusion is slow

Some special rules for ceramics:

- The number of sites is constant. This is the same as in metals, but we often have very different sites available in ceramics.
- The total charge is still zero.
- Intrinsic defect concentrations are often very much lower than impurity concentrations.

11.4 WHAT TYPE OF DEFECTS FORM?

Crystals with open lattice structures will tend to favor Frenkel defects. These crystals have structures with low coordination numbers and large interstitial sites, for example, the würtzite structure has CN = 4. If there is a large difference in size between the cations and anions, Frenkel defects are usually formed by displacing the smaller ion. In Al_2O_3, for example, the cation is smaller and we would expect to form cation Frenkel defects. However, anion Frenkel defects will form in UO_2, CeO_2, and ThO_2, which all have large cations. In contrast, we would expect to find Schottky defects in crystals with high coordination numbers such as MgO.

Vacancies may be present on different sites for the same type of ion. For example, $NiFe_2O_4$ is an inverse spinel $[Fe(NiFe)O_4]$; Fe^{3+} ions sit in both tetrahedral and octahedral sites. Thus, we can have Fe^{3+} vacancies on the tetrahedral or octahedral sites, but these two types of vacancy are not the same. You can imagine how complex the situation becomes for more complex crystals. We usually ignore such complications because they are too difficult to handle, but be aware that we are making this simplification.

SCHOTTKY DEFECTS: THE CALCULATION

Show that the number of Schottky defects is

$$n_s \approx N\exp(-\Delta E_s/2kT) \qquad \text{Box 11.1}$$

The units of ΔE_s are J/defect. Then determine entropy using probability theory.

$$S = k\ln W \qquad \text{Box 11.2}$$

$$W = N!/[(N-n)!n!] \qquad \text{Box 11.3}$$

$$W_A = N!/[(N-n_s)!n_s!] \qquad \text{Box 11.4}$$

$$W_B = N!/[(N-n_s)!n_s!] \qquad \text{Box 11.5}$$

$$W_A = W_B \qquad \text{Box 11.6}$$

$$W_T = W_A W_B \qquad \text{Box 11.7}$$

$$\Delta S = k\ln W_T = k\ln\{N!/[(N-n_s)!n_s!]\}^2 \qquad \text{Box 11.8}$$

$$\Delta S = 2k\ln\{N!/[(N-n_s)!n_s!]\} \qquad \text{Box 11.9}$$

Next use math:

$$\ln N! \approx N\ln N - N \qquad \text{Box 11.10}$$

$$\Delta S = 2k\{N\ln N - (N-n_s)\ln(N-n_s) - n_s\ln n_s\} \qquad \text{Box 11.11}$$

Finally do the substitutions:

$$\Delta G = n_s\Delta E_s - 2kT\{N\ln N - (N-n_s)\ln(N-n_s) - n_s\ln n_s\} \qquad \text{Box 11.12}$$

$$(\partial\Delta G/\partial n_s)_{T,P} = 0 \qquad \text{Box 11.13}$$

$$\Delta E_s = 2kT\ln[(N-n_s)/n_s] \qquad \text{Box 11.14}$$

$$n_s = (N-n_s)\exp(-\Delta E_s/2kT) \qquad \text{Box 11.15}$$

$$n_s \approx N\exp(-\Delta E_s/2kT) \qquad \text{Box 11.16}$$

11.5 EQUILIBRIUM DEFECT CONCENTRATIONS

We need to know how many point defects are present in thermal equilibrium. One of the simplest point defects that can occur in a crystal is a vacancy. In a binary compound, e.g., MgO, where only vacancies exist, we must have an equal number of each type of vacancy to maintain the stoichiometric formula.

At a given temperature there is an equilibrium concentration of point defects in a crystal. The implication of this statement is that a crystal containing point defects must have a lower free energy, G, than a corresponding crystal without any defects. From Chapter 3 we know that the change in free energy accompanying a process, such as the creation of vacancies, is given by

$$\Delta G = \Delta E - T\Delta S \qquad (11.1)$$

We can use ΔH or ΔE in Eq. 11.1. ΔH would refer specifically to the enthalpy to form a point defect. ΔE is the change in the internal energy on forming the defect. The internal energy of the disordered crystal is the energy of the perfect crystal plus nE.

E is strongly affected by the nearest neighbors and the interatomic bonding. It costs energy to produce point defects; we have to break bonds. The internal energy increases when the number of point defects increases. At the same time the entropy (randomness of the structure) increases, and the product $T\Delta S$ also increases. The change in entropy is complex and consists of terms due to the vibration of the atoms around the defects and terms due to the arrangement of defects in the crystal. This configurational entropy relies on statistics (how many) and mechanics (how it is moving); the subject is statistical mechanics. So even though it requires energy to create vacancies, overall G may decrease because of

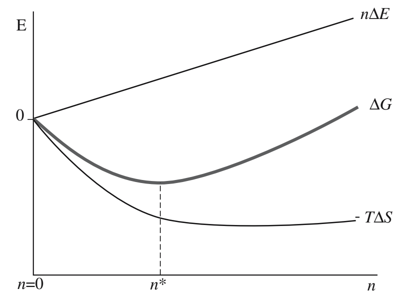

FIGURE 11.3 Schematic showing how the free energy decreases to a minimum as the number of vacancies increases.

the entropy contribution. The challenge is that to know ΔG, we have to know S, the entropy.

Figure 11.3 is a plot of $n\Delta E$, ΔS, and ΔG. From this plot you can see that introducing vacancies lowers the free energy of the crystal until an equilibrium concentration is reached; adding too many vacancies increases G again. At higher temperatures the equilibrium number of vacancies increases. The implications are important. In pure crystals we expect to find point defects at all temperatures above 0 K. Since these defects are in thermodynamic equilibrium, annealing or other thermal treatments cannot remove them.

Usually the concentration of point defects is controlled by impurities.

Dopant-induced defects will also be in thermodynamic equilibrium so that the overall equilibrium is controlled by the dopants.

Schottky defects do not change the composition of the material. The concentration of Schottky defects in a crystal is deduced using standard statistical mechanics that appears in most thermodynamics textbooks (because it is such a clear application of basic thermodynamics). For most ceramics, we just need the result from the calculation. Notice that charge is not mentioned and the derivation assumes a pure simple binary compound, like MgO or NiAl.

ΔE_s is the energy to form a Schottky defect, and k, the Boltzmann constant, is in J/°C. If ΔE_s is given in J/mol, then the number of Schottky defects in a unit volume is the same with R replacing k.

We will review how n_S is derived for a stoichiometric crystal AB. The derivation shows the importance of disorder. In this discussion, we have separated the equations from the text and mainly comment on the logic of the derivation.

If n_S is the number of Schottky defects per cubic centimeter in the crystal at T K, then we have n_S vacant cation sites and n_S vacant anion sites. In addition, in a crystal of this type there are N possible cation sites and N possible anion sites per cubic centimeter. We can determine the entropy change, ΔS, in a system of occupied and unoccupied sites by using the Boltzmann equation (Eq. Box 11.2). (Note that we are neglecting thermal entropy effects.) W is the number of ways of distributing n defects over N sites at random.

The number of ways we can distribute the n_S cation vacancies over the available sites in the crystal will be given by Eq. Box 11.4; we can write a similar expression for the anion vacancies as shown in Eq. Box 11.5.

Because AB is a stoichiometric compound we can write Eq. Box 11.6. The total number of ways, W_T, of distributing these defects is given by the product of W_A and, W_B, so we have Eq. Box 11.7, and therefore the change in entropy caused by introducing Schottky defects is given by Eq. Box 11.8, which we can rewrite as Eq. Box 11.9.

Now we use a standard trick for large N: Stirling's approximation (Eq. Box 11.10) eliminates the factorials, which leaves us with an expression for the entropy change associated with the formation of Schottky defects (Eq. Box 11.11).

Substituting Eq. Box 11.11 into Eq. 11.1 gives the overall change in Gibbs free energy for forming n_s pairs (Eq. Box 11.12). At equilibrium, the free energy change is a minimum with respect to the number of defects and is given by Eq. Box 11.13. We can thus differentiate Eq. Box 11.12 and set the result equal to zero so that after a little rearrangement we arrive at Eq. Box 11.14; in terms of the number of defects we write Eq. Box 11.15. The final step is another approximation: N is usually very much greater than n_S. So we arrive at Eq. Box 11.16 (i.e., Eq. Box 11.1).

We give some experimental values for the enthalpy of formation of Schottky defects in Table 11.4. We can use these numbers to calculate equilibrium defect concentrations as we have for NaCl in Table 11.5. The population of point defects is very low, but it is clear from Eq. Box 11.1 that vacancies are stable in the crystal at any temperature above absolute zero. Because energies for point defect formation in stoichiometric oxides such as

TABLE 11.4 The Formation Enthalpy of Schottky Defects in Some Compounds of Formula MX

Compound	*ΔE_s (10^{-19} J)*	*ΔE_s (eV)*
MgO	10.574	6.60
CaO	9.773	6.10
SrO	11.346	7.08
BaO	9.613	6.00
LiF	3.749	2.34
LiCl	3.397	2.12
LiBr	2.884	1.80
LiI	2.083	1.30
NaCl	3.685	2.30
NaBr	2.692	1.68
KCl	3.621	2.26
KBr	3.797	2.37
KI	2.563	1.60
CsBr	3.204	2.00
CsI	3.044	1.90

TABLE 11.5 Schottky Defect Concentrations in NaCl

Temperature (°C)	Temperature (K)	n_s/N	n_s (cm^{-3})
27	300	4.79×10^{-20}	2.14×10^{3}
127	500	2.56×10^{-12}	1.14×10^{11}
427	700	5.25×10^{-9}	2.34×10^{14}
627	900	3.63×10^{-7}	1.62×10^{16}

MgO and CaO are so high, the concentration of intrinsic defects is going to be very much less than impurity concentrations (usually ~100 ppm or greater). Even close to the melting temperature we find that only one or two sites in one billion are vacant due to intrinsic effects.

To give you an idea of what these numbers mean calculate the equilibrium number of vacancies that would be in a single crystal of MgO the size of the earth (r_{earth} ~6400 km) at room temperature; it is still a pretty small number.

Intrinsic vacancies are much more numerous in metals. For example, in a 1-cm^3 crystal of aluminum at room temperature there are about 9 billion vacancies. In a crystal of silicon in equilibrium at room temperature there are only about 1×10^{-18} intrinsic vacancies per cubic centimeter. This is considerably less than typical concentrations of extrinsic point defects (dopants) in silicon—about 0.0001%: another fortunate fact.

Vacancies are important for dislocation motion at high temperatures: the dislocations then move by climb—they absorb vacancies, which are more numerous at higher temperatures (the vacancies also move more easily at elevated temperatures).

Frenkel defects, like Schottky defects, also involve vacancies in the crystal structure. In this case, though, the vacancies exist on only one sublattice and the atom that should occupy the vacant positions is placed in an interstitial site in the crystal. The interstitial is a very special defect: it can exist on one or more nonequivalent sites, which are not normally crystal sites, so the Frenkel defect is not unique either.

The calculation of the number of Frenkel defects (n_f) in a crystal proceeds along lines that are similar to those for Schottky defects where we consider the number of ways of distributing N vacancies over N^* available interstitial positions.

The values for ΔE_f are usually quoted in J/mol and we therefore use the following equation:

$$n_f \approx (NN^*)^{1/2} \exp(-\Delta H_f/2RT) \tag{11.2}$$

Experimental values of the energy of formation of Frenkel defects in some oxide compounds are given in Table 11.6.

TABLE 11.6 Formation Enthalpy of Frenkel Defects in Some Compounds of Formula MX and MX_2

Material	ΔE_f (10^{-19} J)	ΔE_f (eV)
UO_2	5.448	3.40
ZrO_2	6.569	4.10
CaF_2	4.486	2.80
SrF_2	1.122	0.70
AgCl	2.564	1.60
AgBr	1.923	1.20
β-AgI	1.122	0.70

The actual type of defect found in the crystal will depend on the energy of formation. There may be either more Frenkel defects or more Schottky defects depending on which has the smaller energy of formation. As we have already mentioned, Frenkel defects are more likely to be important in crystals with open structures that can accommodate the interstitials without much lattice distortion.

Frenkel defects are the key to the photographic process. Photographic film contains crystals (called grains) of AgBr or I-doped AgBr. These are dispersed in gelatin, often along with various other chemicals known as sensitizers, on a thin plastic film (it used to be on glass). AgBr has a rocksalt structure, but rather than containing mainly Schottky defects (as we find in NaCl) it contains mostly cation Frenkel defects.

During irradiation with light, electrons are excited from the valence band to the conduction band. The band-gap energy is 2.7 eV, which corresponds to a wavelength of 460 nm (blue light). An electron will neutralize one of the silver interstitial ions. To produce what is known as a *latent image* it is necessary for silver atoms to form clusters of at least four atoms on the surface of the grains. The silver interstitials are mobile and can diffuse to the surface. The exact mechanism for cluster formation is not well understood, but this set of equations shows a possible sequence of reactions. The clusters act as catalysts that lead to the reduction of the AgBr grains in the presence of a suitable developer.

SILVER POINT DEFECTS

$$\begin{aligned}
&\mathrm{Ag}_i^{\bullet} + e' &&\rightarrow \mathrm{Ag}^x\\
&\mathrm{Ag}^x + e' &&\rightarrow \mathrm{Ag}'\\
&\mathrm{Ag}' + \mathrm{Ag}_i^{\bullet} &&\rightarrow \mathrm{Ag}_2^x\\
&\mathrm{Ag}_2^x + e' &&\rightarrow \mathrm{Ag}_2'\\
&\mathrm{Ag}_2' + \mathrm{Ag}_i^{\bullet} &&\rightarrow \mathrm{Ag}_3^x
\end{aligned}$$

11.6 WRITING EQUATIONS FOR POINT DEFECTS

In many instances we have to consider reactions that cannot be expressed within the normal chemical nomenclature. For example, if a dopant is incorporated into a crystal it can have profound effects upon the physical and chemical properties of the substance because of the defects that are necessarily introduced. However, defects do not occur in the balance of reactants expressed in the traditional equations, and so these important effects are lost to the chemical accounting system that the equations represent. Using the Kröger–Vink notation we can build up a *defect chemistry*, provided the normal rules for balancing chemical equations are preserved. In writing equations for point defect reactions we must obey the following rules:

- Maintain the ratio of the number of sites, e.g., in M_aX_b the ratio a/b is constant.
- V_M and M_x require sites; M_i does not require a new site.
- Maintain electrical neutrality overall.
- Surface sites are treated as if they are bulk sites (but could be treated separately).

Example: Point Defects in the Model Oxide M_2O_3

We will illustrate how to formulate the Schottky and cation Frenkel defect reactions using Kröger–Vink notation for the model oxide M_2O_3. For the Schottky defect reaction

$$2M_M^x + 3O_O^x \xrightarrow{M_2O_3} 2V_M''' + 3V_O^{\bullet\bullet} + M_2O_3 \qquad (11.3)$$

We imagine that the molecule of M_2O_3 condenses at the surface of the material or at another interface. In writing Eq. 11.3, we maintain the ratio of M-to-O sites and balance the charge.

For the cation Frenkel defect reaction

$$M_M^x + V_i^x \xrightarrow{M_2O_3} M_i^{\bullet\bullet\bullet} + V_M''' \qquad (11.4)$$

The vacancy on an interstitial site is simply that there is no interstitial initially. Notice that Eq. 11.4 can be written for a single cation. We do not need to consider the number of sites, etc. The anion Frenkel defect reaction would be similar, but is unlikely to occur in a material such as Al_2O_3.

11.7 SOLID SOLUTIONS

We consider solid solutions here because we can think of them as being formed by distributing a large number of point defects in a host crystal. As always, we must balance charge and be sure that the size of the "impurity" (guest) ion is appropriate to fit into the available site. If the impurity ions are incorporated in regular crystal sites the resulting phase is a substitutional solid solution. In an interstitial solid solution the impurity atoms occupy interstices in the crystal structure. The rules for substitutional solid solutions (the Hume–Rothery rules) can be summarized as follows. Note that the last two requirements are really very closely tied to the first two.

Rule	*Example*	*Comments*
Size factor (the "<15% rule")	CdS–CdSe r_S = 0.106 nm r_{Se} = 0.116 nm	If the size difference is outside this range then substitution will be limited or may not occur
Valence factor	Ni^{2+} in MgO	This is not the same in metals where we do not have to consider charge
Chemical affinity	Al^{3+} in MgO $MgAl_2O_4$	The possibility of a new phase can limit solid solubility
Structure type	MgO and NiO (same) SiO_2 and TiO_2 (similar)	The more similar the structures, the greater the solubility

We can summarize the comparison to metals:

- In metals, we have the same rule for the formation of alloys regarding size; the new factor in ceramics is the charge.
- In elemental semiconductors, the rules are approximately the same; these are ceramics with a large covalent component to the bonding. The valence difference is important in forming p-type and n-type semiconductors.
- We will see that the silica-based glasses are somewhere between elemental semiconductors and ionic materials, but introduce other challenges.

Solid-Solution Example 1: FeO_{1-x}

We saw in our discussion of phase diagrams that FeO never exists. Consider wüstite with composition $Fe_{0.95}O$. To compensate the charge of V_{Fe}'' (i.e., for electrical neutrality), we need to replace $3Fe^{2+}$ ions by $2Fe^{3+}$ ions. This is the important point: the Fe ions change their valence from 2 to 3. This point defect reaction happens automatically. You can imagine it occurring by dissolving $Fe_2O_3(s)$ in FeO(s).

$$Fe_2O_3(s) \xrightarrow{FeO} 2Fe_{Fe}^{\bullet} + 3O_O + V_{Fe}'' \qquad (11.5)$$

To exhibit nonstoichiometry, the cation must be able to exist in two different valence states (so this is not the same as nonequimolarity that we saw for spinel). The ions Fe, Co, and, to a lesser extent, Ni can do this; Mg cannot. An alternative way of writing this point defect reaction is to bring oxygen from the gas state into FeO.

$$2Fe_{Fe} + 2O_O + \frac{1}{2}O_2(g) \xrightarrow{FeO} 2Fe_{Fe}^{\bullet} + 3O_O + V_{Fe}'' \qquad (11.6)$$

TABLE 11.7 Composition and Structure of Wüstite

Composition	*Atom % Fe*	*Edge of unit cell (nm)*	*Density (g/cm³)*
$Fe_{0.91}O$	47.68	0.4290	5.613
$Fe_{0.92}O$	47.85	0.4293	5.624
$Fe_{0.93}O$	48.23	0.4301	5.658
$Fe_{0.945}O$	48.65	0.4310	5.728

What we are doing is oxidizing an oxide. Oxides in which the cation can change its valence state, in general, show a variation of composition with oxygen partial pressure. This dependence can even cause a change in the dimensions of the crystal lattice as shown in Table 11.7.

Solid-Solution Example 2: $Co_{1-x}O$

The nonstoichiometry of CoO is similar to that of wüstite but not so pronounced.

$$2\mathrm{Co_{Co}} + \frac{1}{2}\mathrm{O_2(g)} \xrightarrow{\mathrm{CoO}} 2\mathrm{Co^{\bullet}_{Co}} + \mathrm{O_O} + \mathrm{V''_{Co}} \qquad (11.7)$$

We can write this in an alternate form.

$$\frac{1}{2}\mathrm{O_2(g)} \xrightarrow{\mathrm{CoO}} \mathrm{O_O} + \mathrm{V''_{Co}} + 2\mathrm{h^{\bullet}} \qquad (11.8)$$

We can apply equilibrium thermodynamics to Eq. 11.8 and write the equilibrium constant, K, for the reaction

$$K = \frac{[\mathrm{O_O}][\mathrm{V''_{Co}}][\mathrm{h^{\bullet}}]^2}{(p\mathrm{O_2})^{1/2}} \qquad (11.9)$$

The concentration of anions on anion sites, $[\mathrm{O_O}]$, is essentially unity. For each vacancy on the Co sublattice, there are two electron holes; i.e., the concentration of vacancies on cation sites is twice the concentration of holes.

$$2[\mathrm{V''_{Co}}] = [\mathrm{h^{\bullet}}] \qquad (11.10)$$

$$K = \frac{[\mathrm{V''_{Co}}]4[\mathrm{V''_{Co}}]^2}{(p\mathrm{O_2})^{1/2}} \qquad (11.11)$$

Thus, there is a direct relationship between the oxygen partial pressure, $p\mathrm{O_2}$, and the extent of the nonstoichiometry.

$$(p\mathrm{O_2})^{1/6} \propto [\mathrm{V''_{Co}}] \qquad (11.12)$$

We can confirm this relationship by measuring the concentration of cation vacancies as a function of $p\mathrm{O_2}$. Brouwer diagrams (also known as Kröger–Vink diagrams) are used to represent defect concentrations as a function of $p\mathrm{O_2}$. The complete diagram for an oxide of the type MO is shown in Figure 11.4. There are three distinct ranges (I to III) that correspond to different values of the oxygen partial pressure.

It is important to keep in mind that these Brouwer diagrams, while useful, can provide an indication only of what might happen; the equilibrium constants needed for accurate calculations are not widely available. Similar diagrams can be drawn for nonoxide systems.

Solid-Solution Example 3: $Zn_{1+x}O$

The point defect chemistry of ZnO is different partly because there are two charge states of Zn, namely 2+ (as in ZnO) and 1+. When we heat ZnO in Zn vapor, we form a Zn-rich oxide, $Zn_{1+x}O$. Experimentally it is found that the excess Zn sits on interstitial sites (as you would expect from the crystal structure). We can write the basic defect equation for the singly charged interstitial as

$$\mathrm{Zn(g)} \rightarrow \mathrm{Zn}_i^{\bullet} + \mathrm{e'} \qquad (11.13)$$

We then write the equilibrium constant for the reaction in the usual way

$$K = \frac{[\mathrm{Zn}_i^{\bullet}][\mathrm{e'}]}{(P_{\mathrm{Zn}})} \qquad (11.14)$$

From the reaction equation we know that there is one electron for every Zn interstitial.

$$[\mathrm{Zn}_i^{\bullet}] = [e'] \qquad (11.15)$$

$$(P_{\mathrm{Zn}})^{1/2} \propto [\mathrm{Zn}_i^{\bullet}] \qquad (11.16)$$

The excess zinc could be incorporated as a divalent interstitial, which means that there is a second reaction possibility:

$$\mathrm{Zn(g)} \rightarrow \mathrm{Zn}_i^{\bullet\bullet} + 2\mathrm{e'} \qquad (11.17)$$

$$(P_{\mathrm{Zn}})^{1/3} \propto [\mathrm{Zn}_i^{\bullet\bullet}] \qquad (11.18)$$

Experimentally, we can measure the electrical conductivity as a function of the partial pressure of Zn, plot the results, and then look at the exponent. We find that $(P_{\mathrm{Zn}})^{1/2} \propto [\mathrm{Zn}_i]$, which indicates that the Zn interstitial has a charge of +1. When you heat ZnO to high temperatures and control the oxygen partial

POINT DEFECTS AND VARIATION IN pO_2

Range I—Low $p\mathrm{O_2}$

- The number of O vacancies increases.
- The oxide is reduced.
- $n = 2V_{\mathrm{O}}^{\bullet\bullet}$

Range II—Intermediate $p\mathrm{O_2}$

- Schottky defects dominate.

Range III—High $p\mathrm{O_2}$

- Increase in cation vacancies.
- The oxide is oxidized.

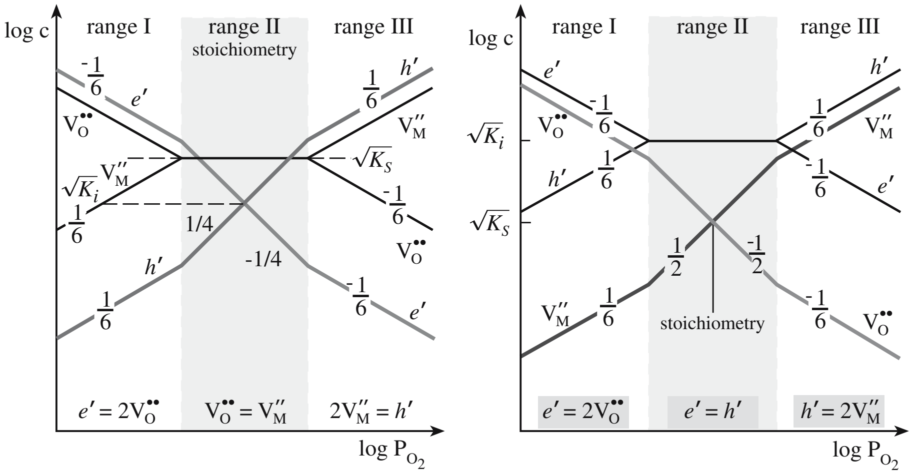

FIGURE 11.4 Examples of Brouwer diagrams in an oxide MO.

pressure the composition of the oxide will change. The equilibrium condition depends on pZn or pO_2.

ZnO is a particularly important ceramic semiconductor: its conductivity decreases with increasing pO_2. A major application for ZnO that makes use of its electrical properties is the varistor (*var*iable res*istor*).

Solid-Solution Example 4: ZrO_2 Doped with CaO

Doping ZrO_2 with Ca is a special example of a ZrO_2 solid solution. We can incorporate ~15% CaO in the structure to form Ca-stabilized cubic zirconia (CSZ). (CZ is the general abbreviation for cubic zirconia.) The special feature here is that the cubic (fluorite structure) phase is not stable at room temperature unless the ZrO_2 is heavily doped. However, we write the point defect equations as if it were always stable. The Ca^{2+} cation substitutes for the Zr^{4+} cation as shown in Figure 11.5. Since the charges are different, we must compensate with other point defects.

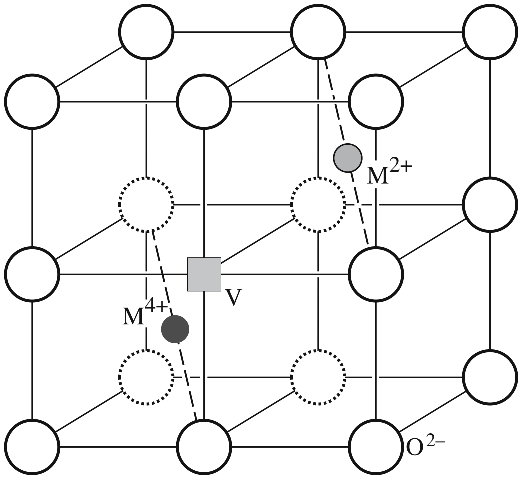

FIGURE 11.5 Accommodation of Ca^{2+} in ZrO_2. Note: this is just half of the unit cell.

$$CaO(s) \xrightarrow{ZrO_2} Ca_{Zr}'' + O_O + V_O^{\bullet\bullet} \qquad (11.19)$$

The cubic structure can also be stabilized using MgO (with essentially the same equation) or with Y_2O_3 (to form YSZ) when the equation is a little different. Later we will see other rare earths partly substituting for Y to give CZ a wide range of colors. The large number of oxygen vacancies makes CZ a material of choice for solid-oxide fuel cells. The material is an electrical insulator but an ionic conductor: oxygen moves quite rapidly through CZ.

11.8 ASSOCIATION OF POINT DEFECTS

This phenomenon often occurs when defect concentrations reach more than about 1 mol%. Then the defects are very close together and defect interactions may result in lowering the overall energy for defect formation and in defect clustering. Clustering is especially important in nonstoichiometric oxides such as $Fe_{1-x}O$ and UO_{2+x}.

This idea is special for ceramics because the effect is greatly enhanced due to the fact that the point defects are charged. In particular, since vacancies behave as if they are charged point defects, cation and anion vacancies can be strongly attracted to one another to form a defect complex denoted in the following equation by the parentheses.

$$V_{Na}' + V_{Cl}^{\bullet} = (V_{Na}', V_{Cl}^{\bullet}) \qquad (11.20)$$

We can write a law of mass action equation relating the fractional molar concentration of the vacancy pairs. At equilibrium

$$\frac{[(V_{Na}', V_{Cl}^{\bullet})]}{[V_{Na}'][V_{Cl}^{\bullet}]} = K \qquad (11.21)$$

We can then relate the equilibrium concentration of the vacancy pair to their free energy of formation, ΔG_{vp}, using

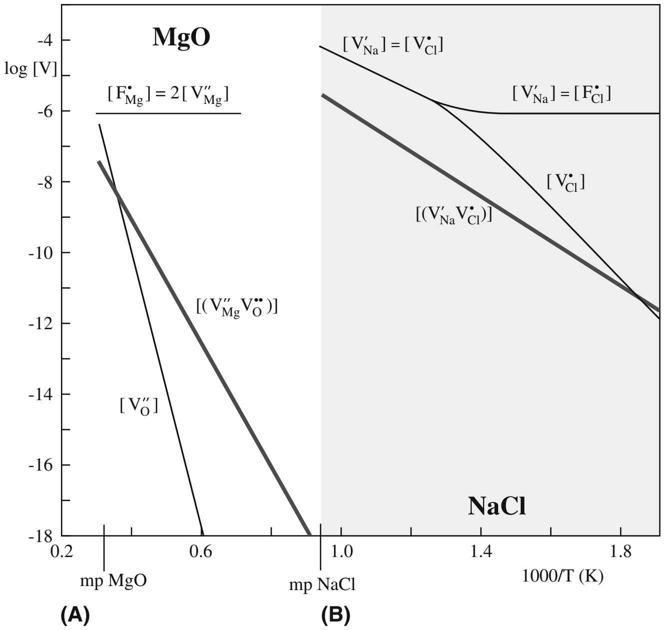

FIGURE 11.6 Plots of the vacancy concentrations varying with T^{-1}: (a) MgO; (b) NaCl.

$$\Delta G = -RT\ln K \tag{11.22}$$

Hence

$$\frac{[(\mathrm{V}_{\mathrm{Na}}^{\prime}, \mathrm{V}_{\mathrm{Cl}}^{\bullet})]}{[\mathrm{V}_{\mathrm{Na}}^{\prime}][\mathrm{V}_{\mathrm{Cl}}^{\bullet}]} = Z\exp\left(\frac{-\Delta G_{\mathrm{vp}}}{kT}\right) \tag{11.23}$$

In Eq. 11.23, Z is the number of distinct orientations of the defect complex; it accounts for the configurational entropy. For example, if the vacancies sit on adjacent sites in NaCl, Z would be 12. Figure 11.6 shows calculated concentrations of Schottky defects and vacancy pairs in NaCl and MgO.

Association of Point Defects due to Doping

Dissolving $\mathrm{CaCl_2}$ in NaCl increases the number of vacancies on the Na sublattice.

$$\mathrm{CaCl_2} \xrightarrow{\mathrm{NaCl}} \mathrm{Ca}_{\mathrm{Na}}^{\bullet} + \mathrm{V}_{\mathrm{Na}}^{\prime} + 2\mathrm{Cl}_{\mathrm{Cl}} \tag{11.24}$$

If the $\mathrm{Ca^{2+}}$ cations form a defect complex with such a vacancy (not now on adjacent sites), we can express this combination as

$$\mathrm{Ca}_{\mathrm{Na}}^{\bullet} + \mathrm{V}_{\mathrm{Na}}^{\prime} \rightarrow (\mathrm{Ca}_{\mathrm{Na}}^{\bullet}, \mathrm{V}_{\mathrm{Na}}^{\prime}) \tag{11.25}$$

A Second Example of Doping

Consider what happens when we dissolve a small amount of $\mathrm{Al_2O_3}$ in MgO.

$$\mathrm{Al_2O_3} \xleftrightarrow{\mathrm{MgO}} 2\mathrm{Al}_{\mathrm{Mg}}^{\bullet} + \mathrm{V}_{\mathrm{Mg}}^{\prime\prime} + 3\mathrm{O}_{\mathrm{O}} \tag{11.26}$$

The $\mathrm{Al^{3+}}$ sits on Mg sites requiring that vacancies be created on the Mg sublattice. Notice that the reaction can lead to precipitation of spinel, but then we are creating a new structure and, thus, new sites.

$$\mathrm{Mg}_{\mathrm{Mg}} + 2\mathrm{Al}_{\mathrm{Mg}}^{\bullet} + \mathrm{V}_{\mathrm{Mg}}^{\prime\prime} + 4\mathrm{O}_{\mathrm{O}} \xleftrightarrow{\mathrm{MgO}} \mathrm{MgAl_2O_4}\,(\mathrm{ppt}) \tag{11.27}$$

This equation emphasizes that four sites are required on each sublattice. We can write the law of mass action equation for the precipitation reaction

$$[\mathrm{Mg}_{\mathrm{Mg}}][\mathrm{Al}_{\mathrm{Mg}}^{\bullet}]^2\,\alpha\exp\left(+\frac{\Delta H_{\mathrm{ppt}}}{kT}\right) \tag{11.28}$$

ΔH_{ppt} is the enthalpy of formation.

$$[\mathrm{Al}_{\mathrm{Mg}}^{\bullet}] \propto 2[\mathrm{V}_{\mathrm{Mg}}^{\prime\prime}] \tag{11.29}$$

$$[\mathrm{V}_{\mathrm{Mg}}^{\prime\prime}]\alpha\left(\frac{1}{4}\right)^{1/3}\exp\left(+\frac{\Delta H_{\mathrm{ppt}}}{3kT}\right) \tag{11.30}$$

In summary, we ask again what is special for ceramics—why don't we do this for metals? The special feature is the formation of defect associates as a result of the strong attraction between oppositely charged point defects. Of course, defect associates and precipitates are related, and both are clearly volume defects.

11.9 COLOR CENTERS

The term color center is now applied to any defect, including an impurity, that produces color in an insulator. The original observation was made in Germany using X-rays to color alkali halides by producing F centers ("F" is for Farben: German for color). The color is the complement of what is absorbed by the color center. Colors are produced in most of the alkali halides following irradiation with ionizing radiation and are characteristic of the crystal not the radiation source used. These F centers are due to an electron or hole trapped at an anion vacancy site. Other types of color center can give other colors.

Several different types of color center have been identified and some of the common ones are illustrated in Figure 11.7 for the alkali halides. V centers are is cation vacancies with trapped holes. Color centers can also form in oxides such as MgO. When a single electron is trapped at an anion vacancy in MgO, the defect is positively charged and is

COLOR OF F CENTERS

The crystal and color:
NaCl—orange-brown/blue
KCl—violet/green
KBr—blue green/orange

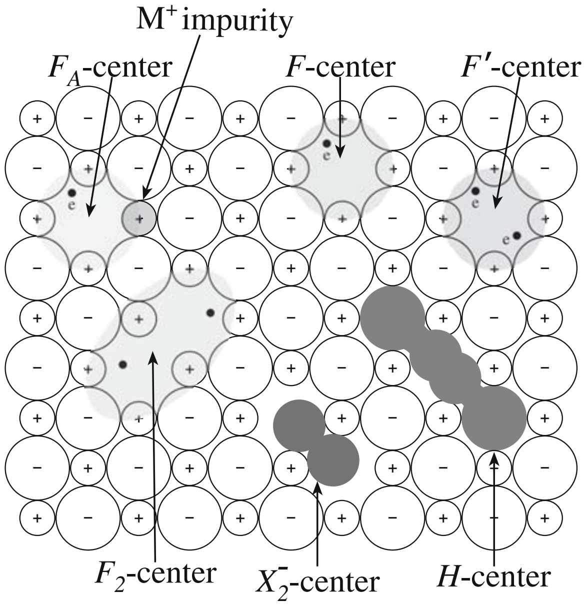

FIGURE 11.7 Schematics of color centers in alkali halides.

referred to as the F$^+$ center. An F center in MgO is a vacancy with two electrons.

IONIZING RADIATION CAUSING POINT DEFECTS

Ultraviolet light	E ~10 eV
X-rays	E = 10–100 keV
γ-rays	E = 1.25 MeV
High-E electrons	E = 100 keV–0 MeV
High-E protons	E = 2.20 MeV
Fast neutrons	E > 0.2 keV

Smoky quartz and amethyst are two natural examples of color centers. In smoky quartz a small amount of Al impurity atoms can substitute for Si. Because of the different valence of the two cations, electron neutrality is maintained by hydrogen ions. Ionizing radiation releases an electron from the $[AlO_4]^{5-}$ group, which is then trapped by the hydrogen ion:

$$[AlO_4]^{5-} + H^+ \rightarrow [AlO_4]^{4-} + H \qquad (11.31)$$

The color center is the $[AlO_4]^{4-}$ group, which is electron deficient, and we can think of it as having a trapped hole. In amethyst the color center is $[FeO_4]^{4-}$, which is due to Fe^{3+} impurities substituting for Si.

We will discuss more examples of how impurities give rise to color in Chapter 32, but it is interesting to note that NaCl crystals, doped to modify the color and increase the density of point defects, can be used for the NaCl color-center laser.

11.10 CREATION OF POINT DEFECTS IN CERAMICS

There are several ways that we can create point defects in ceramics. We have seen already that point defects can be produced in nonstoichiometric oxides, such as ZnO, by heating in either high or low pO_2. In Nd-doped yttrium aluminum garnet (YAG)—the most widely used solid-state laser crystal—the Nd is incorporated into the YAG structure (it substitutes for Y^{3+} in small amounts) when Nd_2O_3 is added to the YAG melt. A single crystal, which contains the Nd ions in solid solution, is pulled from the melt using the Czochralski method (see Chapter 29).

If a crystal is annealed at a sufficiently high temperature and for long enough, then the equilibrium concentration of vacancies will increase. Abrupt quenching of the material can "freeze in" unusually high concentrations of point defects.

Vacancies, usually Schottky defects or Frenkel pairs, can be produced as a result of ion bombardment in a process known as ion implantation. For example, many defects are formed in single crystal MgO when it is bombarded with high-energy Xe$^+$ ions. Figure 11.8 is a TEM image of MgO following implantation with 10^{14} Xe$^+$ cm^{-2} ions at 200 keV in a particle accelerator. The dark regions are typical of numerous point defect clusters. A large amount of work on ion implantation was performed to simulate what happens to materials inside a nuclear reactor when they are bombarded with fast neutrons (energies > 0.1 MeV). Now the topic is again important as ceramics are used to contain radioactive waste and we start to plan for fusion (rather than fission) reactors. Ion implantation is also the technique that is

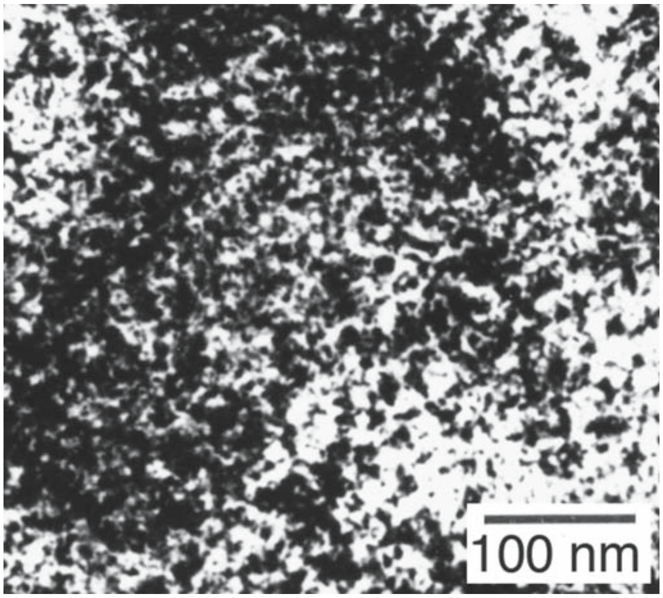

FIGURE 11.8 TEM image of defect clusters formed by ion implantation into MgO.

used to introduce dopant atoms to defined depths during the formation of semiconductor devices or to harden the surface of metals and ceramics.

Any form of ionizing radiation can be used to produce point defects in materials. The term ionizing radiation is used to describe radiation sources that generate electron or holes when interacting with matter.

THE DRIVING FORCE

If there is no bias (gradient in the chemical potential), then there is no driving force.

11.11 EXPERIMENTAL STUDIES OF POINT DEFECTS

It is very difficult to see individual point defects in a material. It is not that they are too small [transmission electron microscopy (TEM) has the resolution], but they have to be surrounded by atoms so we tend to see the surrounding atoms instead! Because point defects (especially substitutional atoms) are so important in determining the electrical properties of semiconductors, a great deal of work has been done to understand these defects. Figure 11.9 is a high-resolution scanning transmission electron microscopy (STEM) image showing columns of silicon atoms in a sample of Sb-doped Si. The brightest columns contain one Sb atom.

In Table 11.7 we showed how the lattice parameter of iron oxide changes as a function of the oxygen-to-iron ratio. The lattice parameter can be determined to a high precision using X-ray diffraction and the precise lattice-parameter method. It can then be correlated with the oxygen/iron ratio determined by chemical analysis.

We can also correlate point defect concentrations with a property measurement such as electrical conductivity, σ, or density, ρ. The density will show whether the dopant enters interstitially or is a replacement for the host ion.

NOTATION FOR DIFFUSION ACTIVATION

AB*	The activated complex
*K**	*K* for the activated complex
ν	The Debye atomic jump frequency, 10^{13} Hz
χ	The potential gradient per atom
N	Avogadro's number

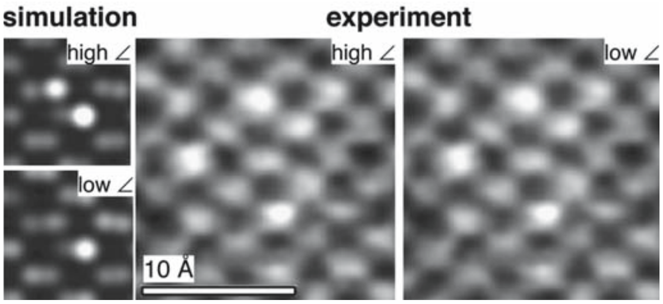

FIGURE 11.9 Atoms of Sb in Si; the *Z*-contrast image does not show where the atom sits in the column.

- An interstitial atom increases density.
- If a substitutional atom replaces a heavier atom the density may be reduced.

11.12 DIFFUSION

Diffusion occurs by atomic defects moving through the crystal; it is not a continuum process. We thus analyze diffusion as a statistical process. This is the kinetics part of the story. Diffusion of point defects is the key to understanding their properties. There are four basic mechanisms that can, in principle, occur; these are illustrated in Figure 11.10.

- Direct exchange (difficult and not energetically probable)
- Ring mechanism (cooperative motion: possible but not demonstrated)
- Vacancy mechanism (most common, dominant in metals)
- Movement of interstitials (occurs when you can put interstitials in); an example is Zn in ZnO.

Diffusion is a thermally activated process so we plot the free energy versus distance. The region of high energy in this plot corresponds to an "activated complex." It occurs because there is an activation energy barrier to diffusion. We can use reaction rate theory to understand diffusion behavior. First we consider two features of the process.

- Each step in a difficult process is relatively simple; if it appears more difficult, we break it into smaller steps.

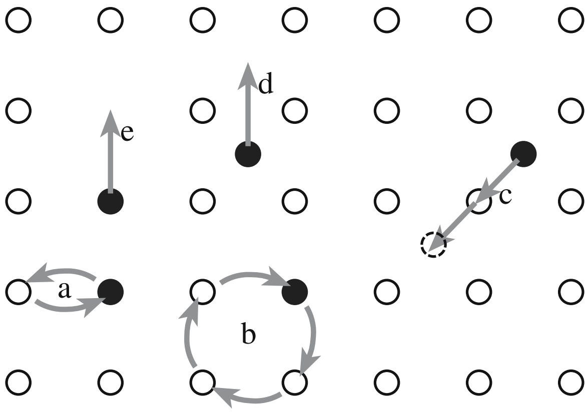

FIGURE 11.10 Possible mechanisms occurring during diffusion. (a) Exchange, (b) ring, (c) vacancy, (d) interstitial, (e) knock-on.

- The reaction path of each step, such as an individual atom jump, involves an activated complex, which gives an energy maximum.

For the chemical reaction (the route of transition of the activated complex, AB*, into the product)

$$A + B \rightarrow AB^* \qquad (11.32)$$

The concentrations are related by the equilibrium constant *K**.

$$\frac{[AB^*]}{[A][B]} = K^* \text{ and } \Delta G^* = -RT \ln K^* \qquad (11.33)$$

The reaction rate, k, is $\nu[AB^*]$, which can be written as $\nu K^*[A][B]$. So the reaction rate for unit concentration is just νK^*. If there is no net flow to the right or left, because each time a point defect jumps to the right one also jumps to the left, we need to add a driving force, i.e., a potential gradient, to produce a flux. Now the flux of atoms is the rate in the forward direction minus the rate in the backward direction. ΔG forward is $\Delta G^* - \frac{1}{2}\chi\lambda$ for unit concentration forward. The analysis makes two assumptions:

1. The activity coefficient is unity.
2. *D* does not vary with the composition.

The diffusion coefficient does have a strong dependence on temperature.

$$D = D_0 \exp\left(\frac{-Q}{RT}\right) \qquad (11.34)$$

The diffusion equations are summarized by Fick's laws. One assumption that is explicit in the derivation of these equations is that there is a well-defined value of *D*. We will see later that because of the high concentrations of defects in ceramic materials, there will be many important situations in which this assumption is not valid; i.e., *D* often does depend on the composition.

FICK'S LAWS

Fick's first law	$J = -D\left(\frac{\partial c}{\partial x}\right)$
Fick's second law	$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}$

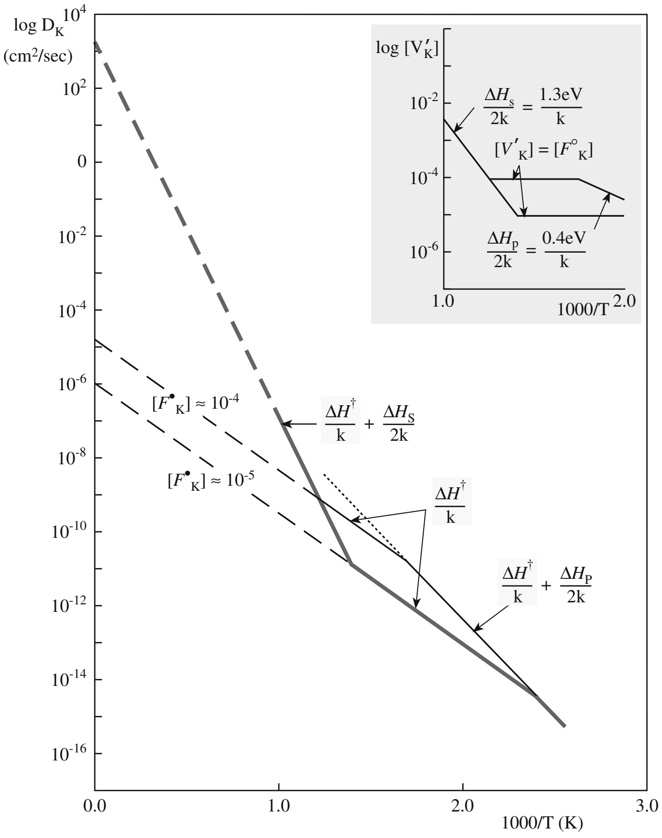

FIGURE 11.11 Diffusion coefficients in the intrinsic and extrinsic ranges. KCl with 10^{-4} and 10^{-5} atomic fraction divalent cation impurities.

11.13 DIFFUSION IN IMPURE, OR DOPED, CERAMICS

Now you can compare diffusion in stoichiometric oxides with diffusion in nonstoichiometric oxides. The first question concerns the diffusing species when two species are present. For many ceramics ΔE_s is large. For MgO, Al_2O_3, and B_2O_3, ΔH_s is ~600 kJ/mol (~6 eV/formula unit). Thus the intrinsic concentration of vacancies, [V], is ~10^{-4} at 2000°C. As we lower the temperature, *T*, the equilibrium concentration of vacancies, [V], decreases too.

Figure 11.11 shows a plot of intrinsic and extrinsic diffusion coefficients. The first part of this plot is almost never observed: most ceramics are never this pure (the exceptions are the semiconductors, Si and Ge, and a few others). This means that an impurity concentration of between 1 part in 10^5 (10 ppm) and 1 part in 10^6 (1 ppm) is sufficient to dominate the vacancy concentration. There is always an attraction between oppositely charged point defects so that they will associate with one another, which will in turn affect their effective concentrations. (Consider how far apart point defects can be if the concentration is 1 in 10^6—~20 nm, maximum.)

We will now consider examples that illustrate special features of diffusion in ceramics.

$CaCl_2$ in KCl shows the effect of changing the charge on the cation.

ZrO_2 doped with CaO is special because the structure of CSZ requires the Ca to be present and CZ is a fast conductor of oxygen.

$Zn_{1+x}O$ is special because interstitials dominate and can have different charges.

FeO_{1-x} and CoO_{1-x} are special because neither oxide is ever stoichiometric because the cation is always present in two charge states.

CuO_{1-x} is special because it is oxygen deficient but behaves differently from $Zn_{1+x}O$; it shows the importance of crystal structure.

In each case, we clearly have to understand the defects before we can understand their behavior.

Diffusion Example 1: $CaCl_2$ in KCl

When we dope KCl with $CaCl_2$ diffusion of K^+ cations occurs by an interchange between K^+ cations and cation vacancies. The concentration of cation vacancies is given by

$$[V_K'] = \exp\left(\frac{\Delta G_S}{2kT}\right) \qquad (11.35)$$

ΔG_S is the Schottky formation energy. The diffusion coefficient can then be written as

$$D_K^\bullet = [V_K']\gamma\lambda^2\nu\exp\left(-\frac{\Delta G}{kT}\right) \qquad (11.36)$$

Here, γ is an orientation factor. The Gibbs free energy can then be rewritten as usual:

$$\Delta G = \Delta H - T\Delta S$$

So the diffusion coefficient is

$$D_K^\bullet = \gamma\lambda^2\nu\exp\left(\frac{(\Delta S_s/2)+\Delta S^*}{k}\right)\exp\left(\frac{(-\Delta H_s/2)+-\Delta H^*}{kT}\right) \qquad (11.37)$$

In this equation, $-\Delta H^*$ is the enthalpy of motion of an atom while ΔH_S is the enthalpy of formation of a Schottky defect. The constants are all well known: λ is ~0.2 nm, γ is ~0.1, and the Debye frequency, ν, is $10^{13}\,s^{-1}$. $\Delta S_s/2k$ and $\Delta S^*/2k$ are small positive numbers. Hence we can estimate that

$$\gamma\lambda^2\nu\exp\left(\frac{(\Delta S_s/2)+\Delta S^*}{k}\right) \cong 10^2 \quad \text{or} \quad 10^3 \qquad (11.38)$$

Why is diffusion controlled by a small amount of impurity in KCl? We can write an equation describing what happens when we dissolve $CaCl_2$ in KCl.

$$CaCl_2(s) \xrightarrow{KCl} Ca_K^\bullet + V_K' + 2Cl_{Cl} \qquad (11.39)$$

Once $CaCl_2$ is dissolved and you change T, nothing much will change because you have fixed the concentration of vacancies.

$$[V_k'] = [Ca_K^\bullet] = \text{fixed} \qquad (11.40)$$

At high T we have a high intrinsic vacancy concentration; at low T we have a low intrinsic vacancy concentration. At low T the diffusion coefficient is controlled by the starred (dopant) quantities.

$$D_K^\bullet = \gamma\lambda^2\nu[Ca_K^\bullet]\exp\left(\frac{\Delta S^*}{k}\right)\exp\left(\frac{-\Delta H^*}{kT}\right) \qquad (11.41)$$

The slope is determined only by $\Delta H/kT$. If we increase [Ca] to 10^{-4}, the curve will move up.

In general, even at high T, $[V_k']$ due to thermal equilibrium is less than $[V_k']$ due to $[Ca_K^\bullet]$. This is the case for most ceramics.

- Intrinsic diffusion is controlled by defects that are present at thermal equilibrium.
- Extrinsic diffusion is controlled by dopants (impurities) that are present.

The intrinsic part is not usually observed because an impurity concentration of between 1 in 10^5 and 1 in 10^6 is sufficient to control the vacancy concentration. There is another complication we should consider: as we saw in Section 11.8, there is attraction between $[Ca_K^\bullet]$ and $[V_k']$ that results in defect complexes. The concentration of these complexes depends on the strength of this attraction; formation of these complexes will probably reduce the mobility of the point defects and thus decrease diffusion rates.

Diffusion Example 2: ZrO_2 Doped with CaO

We noted above that if ZrO_2 is doped with 15 at% CaO we can stabilize the cubic phase. The resulting material can be described as $Zr_{0.85}Ca_{0.15}O_{1.85}$. Experimentally we find that D_O decreases as the concentration of oxygen vacancies increases. Hence we know that oxygen diffuses by a vacancy mechanism.

In UO_2, which has the same fluorite structure, oxygen diffuses by an interstitial mechanism. (Look back to the discussion of the structure of fluorite to see why this might be so.) D_O increases as the concentration of oxygen vacancies increases. In such an interstitial mechanism, the interstitial moves onto a regular site and *bumps* the ion occupying that site onto an interstitial site. If you study diffusion using tracer atoms, the tracer atom has now temporarily stopped moving, but the interstitial ion may repeat the process and move on through the crystal. The process is thus rather complicated. A typical activation energy for this process might be ~120 kJ/mol.

Notice that the crystals in these two examples both have the fluorite structure, but the diffusion mechanism that operates is very different because the point defects involved are different. Experimental data for diffusion in several oxides are plotted in Figure 11.12.

Diffusion Example 3: $Zn_{1+x}O$

In our discussion of point defects in nonstoichiometric ZnO, we noted that excess zinc is present as interstitials; the oxide is metal rich.

$$Zn(g) \rightarrow Zn_i^{\bullet} + e' \qquad (11.13)$$

Figure 11.13 shows a plot for D_{Zn}: the steeper the slope, the larger the activation energy. Varying the partial pressure of Zn, P_{Zn}, has a large effect on D_{Zn} in ZnO because, as we saw above, P_{Zn} is proportional to the interstitial concentration. Thus, as P_{Zn} increases, the density of interstitials increases and the diffusion of Zn through the structure increases. This result again emphasizes the difference between a metal and its oxide: in metallic Zn, we do not care about the partial pressure of the Zn; the Zn atoms diffuse by a vacancy mechanism, but the concentration of vacancies does not depend on P_{Zn}.

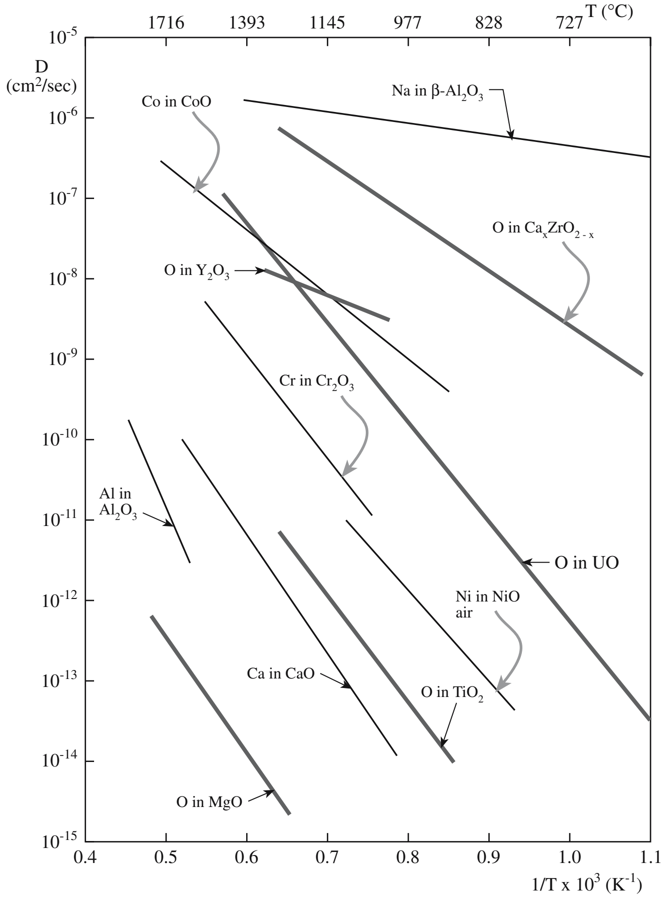

FIGURE 11.12 Arrhenius plots for different oxides: diffusion coefficients varying with T^{-1}.

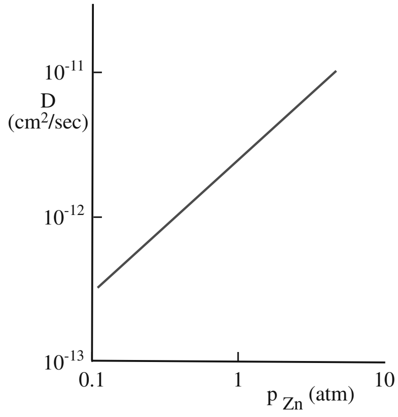

FIGURE 11.13 Plot for D_{Zn} as it varies with T^{-1}.

Diffusion Example 4: FeO_{1-x} and CoO_{1-x}

As we noted above, the group of binary oxides, FeO, NiO, CoO, and MnO (all have rocksalt structure), is metal deficient. The most important of these is $Fe_{1-x}O$. FeO does not exist (i.e., it is not stable) and can contain as many as 15 at% vacancies due to the equilibrium concentration of Fe^{3+} cations; we saw earlier that this equilibrium occurs by the cation changing its valence: Fe^{2+} changes to Fe^{3+}. We can write an equation to dissolve excess oxygen in $Fe_{1-x}O$. We can generalize the equation we wrote for FeO and CoO as

$$2M_m + \frac{1}{2}O_2(g) \rightarrow O_0 + V_M'' + 2M_M^{\bullet} \qquad (11.42)$$

Here M is the metal and $M_M^{\bullet}$ represents an Fe^{3+} cation sitting on an Fe^{2+} site. The rest of the analysis follows from our discussion of CoO. The important point is that diffusion depends on pO_2 as illustrated for several oxides in Figure 11.14. There is an energy associated with the reaction and an energy associated with the motion of defects, but point defects are always present in FeO.

Good experimental observations are available for CoO. We can show that the diffusion coefficient for Co can be expressed as

$$D_{Co} = \gamma\nu\lambda^2 \left(\frac{1}{4}\right)^{1/3} pO_2^{1/6} \exp\frac{\Delta S_0}{3k} \exp\frac{\Delta S^*}{k} \exp\left(-\frac{\Delta H_0^{\bullet}}{3kT} - \frac{\Delta H^*}{kT}\right) \qquad (11.43)$$

Experimentally, we find that D_{Co} is proportional to $pO_2^{1/n}$ with the value of n depending on T.

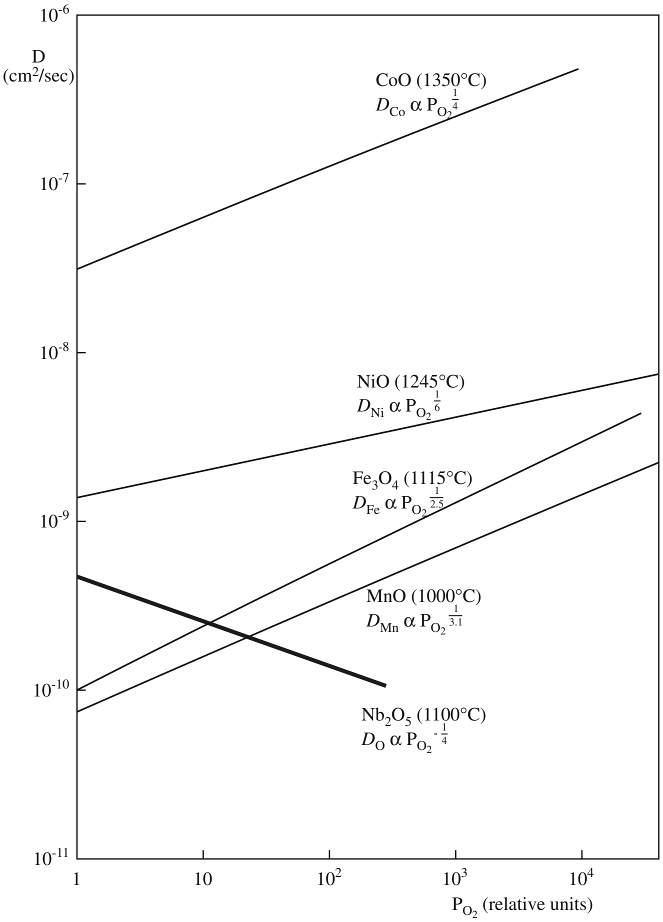

FIGURE 11.14 Diffusion coefficients varying with pO_2.

- $n = 3.1$ at 1000°C
- $n = 3.3$ at 1200°C
- $n = 3.6$ at 1350°C

We expected D_{Co} to be proportional to $pO_2^{1/6}$. This means that as we increase pO_2, the value of D_{Co} increases faster with vacancy concentration than we had predicted. The reason is that the defects are charged and that there is an association of defects (see the discussion of KCl above). We must consider two possibilities. A single cation vacancy might form a complex with one hole or with two holes. In the first case we write

$$\frac{1}{2}\mathrm{O}_2^* = [\mathrm{h}^\bullet] + [(\mathrm{V}_{\mathrm{Co}}''\mathrm{h}^\bullet)] + \mathrm{O}_0 \tag{11.44}$$

$$\frac{[\mathrm{h}^\bullet][\mathrm{V}_{\mathrm{Co}}''\mathrm{h}^\bullet]}{p\mathrm{O}_2^{1/2}} = K_1 \tag{11.45}$$

$$[\mathrm{h}^\bullet] = [(\mathrm{V}_{\mathrm{Co}}''\mathrm{h}^\bullet)] \tag{11.46}$$

so that $[(\mathrm{V}_{\mathrm{Co}}''\mathrm{h}^\bullet)]$ is proportional to $p\mathrm{O}_2^{1/4}$ and n is 4.

If two holes are involved in the complex, then

$$[(\mathrm{h}^\bullet\mathrm{V}_{\mathrm{Co}}''\mathrm{h}^\bullet)] \propto p\mathrm{O}_2^{1/2} \tag{11.47}$$

Now, $D_{Co} \propto pO_2^{1/2}$ and n is 2. Experimentally we find n is ~3 so we can thus conclude that the vacancy must be associated with between one and two electron holes. It is not obvious that we actually make another assumption in this treatment. We perform a diffusion experiment, measure D_{Co}, and then assume that a vacancy in a complex moves at the same velocity as the free vacancy. So we should say that if the association of a vacancy and holes does not affect the movement of vacancies, then each vacancy must have between one and two electron holes associated with it.

Finally, we comment on the mobility (B_V) of the vacancy; B_V is proportional to Γ_V, the jump frequency. Similarly, for holes B_h is proportional to Γ_h. In CoO all the electrical current is carried by the electron holes (i.e., it is a p-type semiconductor) so that the mobility of holes is much greater than the mobility of the vacancies. In CoO we can say that

$$B_h \gg B_V$$

$$\Gamma_h \gg \Gamma_V$$

So when a vacancy jumps, holes follow. Thus a hole/vacancy pair will move at a rate determined by the jump frequency of the vacancy, the slower moving defect.

Diffusion Example 5: CuO_{1-x}

CuO is an example of a group of oxygen-deficient oxides, MO_{1-x}, where the nonstoichiometry is accommodated by oxygen vacancies. We write the equilibrium equation in the usual way.

$$\mathrm{O}_0 \rightarrow \frac{1}{2}\mathrm{O}_2(\mathrm{g}) + \mathrm{V}_{\mathrm{O}}^{\bullet\bullet} + 2\mathrm{e}' \tag{11.48}$$

We can determine the equilibrium constant and relate this to the Gibbs free energy ΔG.

$$[\mathrm{V}_{\mathrm{O}}^{\bullet\bullet}] \cong \left(\frac{1}{4}\right)^{1/3} (p\mathrm{O}_2)^{-1/6} \exp\left(-\frac{\Delta G_{\mathrm{O}}}{3kT}\right) \tag{11.49}$$

$$D_{\mathrm{O}} = \gamma\lambda^2\nu[\mathrm{V}_{\mathrm{O}}^{\bullet\bullet}]\exp\left(-\frac{\Delta G^*}{kT}\right) \tag{11.50}$$

$$D_{\mathrm{O}} = \gamma\lambda^2\nu\left(\frac{1}{4}\right)^{1/3} (p\mathrm{O}_2)^{-1/6} \exp\left(-\frac{\Delta G_{\mathrm{O}}}{3kT}\right)\exp\left(-\frac{\Delta G^*}{kT}\right) \tag{11.51}$$

If you increase the pO_2, then the diffusion coefficient will decrease. This is found for a very wide temperature range and implies that diffusion is a complex process. It is interesting to compare this case to that of ZnO. Both cations exist in the 2+ or 1+ charge state, but the crystal structures and point defect chemistry are very different.

11.14 MOVEMENT OF DEFECTS

We often think that only cations move in ionic materials because so much work has been carried out on oxides or halides where the anion is much larger than the cation, for example, MgO, Al_2O_3, and NaCl. Ceramics that are being used for electronic applications often contain heavier (hence larger) cations so that the assumption is not necessarily valid. In solid-oxide fuel cells, it is the oxygen ion that diffuses.

Most perfect ceramic crystals have more unoccupied volume than there is in metals, so they tend to be less dense. There is thus more open volume through which point defects can move.

Covalent bonds are directional in character and must be "broken" as defects move. We will see that dislocations in Si do not move easily at room temperature so the material does not easily deform plastically and is brittle at this temperature. This simple fact has many consequences for the electronics industry. Where would silicon technology be if dislocations moved readily under an applied stress? The same consideration applies to point defects except that interstitials in Si do not need to break bonds!

The charge can affect the concentration of point defects. FeO contains at least 5% vacancies on the cation sublattice so the formula should be written as $Fe_{1-\delta}O$ [or better $Fe(II)_{1-2\delta}Fe(III)_{\delta}O$]. The superconductor YBCO must be off stoichiometric to optimize its superconducting properties ($YBa_2Cu_3O_{7-\delta}$ or $YBa_2Cu_3O_{6+x}$).

CORROSION

Diffusion of oxygen in oxidation scales occurs along grain boundaries. Corrosion of metals is controlled by formation and diffusion of point defects in the ceramic. Corrosion of polycrystalline ceramics also occurs most quickly along grain boundaries.

The movement of grain boundaries is very important in ceramics because of the method of processing and their high-temperature applications. However, the properties of grain boundaries are very different from the bulk material and they may contain a second phase. Grain boundaries in ceramics are not as dense as in metals because the bonding is different (remember the open space in Si and Al_2O_3) so diffusion of point defects along, and across, these interfaces is much faster than in the bulk. There may also be an excess charge of one sign at the grain boundary plane, which must be compensated in the bulk (the space-charge region) and will again affect the diffusion of charged point defects.

11.15 DIFFUSION AND IONIC CONDUCTIVITY

Some ceramics are important because they are ionic conductors. The band gap energy, E_g, of these materials is large (typically > 5 eV) and the only mechanism for conduction of charge is by the movement of ions. In some cases the rate of movement of the ions is very rapid and large conductivities are possible. In this section we will mention one application of ionic conductivity; we will consider more examples and details in Chapter 30. The ionic conductivity, σ, is

$$\sigma = \sum_i q_i N \mu_i \tag{11.52}$$

Here, q_i is the effective ionic charge, μ_i is the mobility, and N is the number of mobile defects.

Mobile defects contributing to ionic conductivity include Schottky defects. Figure 11.15 shows the example of Na^+ motion in NaCl. Two different routes are shown for how the Na^+ ion might reach the vacant site. The direct route, although shorter, is unlikely because the two Cl^- ions will be very close together ("touching") creating a large energy barrier for the migrating Na^+ ion to overcome. Route 2 is more likely. The Na^+ ion first passes through the face of the octahedron, then into the vacant tetrahedral site between the octahedra, then through the face of the opposite octahedron and into the vacant site. We cannot prove that this is what happens, but we expect the ion to follow the lowest energy pathway, which looks like route 2.

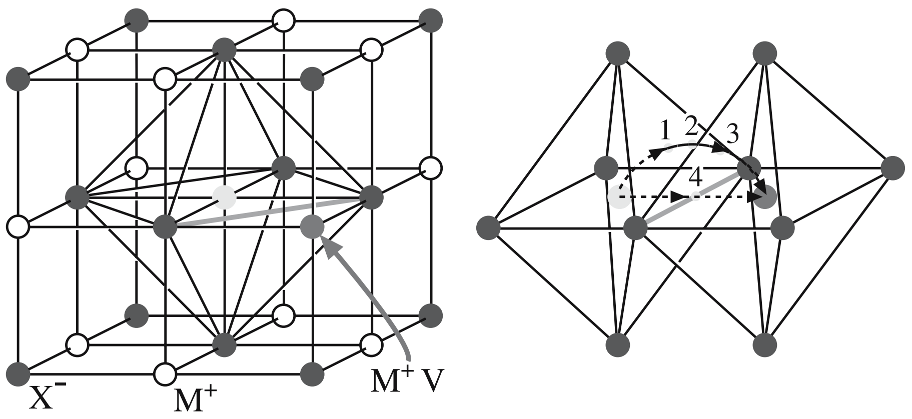

FIGURE 11.15 A probable diffusion mechanism in NaCl.

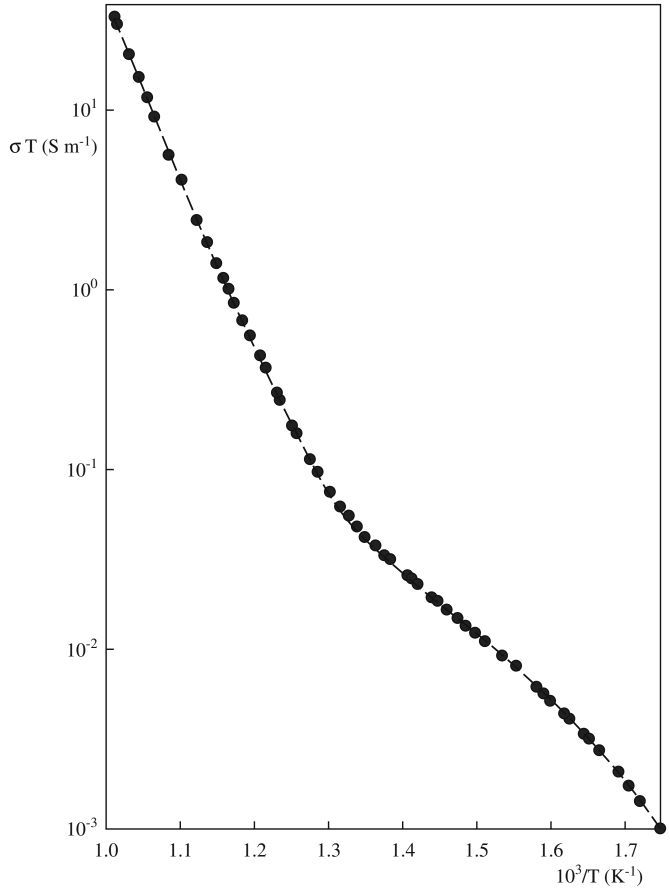

FIGURE 11.16 Ionic conductivity of NaCl varying with T^{-1}.

If ionic conductivity and diffusion occur by the same mechanism then μ and σ are related to D by the Nernst–Einstein equations

$$\mu = \left(\frac{q}{kT}\right)D \qquad \sigma = \left(\frac{Nq^2}{kT}\right)D \tag{11.53}$$

In using Eq. 11.53 we have to be aware of the following points:

- Some defects may contribute to D but not to σ (paired cation and anion vacancies in the alkali halides, for example).
- Diffusion may occur along grain boundaries and dislocations (these are more rapid paths than bulk diffusion).
- Electronic contributions to σ are low (this is the case in wide band gap materials at low T).

For a simple vacancy transport mechanism both σT and D would obey Arrhenius relationships

$$\sigma T = A\exp(-Q/kT) \tag{11.54}$$

$$D = D_0\exp(-Q/kT) \tag{11.55}$$

A and D_0 are temperature-independent constants and Q is the activation energy.

Figure 11.16 shows the variation in σT with T for a relatively pure sample of NaCl. As you can see, the relationship is complicated. The activation energy Q consists of two components, one due to the creation of the defect and one due to its motion, and these may have different temperature behaviors. For NaCl at temperatures above 550°C there is good correlation between D and σ. Below 550°C the trend is similar, i.e., there is the same change in slope, but the D values are greater than σ because of the presence of impurities. We can summarize some important characteristics of the plot.

- Stage I′—Anion disorder
- Stage I—Intrinsic conductivity; cation migration dominates
- Knee region—where the slope changes: separates the intrinsic dominated region from the impurity region
- Stage II—Extrinsic conductivity dominated by cation vacancies
- Stage III—Vacancy-impurity association

An application that makes use of ionic conductivity is the ZrO_2 oxygen sensor illustrated in Figure 11.17. These sensors are used, for example, to measure the pO_2 in automobile exhaust gases or in the gases present in the carburizing process used to harden steel components. Polycrystalline CZ (stabilized with either Ca^{2+} or Y^{3+} substituting for Zr^{4+}) is the electrolyte. At the operating temperature of the sensor (~900°C), CZ is a good oxygen ion conductor. The substitution for Zr^{4+} results in the

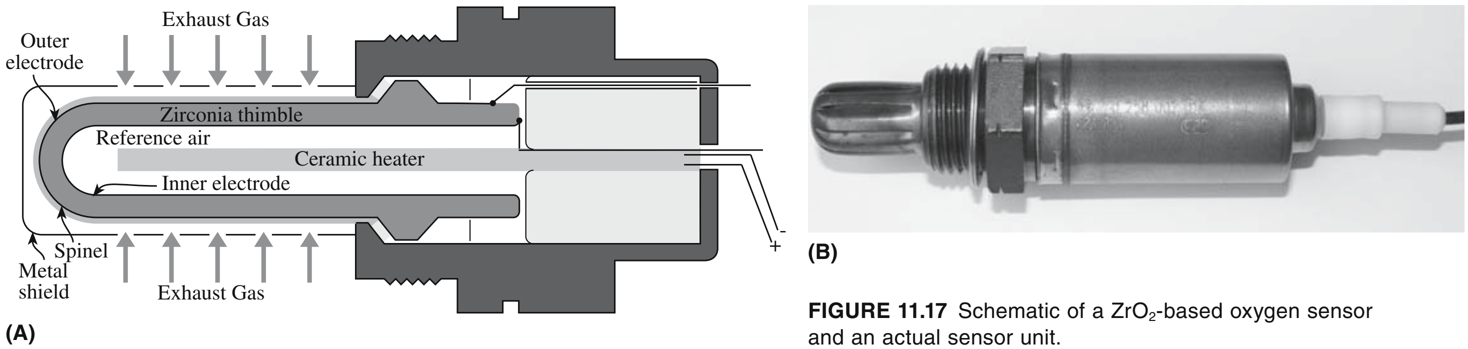

FIGURE 11.17 Schematic of a ZrO_2-based oxygen sensor and an actual sensor unit.

formation of oxygen vacancies and these lead to high ion mobilities on the oxygen sublattice. Solid-oxide fuel cells use the same electrolyte or another with the same structure and work on a similar principle.

11.16 COMPUTING

Computer modeling is becoming a more widespread method of studying point defects because computers are now capable of addressing large numbers of atoms. In ceramics, although the structural aspects of point defects may be short-range, interactions tend to be long-range because of Coulombic forces. Atomistic simulation calculations usually require several steps, and the actual energy may depend on how the structure is able to relax around the defect.

- Choice of a modeling approach
- Choice of an interatomic potential describing the atoms or ions
- A starting model for the point defect or defect cluster

This modeling approach is discussed in more detail later. The size of the modeled crystal must be considered and its other physical properties modeled. Modeling the potential is already much more difficult when the material is a compound rather than an elemental solid and we add charge. The choice of the starting model is important since calculations may suggest structures that are metastable but are not the lowest-energy configuration. Table 11.8 compares calculated values and experimental values of defect formation energies for which the agreement is good.

TABLE 11.8 Comparison of Calculated and Observed Defect Formation Energies in Ionic Crystals

Compound	*Defect*	*Calculated energy (eV)*	*Experimental energy (eV)*
NaCl	Schottky pair	2.32	2.30
CaF_2	Frenkel pair	2.75	2.7
MgO	Schottky pair	7.5	5.7
ZnO	Cation Frenkel pair	2.51	—
AgCl	Cation Frenkel pair	1.4	1.45

CHAPTER SUMMARY

This chapter is one of the most basic in ceramics and for the student, the easiest to learn. You need to know the different types of point defects and their names, the thermodynamic principles leading to the calculation of point defect concentrations, and how point defects make diffusion possible. The next level of complexity concerns how point defects interact with one another. We touch on much of this in the examples of real materials where we bring in the fact that the point defects are often (usually) charged. The importance of the topic is that point defects not only *affect* the properties of materials but, in many cases, *determine* and *control* the properties that interest us. An interesting question is how does this discussion change for nanoceramics—are point defects in nanoceramics important? The initial answer is that the surface (a two-dimensional defect) dominates everything when the particle is small, but one point defect in a nanoparticle can be a very high concentration!

PEOPLE IN HISTORY

Boltzmann, Ludwig (1844–1906) lived in Vienna and eventually committed suicide supposedly because his theory was ridiculed by some colleagues. He died before his theories were experimentally proven correct.

Fick, Adolf Eugen (1829–1901) is best known for his laws, which he first applied in 1855. He was actually also a physiologist and developed a technique for measuring cardiac output, but also was one of the first to experiment with contact lenses in 1887.

Frenkel, Jacov Il'ich was born 10 February 1894 in Rostov-on-Don, Russia and died 23 January 1952 in Leningrad, Russia. He worked on dislocations, twinning, and much more.

Kröger, Ferdinand Anna born Amsterdam (1915–2006) was the lead author on the paper that gave the Kröger–Vink notation. He retired as Professor from USC in ~1990.

Schottky, Walter was born July 23, 1886 in Zürich and died March 4, 1976 in Pretzfeld, Germany. He worked under Max Planck in Berlin. He discovered both the electron hole and the Schottky defect.

Wedgwood, Thomas (son of the famous potter) and Davy, Sir Humphrey, in 1802, reported in a communication to the Royal Institution in London a photographic process using AgCl to record an image.

GENERAL REFERENCES

Hayes, W. and Stoneham, A.M. (1985) *Defects and Defect Processes in Nonmetallic Solids*, John Wiley & Sons, New York. Chapter 3 (pp.106–168). Very nice but more advanced than our treatment.

Kröger, F.A. (1964) *The Chemistry of Imperfect Crystals*, North-Holland, Amsterdam.

Kröger, F.A. and Vink, H.J. (1956) "Relations between the concentrations of imperfections in crystalline solids," *Solid State Phys.* **3**, 307. The original proposal of the notation that is now universally used to describe charged point defects. This is an invaluable paper when you have time to study it. The official notation is given in the IUPAC *Red Book on the Nomenclature of Inorganic Chemistry*, Chapter I-6.

Smyth, D.M. *The Defect Chemistry of Metal Oxides*, Oxford University Press, Oxford 2000. Clear and at the right level.

Swalin, R.A. (1972) *Thermodynamics of Solids*, 2nd edition, John Wiley & Sons, New York. Chapter 11.

Tilley, R.J.D. (1987) *Defect Crystal Chemistry and Its Applications*, Blackie, Glasgow. Chapter 1. A very readable text.

Wagner, C. and Schottky, W. (1929) *Thermodynamik*, Springer, Berlin. The book by the masters—not often read.

Wagner, C. and Schottky, W. (1930) "Theory of regular mixed phases," *Z. Phys. Chem. B***11**, 163. An original paper.

WWW

MARVIN was originally authored by David H. Gay and Andrew L. Rohl. The program is "an advanced computer code for performing static and dynamic modeling of surfaces and interfaces." http://www.ri.ac.uk/DFRL/research_pages/resources/MARVIN_data/main.html

EXERCISES

11.1 What, if any, is the difference between a Schottky defect in Mg and a Schottky defect in MgO?

11.2 Would you expect Frenkel defects to be more likely in Al or Al_2O_3? What factors do we need to consider in answering this question?

11.3 How might you expect Si to diffuse in an alumina furnace tube at high temperatures? Would the mechanism be different if the alumina were in the form of a single crystal?

11.4 We form a Schottky defect in KCl. What is the difference in energy of the system if the two vacancies are $5.5a$ apart or $3.5a$ apart (a being the lattice parameter). Why do we choose 3.5 rather than 3, etc.? Discuss.

11.5 We dissolve (add) 100 ppm of Al_2O_3 to a cube of MgO (1 mm on the side). How many vacancies are formed and how much bigger is the cube?

11.6 Construct a table of defects and discuss the different types of point defects in Fe_2O_3. Compare this table with the table you produce for Fe_3O_4.

11.7 Write down a point defect reaction for ZnO in which you produce twice as many singly charged defects as doubly charged ones.

11.8 Write a balanced defect reaction equation using Kröger–Vink notation for substitution of Ca^{2+} in CaF_2 by Y^{3+}.

11.9 During the movement of the Na^+ ion in Figure 11.15, how does the CN of the Na^+ ion change as it moves through route 2.

11.10 Express using Kröger–Vink notation the formation of Frenkel defects in AgBr and explain why you think these are more numerous than Schottky defects. How might you increase the concentration of Schottky defects in AgBr?