

DEFECTS IN INORGANIC COMPOUNDS

D.M. SMYTH

1. INTRODUCTION

It is a great personal pleasure to contribute to this volume in honor of Per Kofstad on the occasion of his 60th birthday. Professor Kofstad has been one of the leading figures in the field of defect chemistry. I have followed his work with interest for over 30 years, and have benefited greatly from his insight and originality. It is to be hoped that we can look forward to many more years of contributions from Professor Kofstad and his students. In order to participate in the celebration of this notable anniversary, I would like to review some of the basic principles that underlie the field of defect chemistry. It is sometimes useful to step back from the daily battle, and to take a broader look at what we understand about a subject. Since Professor Kofstad is a fellow chemist, there will be an emphasis on the relationship between the basic chemical properties of the elements that are contained in a compound, and the nature of its defect chemistry. A little chemical knowledge can be very helpful in understanding or even predicting the types of disorder and the behavior that will be favored in a particular system.

This chapter will begin with a review of the sources of lattice and electronic defects in crystalline compounds, with an emphasis on the relationship between the known chemical properties of the constituent atomic components and the energetically most-favored types of disorder. This will be followed by a discussion of an appropriate definition and determination of the stoichiometric composition, no longer a trivial point as increasingly complex materials are considered. Following some considerations of nonstoichiometry, the use of the traditional mass-action approach for modeling defect chemistry will be discussed, and the chapter will conclude with some thoughts on the increasingly important phenomenon of defect ordering. Throughout the chapter, a purely ionic model will be used. While not accurate in terms of the actual nature of the chemical bonds in compounds, it is an adequate approximation for considerations of defect chemistry, since covalent contributions mostly cancel out because of the requirements of bulk charge neutrality.

2. DEFECTS IN CRYSTALLINE COMPOUNDS

2.1 General Considerations

Both lattice disorder and electronic defects are intrinsic parts of the

state of internal equilibrium in crystalline compounds. Intrinsic disorder may or may not be the major source of defects, depending on the energetics of their formation, and the state of purity and stoichiometry of the material. Aliovalent substitutional impurities also result in the presence of defects, and, in many cases, represent the major source. Deviation from the stoichiometric composition is the third principal source of defects, and can control important properties, especially electrical conductivity. In all of these cases, the defects are defined as deviations from a thermodynamic standard state, which is the ideally pure, stoichiometric composition with the perfect crystalline lattice, with all atomic constituents in their proper places and nowhere else, and with all electrons in the lowest available energy states. This is a fictional state that is never achieved in an absolute sense, but is approached in some cases. One of the convenient features of the field of thermodynamics is that it does not require that considerations be restricted to achievable states and processes. The standard state is obvious for simple systems such as NaCl, MgO, and Al_2O_3 , but its definition is nontrivial for complex systems, particularly when there are mixed valence states for some of the constituents, as in Fe_3O_4 and the recently discovered high temperature superconductor, $\text{YBa}_2\text{Cu}_3\text{O}_x$. These aspects will be discussed in a later section.

The creation of disorder in bulk crystalline solids is restricted by certain laws of conservation that must be reflected in the writing of properly balanced defect equations related to an appropriate standard state:

1. Conservation of mass: ions do not spontaneously appear or disappear in the system.
2. Conservation of charge: defect charges must balance to maintain bulk charge neutrality.
3. Conservation of structure: the lattice site ratios must conform to those of the ideal structure, and lattice sites can appear or disappear only in stoichiometric combinations.
4. Conservation of electronic states: the electronic states in the solid derive from the electronic states of the constituent atoms.

While these rules may seem obvious, their violation occurs with distressing frequency in the scientific literature, particularly for the case of the conservation of structure in complex compounds.

2.2 Intrinsic Ionic Defects

Intrinsic defects can be formed from the standard state by moving ions within a structure, as in the formation of cation or anion Frenkel disorder, or in the form of Schottky disorder by movement of stoichiometric combinations of cations and anions from normally occupied sites to new sets of lattice sites,

created in proper ratios on external or internal surfaces. Place exchange between cations and anions is not an important type of disorder in ionic materials because of the highly unfavorable electrostatic repulsions that result, but can be significant in highly covalent compounds such as GaAs and InSb.

Intrinsic defects form in order to minimize the free energy of the system as the configurational entropy, or entropy of solution, balances the enthalpic cost of the disorder. Since the configurational entropy as a measure of the randomness of the system is strictly statistical in nature, it is the same function of the defect concentration for a specific type of disorder in a given crystal structure. The relative amount of intrinsic disorder is then determined primarily by the enthalpy of disorder, which reflects the energetic cost of the disruption of chemical bonding relative to the standard state where it is always optimal. For convenience, consider the formation of Schottky disorder in a binary compound MX in which the oxidation states of the cations and anions are $+n$ and $-n$, respectively



The self-consistent defect notation of Kroger and Vink has been used, and nil refers to the defect-free standard state. The mass-action expression for this reaction is

$$\begin{aligned} [v_M^{n'}] [v_X^{n*}] &= K_s = e^{-\Delta G_s^0/kT} \\ &= e^{\Delta S_s^0/k} e^{-\Delta H_s^0/kT} \end{aligned} \quad [2]$$

where ΔG_s^0 , ΔH_s^0 , and ΔS_s^0 are the standard free energy, enthalpy, and entropy of the defect reaction, and are contained in the mass-action constant, K_s . The concentrations are in dimensionless site fractions, and the entropy term contains only vibrational contributions, since the configurational part of the entropy is contained in the form of the mass-action expression. When this is the only significant source of cation and anion vacancies, their concentrations must be equal

$$[v_M^{n'}] = [v_X^{n*}] = K_s^{1/2} = e^{\Delta S_s^0/2k} e^{-\Delta H_s^0/2kT} \quad [3]$$

This is an expression of charge neutrality. It is the exponential term containing the enthalpy per defect that primarily determines the concentration of defects. In a representative example of an MX with monovalent ions, NaCl, the enthalpy of formation for Schottky defects is 2.45 eV (236 kJ/mol), and the exponential enthalpy term per defect is 1.8×10^{-6} at 800°C, one degree below

the melting point (1). The standard entropy term, $\Delta S_g^0/k$, is 9.3, so that the exponential entropy term per defect contributes a factor of 105 to the defect concentrations, which are then 1.9×10^{-4} ; 0.019 % or 190 ppm of the lattice sites are vacant in equilibrium just below the melting point. This is a sufficient concentration to control transport properties, such as ionic conductivity and diffusion, in NaCl of readily achieved purity.

The factors that contribute to the enthalpy term include the question of adequate space for the defect, of importance for interstitial defects but not for vacancies; the amount of disruption of the chemical bonding, roughly indicated by the magnitude of the effective charge on the defect; and the coordination environment of the defect, again of more significance for interstitials than for vacancies. In the example of Schottky disorder in MX, the space available and the coordination environment are of little consequence, so that the magnitude of the defect charge is of prime importance. This is demonstrated by the properties of the isostructural MgO, where the ions are divalent and the vacancies carry an effective double charge. Because of the greater disruption in local charge when vacancies are created in this material, the enthalpy for Schottky disorder in MgO is much higher; a theoretical calculation indicates a value of 7 eV (670 kJ/mol) (2). Thus at 800°C the exponential enthalpy term per defect is 4×10^{-17} . If the entropy term is similar to that in NaCl, the concentrations of the cation and anion vacancies at 800°C will be 11 orders of magnitude less than those in NaCl, and concentrations at this level are of no practical significance. At the melting point of MgO, 2800°C, the exponential enthalpy term per defect is 1.9×10^{-6} , almost exactly the same as the value for NaCl at its melting point. This illustrates the commonly observed scaling of the defect formation energies with the melting points, since both processes involve disruption of the same chemical bonds.

The alternative cation and anion Frenkel disorders are less favorable than Schottky disorder in NaCl because of space and environment problems. It is difficult to fit the rather rigid Na^+ ion into the interstitial tetrahedral site in the rock-salt structure, since this site is only about half as large as the octahedral site that is normally occupied by the cations. The larger Cl^- causes an even more severe space problem, and in the interstitial site it would also find itself surrounded by four other nearest-neighbor anions, a highly unfavorable electrostatic environment. In the case of AgBr, which also has monovalent ions in the rock-salt structure, the unusually polarizable Ag^+ ion is able to squeeze into the tetrahedral site, and cation Frenkel disorder is favored over Schottky disorder. In fact the silver halides are remarkable in both the amount of ionic disorder that they allow, and the ease with which the Ag^+ can move through the lattice by an interstitialcy mechanism that results in extremely high ionic conductivity. Some years ago, this even led to the

development of highly stable, miniature batteries in which a thin film of silver halide functioned as the solid electrolyte (3). In contrast to the NaCl structure, in the fluorite structure the anions sit in tetrahedral sites in a face-centered array of cations. The larger octahedral sites can then serve as interstitial sites, and have a first coordination sphere of six cations, so they are quite hospitable to anions. Thus in a material such as CaF_2 , with monovalent anions and the fluorite structure, anion Frenkel disorder is the most favored type, and the ionic conductivity is primarily due to anionic defects.

While the effective charge on a defect can be used as a guide in estimating the probability of its existence in significant amounts, it must be applied with caution. Recent structural studies indicate that Li_2O -deficiency in LiNbO_3 is accommodated by the formation of Nb vacancies, $V_{\text{Nb}}^{5'}$, and Nb on Li-sites, $\text{Nb}_{\text{Li}}^{4\bullet}$ (4). Based on effective charges, these would seem to be extremely unfavorable defects. However, it has been proposed that these defects occur in pairs on adjacent cation sites, $(V_{\text{Nb}}^{5'} \text{Nb}_{\text{Li}}^{4\bullet})'$, and this results in a local cation sequence that corresponds to the closely related ilmenite structure with a Li vacancy, V_{Li}' (5). In both the LiNbO_3 and ilmenite structures, the cations fill 2/3 of the octahedral sites in an hcp oxygen lattice, and these sites form linear strings perpendicular to the basal planes. The filling sequence is AB-AB-AB-AB- in the LiNbO_3 structure and AB-BA-AB-BA- in the ilmenite structure, where - designates an unoccupied octahedral site. The complex defect differs from a Li vacancy in the LiNbO_3 structure only by movement of the Nb by one cation site



It has since been shown by theoretical calculations that the complex defect is indeed the favored configuration (6). The lesson is that the entire picture must be carefully examined before conclusions can be drawn about the most favorable types of defects. In this case the two highly charged point defects combine to give a singly charged complex defect, so one is really looking at the stability of a species with a single net charge.

In principle, all possible kinds of intrinsic disorder should occur simultaneously in a material in amounts determined by their thermodynamic parameters. In practice, one form is usually energetically favored to an extent that it dominates the defect structure. The minority defects are also suppressed by a common defect effect. Thus in NaCl, the high intrinsic cation vacancy concentration will enter the mass-action expression for the less-favored cation Frenkel disorder, and will suppress the cation interstitial concentration below the already low level it would have had in the absence of Schottky disorder. Even when they are not the major source of defects, the intrinsic defect reactions and their mass-action expressions serve as valid relationships

between the various defect concentrations. This can be of practical importance when a minority defect controls a process of interest. An example of this would be anion vacancy diffusion in AgBr; the effects of impurities and nonstoichiometry on the anion vacancy concentration will be linked to their effects on the major defects through the mass-action expression for Schottky disorder, even though this is not the preferred type of intrinsic disorder.

As a practical matter, intrinsic ionic disorder is the major source of lattice defects in only a limited number of materials. These include those cases where the defects are singly charged, and where the defect can fit into the structure, i.e. vacancies or small or highly polarizable interstitial cations, and in structures that have large, compatible interstitial sites as in the case of anion interstitials in the fluorite structure. Otherwise, impurities are the most likely source of the major lattice defects. However, even then the mass-action expressions for intrinsic disorder link together the effects of the impurities on a variety of defect species.

2.3 Intrinsic Electronic Disorder

All insulating and semiconducting compounds share a single, common form of electronic disorder, ionization across the band gap



The mass-action expression for this reaction is

$$np = N_c N_v e^{-E_g/kT} \quad [6]$$

where n and p are the electron concentration in the conduction band and hole concentration in the valence band, respectively, and N_c and N_v are the effective density of states at the band edges. E_g is the band gap between the valence and conduction bands, and may be equated with the standard free energy of the intrinsic ionization reaction. Thus n/N_c and p/N_v are the fractional populations of available states at the bottom of the conduction band, and at the top of the valence band, respectively. If this intrinsic disorder is the major source of electronic defects, then n and p must be equal

$$n = p = (N_c N_v)^{1/2} e^{-E_g/2kT} \quad [7]$$

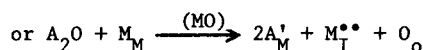
In a simple compound such as NaCl, the valence band is derived from atomic Cl 3p states and is exactly filled at 0 K, as are the atomic 3p states in the isolated chloride ion. The conduction band is derived from Na 3s states and is nominally empty. Thus the elevation of an electron from the valence band to the conduction band is related to taking an electron away from the chloride ion and giving it back to the sodium, to make elemental Cl and Na. This is highly unfavorable in the matrix of the ionic lattice, and this is reflected in the

wide band gap, 8-10 eV (800-1000 kJ/mol), and the insulating properties of such materials. Thus there is a relationship between the chemistry of the constituent elements, in the form of the existence of achievable adjacent oxidation states, and the magnitude of the band gap that determines the electrical nature of the resulting compounds. Compounds of the transition metals, which generally have multiple oxidation states, typically have modest band gaps of the order of 3-4 eV (300-400 kJ/mol).

2.4 Extrinsic Disorder

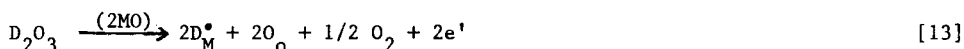
As just described, aliovalent impurities are commonly the major source of defects. In this case the defects are required to maintain charge neutrality in the presence of charged impurity centers. Unfortunately, in most oxides of interest, the metallic impurity content is commonly in the range of 10-1000 ppm atomic, and this is enough to dominate the defect chemistry under most conditions. It is appropriate to mention that Per Kofstad was one of the first to emphasize the dominant role of impurities in the defect chemistry of transition metal oxides (7,8). That idea is now generally accepted, but previously there was a misplaced emphasis on the role of intrinsic disorder in oxides of multiply charged cations.

There are three classical options for the charge compensation of aliovalent substitutional impurities. For acceptor impurities, which represent negatively charged impurity centers, these are interstitial cations, anion vacancies, or holes. For a generic oxide MO , the options for incorporation of the oxide of an acceptor impurity, A_2O include



In each reaction, the number of units of MO to be replaced by the designated amount of the impurity oxide is indicated in parentheses above the arrow. This is a convenient form of book-keeping to help maintain the proper site balance. The two choices in Eq. (8) depend on whether the impurity ion or the much more prevalent host cation is preferred in the interstitial site. The hole option requires an interaction with the atmosphere whereby the anion vacancy is filled by the introduction of neutral oxygen atoms that then acquire two electrons from the valence band in order to become normal oxide ions. This represents a true oxidation of the system, in that the starting materials were the host oxide and the impurity oxide in their normal, stable oxidation states, and then extra

oxygen was added to the system, raising the apparent average oxidation state of the cations. In the well-known case of NiO doped with Li_2O , where the favored compensation is by holes, the oxidizing power of the material will liberate free chlorine when it is dissolved in aqueous HCl. The possible defects for charge compensation of positively charged donor centers are cation vacancies, anion interstitials, and electrons. Thus for the solution of D_2O_3 in MO the options are



The electron option involves the loss of neutral oxygen from the lattice with the two electrons from each oxide ion being left behind, and this situation represents a real chemical reduction of the system.

There should be continuity between the preferred type of intrinsic defects and the defects that compensate charged impurity centers. Fig. 1 shows a log-log plot of the defect concentrations as a function of donor impurity for an MO in which Schottky defects are the preferred type of disorder. It is obvious from this that as the impurity concentration surpasses the level of the intrinsic disorder that the impurity will pick up the oppositely charged intrinsic defect and carry it upward to maintain charge neutrality. For a system of the type shown in Fig. 1 there is no basis for expecting minority defects, such as anion interstitials, to suddenly spring from the depths and become the major extrinsic defect.

The reactions in Eqs. (8) - (13) have been deliberately written with an arrow toward the right, rather than with double arrows that imply equilibrium, because the impurity content is seldom in equilibrium with anything. Impurities are usually incorporated by inclusion in the starting mixture of binary compounds, by dissolution in the melt, or by diffusion. Only rarely are the impurities introduced into the host compound by equilibration with a source of fixed activity. Thus the impurity content is treated here as a fixed concentration irreversibly incorporated into the lattice. However, there is a potential equilibrium link between the alternative modes of accommodation. Thus we can obtain an equation for the shift from compensation of an acceptor impurity by anion vacancies, Eq. (9), to holes, Eq. (10), by subtracting the two equations to obtain



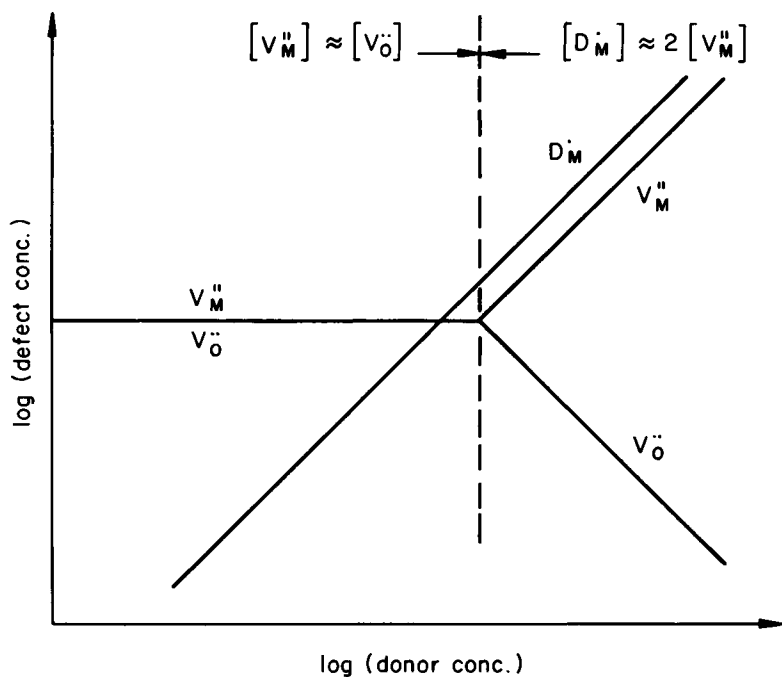


Fig. 1. Defect concentrations as a function of a donor impurity content, $[D_M^\bullet]$, for an oxide MO with Schottky disorder. Regions with intrinsic and extrinsic behavior are shown.

This leads to a mass-action expression at fixed acceptor content that emphasizes the dependence of the mode of compensation on the nonmetal activity

$$\frac{p^2}{[V_O^{\bullet\bullet}]} = K_p P(O_2)^{1/2} \quad [15]$$

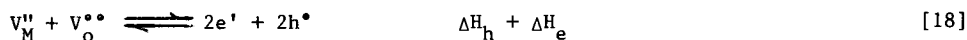
A similar linkage can be obtained for donor impurities by combination of Eqs. (12) and (13)



and its mass-action expression

$$\frac{n^2}{[V_M'']} = K_n P(O_2)^{-1/2} \quad [17]$$

If we assign ΔH_h as the enthalpy for the shift from anion vacancies to holes in Eq. (14), and ΔH_e as the enthalpy for the shift from cation vacancies to electrons in Eq. (16), and add the two equations, we obtain



This is also the difference between twice the intrinsic electron ionization reaction, Eq (5), and the formation of Schottky disorder, Eq (1), so that

$$\Delta H_h + \Delta H_e = 2E_g - \Delta H_s \quad [19]$$

This emphasizes the relationship between the enthalpies for intrinsic defect formation and the general tendency to compensate charged impurity centers by ionic or electronic defects. Unfortunately, the enthalpies for impurity compensation for acceptors and donors are linked as a pair, and their individual contributions cannot be separated, but the relationship does give some indication of the behavior to be expected. Eq. (19) states that the shift to electronic compensation involves the enthalpic cost of creating the electronic defects, but that this is reduced by the enthalpic advantage of eliminating the ionic defects. The strong role of the band gap in determining the behavior is apparent. For two isostructural oxides such as NiO and MgO, the enthalpies for Schottky disorder should be comparable, perhaps 7 eV (670 kJ/mol) (2), but the band gaps may differ by as much as 6-7 eV (580-670 kJ/mol). As a result, the sum of the two enthalpies for a shift in the compensation mode from ionic to electronic defects will be 12-14 eV larger in the case of MgO. Thus NiO can be nonstoichiometric by about 0.1% excess oxygen, while MgO exhibits no detectable nonstoichiometry at all. The mode of impurity compensation is thus directly linked to the chemistry of the cation, since the magnitude of the band gap is linked to the stability of adjacent oxidation states, as described earlier.

The relative stability of adjacent oxidation states can be used as a more specific chemical test for the probable mode of impurity compensation. In the case of TiO₂ doped with a donor such as Nb₂O₅, the possible compensating defects are interstitial anions, cation vacancies, or electrons. As described earlier in this section, compensation by electrons results from the loss of oxygen from the combination of the stoichiometric binary oxide components, and represents a chemical reduction of the system. This possibility will be enhanced if there is a chemically reducible species in the solid solution. Lower oxidation states of Ti are achievable, e.g. Ti₂O₃ and TiO, and electrons are therefore chemically compatible with the system. Donor impurities in TiO₂ are indeed compensated by electrons, at least at low concentrations, and the resulting materials are n-type semiconductors near room temperature. However, for the case of an acceptor impurity such as Al₂O₃, compensation by holes requires an addition of oxygen to the solid solution of stoichiometric binary constituents, and represents a chemical oxidation of the system. Oxidation of Ti⁺⁴ would require the removal of an electron from the stable rare gas electronic configuration, and this is energetically unfavorable. Likewise, oxidation of O²⁻ to O⁻ is also unfavorable. Accordingly, acceptor impurities in TiO₂ are compensated by ionic defects, resulting in insulating behavior. In a material in which the cations

have achievable higher oxidation states, e.g. NiO, the situation is reversed. Acceptor impurities are compensated by holes and give p-type semiconductors, as in NiO doped with Li_2O , while donor impurities, such as Cr_2O_3 , are primarily compensated by ionic defects and give insulating material.

3. THE STOICHIOMETRIC COMPOSITION

Any discussion of nonstoichiometry requires a reference composition that is considered to be stoichiometric. In other words, nonstoichiometric relative to what? This seems to be a trivial question for simple binary compounds such as NiO, TiO_2 , or Nb_2O_5 , where the stoichiometric composition corresponds to the simple, integral atomic ratios with the cation having a single, integral oxidation state. However, the question has come to the forefront again with the discovery of the complex and highly nonstoichiometric superconducting oxide $\text{YBa}_2\text{Cu}_3\text{O}_x$, where $6 < x < 7$. This has usually been described as $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$, and sometimes as $\text{YBa}_2\text{Cu}_3\text{O}_{6+y}$. These representations imply reference states with average oxidation states for the Cu of +2.333 and 1.667, respectively. It is tempting to choose $\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$ as the reference, since this would correspond to all of the Cu being divalent. One might also view the completely filled perovskite-related structure, $\text{YBa}_2\text{Cu}_3\text{O}_9$, as the proper reference. The correct choice is important since it will determine which compositions should be oxygen-deficient, n-type conductors and which should be oxygen-excess, p-type conductors. The stoichiometric composition separates these regions of quite different behavior. Before dealing with such a complex material, it is instructive to review the defect chemistry of a simple compound, where the situation is intuitively clear.

Consider a hypothetical binary compound of a monovalent cation and a monovalent anion, MX, in which Schottky disorder is the preferred type of intrinsic ionic disorder. In the ideally pure, stoichiometric state, the only defects present will result from intrinsic ionic and electronic disorder. The former will correspond to Eq. (1) with $n = 1$, and the latter to Eq. (5). At the stoichiometric composition, the concentrations of cation and anion vacancies will be equal, $[V_M^\bullet] = [V_X^\bullet]$, and the concentrations of free electrons and holes will also be equal, $n = p$. Under these circumstances, the conduction band is nominally empty and the valence band is nominally filled. This is a useful definition for the stoichiometric composition.

Any decrease in the equilibrium nonmetal activity $P(X_2)$ will cause an increase in V_X^\bullet and n



$$[V_X^\bullet] n = K_n P(X_2)^{-1/2} \quad [21]$$

while an increase in $P(X_2)$ above that in equilibrium with the stoichiometric composition will result in an increase in V_M' and p



$$[V_M'] p = K_p P(X_2)^{1/2} \quad [23]$$

If in MX , K_s is considerably larger than K_i , charge neutrality will be approximately satisfied by $[V_M'] = [V_X^\bullet] = K_s^{1/2}$ for small deviations from stoichiometry. Combination of this condition with the mass-action expressions for the reduction and oxidation reactions, Eqs (21) and (23), indicates that n-type conductivity will increase with increasing deficiency of X , while p-type conductivity will increase with increasing excess of X . The equilibrium conductivity will then have a minimum value at the stoichiometric composition, neglecting any inequality between the electron and hole mobilities. This is a convenient way to determine the equilibrium conditions for the stoichiometric composition. If the composition of the material is known as a function of temperature and $P(X_2)$, then the stoichiometric composition can be determined from the conditions at which the conductivity has its minimum values.

If K_i is larger than K_s in our hypothetical compound MX , the situation is experimentally more complex. Charge neutrality in the near-stoichiometric region will be approximated by $n = p = K_i^{1/2}$. While in principle there will still be a minimum value in the conductivity as a function of $P(X_2)$, it will generally be much too shallow to be located accurately. The ionic defects will then be a strong function of $P(X_2)$, and a diffusion experiment may be helpful in determining the conditions under which the material is stoichiometric. For example, for cation self-diffusion, the motion may be by a vacancy mechanism for nonmetal-excess compositions, but by an interstitial mechanism for nonmetal-deficient compositions. If the predominant ionic disorder is of the cation Frenkel type, then the minimum in the cation self-diffusion constant as a function of $P(X_2)$ will correspond to the stoichiometric composition, if mobility differences for the two mechanisms can be ignored. In fact, the mobility difference may be substantial, and the diffusion minima may be correspondingly displaced from stoichiometry. The situation would be even more complex if Schottky disorder is the major source of ionic defects. As shown in Fig. 2, for this case the position at which $[M_i^\bullet] = [V_M']$ will be displaced from the stoichiometric composition by an amount that will depend on the relative mass-action constants for the two types of ionic disorder, even without consideration of the relative mobilities.

In a compound with a cation in mixed oxidation states, such as Fe_3O_4 , electronic disorder is expected to prevail, and this represents a case where

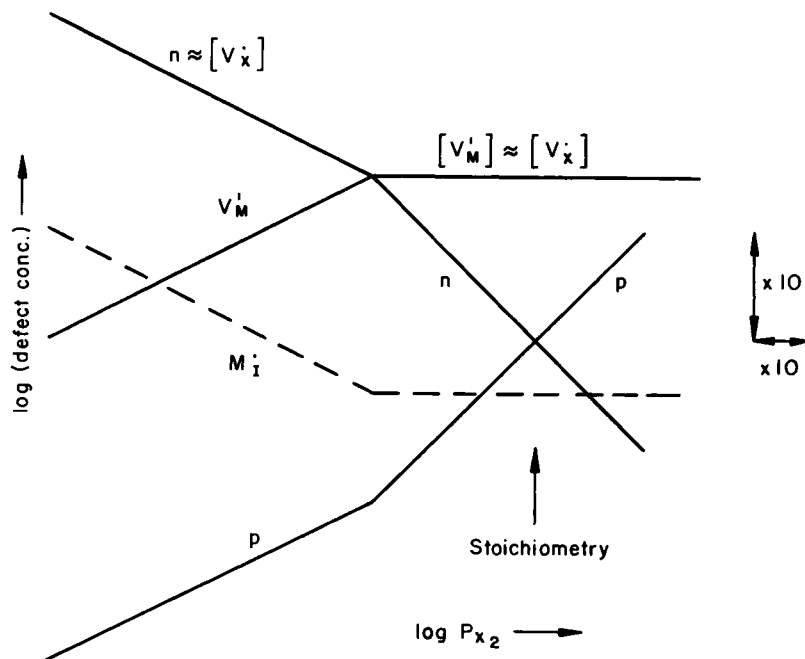


Fig. 2. Defect concentrations as a function of nonmetal activity for MX with monovalent ions and $K_S = 10^4 K_I = 10^3 K_F$. $P(X_2)$ for $[M_I^\cdot] = [V_M^\cdot]$ is displaced by 10^{10} from $P(X_2)$ at the stoichiometric composition.

diffusion studies should be more informative than conductivity studies. In the elegant measurements by Dieckmann of Fe self-diffusion as a function of $P(O_2)$ in Fe_3O_4 , reproduced in Fig. 3, clear minima are seen (9). The results were interpreted as resulting from predominant cation Frenkel disorder in the inverse spinel structure, with excess cation vacancies being formed by oxidation, and excess cation interstitials being formed by reduction. The resulting defect model is in good agreement with the observed dependences of the diffusion constants on $P(O_2)$. In this case, the stoichiometric composition will be displaced from the diffusion minima only because of the different mobilities for the two mechanisms. By comparison of the diffusion data with compositional data obtained by coulometric titration (10), it was determined that at $1200^\circ C$, the stoichiometric composition occurs at 6.8×10^{-8} atm, a value that is 15 times lower than the $P(O_2)$ at the minimum in the diffusivity.

In Fe_3O_4 , the stoichiometric composition corresponds to a mixed oxidation state for the Fe, with one Fe^{+2} and two Fe^{+3} , or an average oxidation state of 2.67. At stoichiometry, $[Fe_I^{\bullet\bullet}] = [V_{Fe}^{\bullet\bullet}]$ and $n = p$, so that all of the electronic bands must be either nominally filled or empty. This implies that the energy

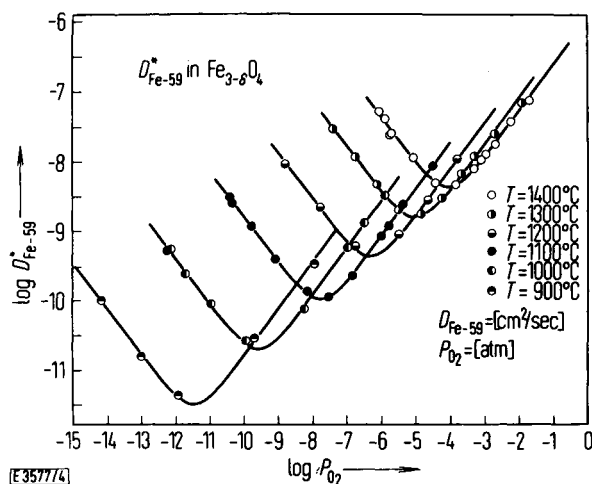


Fig. 3. Fe tracer diffusion for Fe_3O_4 as a function of oxygen partial pressure according to Dieckmann and Schmalzried (9) showing evidence of an interstitial mechanism at low $P(\text{O}_2)$ and a vacancy mechanism at high $P(\text{O}_2)$. (Reprinted from (9) with permission from the authors and from Verlag Chemie GMBH.)

bands associated with the two oxidation states of the Fe must be distinct, since if they were combined into a single Fe band with a noninteger number of electrons, one would expect partial band-filling and metallic behavior. This is in accord with the usual view of Fe_3O_4 as an inverse spinel with discrete integer oxidation states assigned to the Fe. Stoichiometric Fe_3O_4 is a combination of the two stoichiometric binaries, $\text{FeO} + \text{Fe}_2\text{O}_3$. This is more obvious when two different cation species are present as in the case of the normal spinel Mg_2AlO_4 .

We are now in a position to better understand the nonstoichiometry of $\text{YBa}_2\text{Cu}_3\text{O}_x$. The equilibrium conductivity of a sintered, polycrystalline sample is shown in Fig. 4. The range of $P(\text{O}_2)$ is severely limited because of the easy reducibility of the Cu content. The behavior is primarily that of an oxygen-excess, p-type material, but a conductivity minimum is detected at the highest temperatures and lowest $P(\text{O}_2)$, just before the material irreversibly decomposes. The shape of the minimum is consistent with the situation where ionic disorder predominates, as might be expected in such an oxygen-deficient version of the perovskite structure. Comparison of the conditions at the minima with published equilibrium compositional data (11) indicates that the composition at the minima is $\text{YBa}_2\text{Cu}_3\text{O}_6$. This is then the proper stoichiometric reference composition, and material with higher oxygen contents is correctly referred to as $\text{YBa}_2\text{Cu}_3\text{O}_{6+y}$. As in the case of Fe_3O_4 , the stoichiometric composition involves a cation in a mixed oxidation state, in this case two Cu^{++} and one Cu^+ . Since the stoichiometric state is determined by the condition that bands be nominally either filled or empty, it is also implied for this material that the two Cu states are represented by two different sets of band states, and this is supported by a variety of theoretical calculations of the band structure. As in

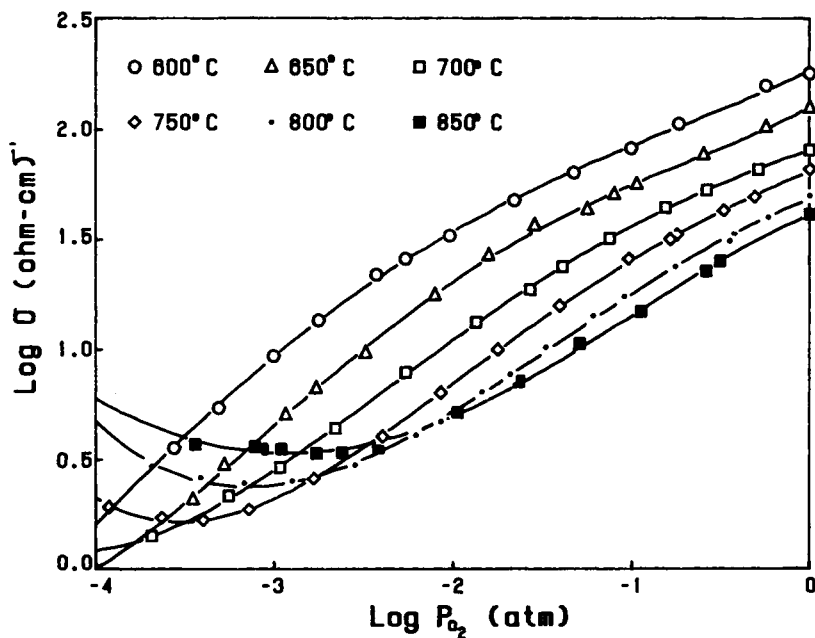


Fig. 4. The equilibrium conductivity of $\text{YBa}_2\text{Cu}_3\text{O}_x$ as a function of the oxygen partial pressure.

the case of Fe_3O_4 , the stoichiometric composition corresponds to a combination of stoichiometric binaries, $1/2(\text{Y}_2\text{O}_3) + 2\text{BaO} + 1/2(\text{Cu}_2\text{O}) + 2\text{CuO}$. It should be noted that this compound behaves as a semiconductor at high temperatures, over the entire compositional range $0 < y < 1$. Only at lower temperatures do the more highly oxidized compositions show metallic behavior, and this must reflect some change in the band structure.

When impurities play a major role in the defect chemistry of a material, the definition of stoichiometry is the same as that for the ternary and quaternary compounds described above. It corresponds to a combination of stoichiometric binary compounds, including the impurity compound, and to the presence of nominally filled and empty bands. Thus the stoichiometric composition for alumina-doped TiO_2 is $(1-x)\text{TiO}_2 + x/2(\text{Al}_2\text{O}_3)$, and for Nb-doped BaTiO_3 it is $\text{BaO} + (1-x)\text{TiO}_2 + x/2(\text{Nb}_2\text{O}_5)$. The equilibrium conductivities as a function of $P(\text{O}_2)$ will have their minimum values at these compositions. Oxidation of alumina-doped TiO_2 will correspond to the filling of oxygen vacancies in the rutile structure, or perhaps more accurately, the removal of interstitial Ti from the rutile structure. Thus the reference structure is the ideal rutile structure that corresponds to a composition that is different from the stoichiometric reference composition. It is important to make a distinction

between these thermodynamic and structural reference states. In $\text{YBa}_2\text{Cu}_3\text{O}_x$, for which the stoichiometric composition was found to correspond to $x = 6$ where the structure is tetragonal, it has been shown that the oxidation reaction for $x > 6.3$ corresponds to the filling of oxygen vacancies in the orthorhombic reference structure of $\text{YBa}_2\text{Cu}_3\text{O}_7$ (12).

4. NONSTOICHIOMETRY

There is an obvious relationship between nonstoichiometry and extrinsic disorder, since the former can be viewed as self-doping with a constituent cation in a different oxidation state. A thermodynamic thought experiment could go as follows: starting with a sample of MO , for which Schottky defects are the preferred form of intrinsic ionic disorder, break off a small piece and reduce it to M_2O



Then redissolve the resulting M_2O into the remaining MO



M'_{M} is equivalent to M^{+2} plus an electron



Addition of these three equations leads to the familiar reduction reaction



and its corresponding mass-action expression

$$[\text{V}_{\text{O}}^{\bullet\bullet}]^2 n = K_n P(\text{O}_2)^{-1/2} \quad [28]$$

A similar exercise can be done for self-doping with the cation in a higher oxidation state to lead to an equilibrium oxidation reaction and its mass-action expression



$$[\text{V}_{\text{M}}''] p^2 = K_p P(\text{O}_2)^{1/2} \quad [30]$$

If we assign ΔH_n and ΔH_p as the enthalpies of the reduction and oxidation reactions, and add the reactions, we obtain



This is also the sum of the formation reaction for Schottky disorder and for twice the intrinsic electronic ionization reaction, so that

$$\Delta H_n + \Delta H_p = 2E_g^\circ + \Delta H_s \quad [32]$$

Once again there is a relationship between the combined tendencies for nonstoichiometry and the enthalpies of intrinsic disorder. In this case it is the sum of the intrinsic enthalpies, because reduction and oxidation result in the formation of both ionic and electronic defects, and there is an enthalpic cost for both of them. The dominant role played in oxides by the band gap is again apparent, and this leads directly to a correlation of the extent of nonstoichiometry with the stability of adjacent oxidation states of the cationic component. The chemical properties of the cations in the system will influence the balance between the tendency for oxygen-excess or oxygen-deficient nonstoichiometry. Those oxides that contain a cation with achievable higher oxidation states, such as NiO, CoO, MnO, FeO, and Cr_2O_3 , will display primarily oxygen-excess, p-type behavior, while those that contain a cation with achievable lower oxidation states, such as TiO_2 , Nb_2O_5 , Ta_2O_5 , and WO_3 , will behave as oxygen-deficient, n-type materials over a wide range of equilibration conditions.

The importance of maintaining continuity in the types of defects formed also holds for nonstoichiometry. One expects to see the same defects, regardless of whether they are intrinsic, extrinsic, or due to nonstoichiometry. In the familiar Kroger-Vink diagrams that depict the concentrations of the various significant defects as a function of nonmetal activity, continuity is shown by the fact that there is always a common defect in adjacent regions defined by different approximations to the condition of charge neutrality.

Measurements of the equilibrium electrical conductivity as a function of nonmetal activity generally follow the changing concentrations of electrons or holes generated by reduction or oxidation reactions such as Eqs (27) and (29). (In some cases there may be an additional ionic contribution.) At the temperatures that are necessary for equilibration, these measurements usually show quite comparable magnitudes for electron and hole conduction, implying that the carriers are free to contribute to conduction, and are not significantly trapped by the oppositely charged ionic defects. However, this situation may change drastically when the materials are cooled to the vicinity of room temperature. Some of the defect levels may be deep enough to serve as effective traps for one or both types of charge carrier at the lower temperatures. Once again, a chemical test can be applied to determine which defects might serve as trap sites. The conduction band of TiO_2 is derived from Ti 3d atomic states, and the placement of an electron in it is analogous to reducing Ti^{+4} toward Ti^{+3} . Since such lower oxidation states are possible for Ti, these electronic states are accessible to electrons. If the electrons are in the conduction band, the related positively-charged donor sites, i.e. oxygen vacancies or

cation interstitials, must be sufficiently close to the conduction band not to serve as traps. Thus in reduced oxides that contain a reducible cation, the donor levels are very shallow, and the electrons generally remain free to contribute to conduction down to room temperature and below. On the other hand, the valence band in TiO_2 is derived from O 2p atomic states, and the presence of a hole would be analogous to the oxidation of O^{--} to O^- , which is not generally possible. Thus the valence band states are not very favorable to holes, which means that they are trapped by the acceptor centers, such as cation vacancies or acceptor impurities, as the materials are cooled toward room temperature, with resulting insulating behavior. In other words, acceptor levels are relatively deep in such materials. For systems that contain oxidizable cations, the situation is reversed. Donor centers are deep, and acceptor levels are shallow, and such materials are either n-type insulators or p-type semiconductors. NiO is a good example of this type of behavior. In this case, the valence band is derived from the partially filled Ni 3d atomic states and the presence of holes is analogous to the oxidation of Ni^{+2} toward Ni^{+3} , which is chemically feasible. Since the chemical nature of the atomic species in the compound influences the tendency of the material to oxidize or reduce, as well as the relative levels of acceptor and donor levels, the chemical test is doubly powerful. When the oxide tends to reduce to an oxygen-deficient, n-type material, the resulting donor levels will also be shallow so that the electrons created by reduction will remain free to contribute to conduction at ambient temperatures. The resulting asymmetry in electrical properties is of considerable practical importance in a variety of electronic and optical devices (13).

5. THE LAW OF MASS ACTION

The traditional approach for the analysis of the defect chemistry of a compound is to write a properly balanced equilibrium equation that expresses the interactions of the various species, and then to write a mass-action expression involving the concentrations and partial pressures of the species involved in that reaction. The analogy to the similar approach so commonly used in aqueous chemistry is obvious. In this case the matrix is the otherwise perfect crystal, the usual standard state, and the defects, both ionic and electronic, are the reactants and products. The usefulness of this approach depends on the validity of the approximations of dilute solution thermodynamics, mainly that concentrations are not significantly different from activities. Whereas this approximation begins to fail for even quite dilute aqueous solutions, e.g. above 0.001 molar, it seems to hold very well for lattice and electronic defects in solids even for concentrations in the percent range. Presumably, this fortunate situation is due to the effective screening of the defect charges by the crystal lattice.

The mass-action relationship can be derived in several ways. One simple approach involves the contribution of the chemical potential, μ_i , of each species i in the equilibrium reaction. The chemical potential, or partial molar Gibbs free energy is defined by the expression

$$\mu_i = \left(\frac{\partial G}{\partial a_i} \right)_{P,T,a_1,a_2,\dots} = \mu_i^0 + RT \ln a_i \quad [33]$$

where R is the gas constant, a_i is the activity of species i , and μ_i^0 is the standard chemical potential at unit activity. μ_i is the change in the Gibbs free energy of the system that results from an incremental change in the activity of species i at constant pressure, temperature, and activities of all other species. Consider the generic equilibrium reaction



where the symbols represent a moles of species A, b moles of species B, etc., and the species may be ionic or electronic defects, or the partial pressure of a gaseous species in equilibrium with the system. The contributions to the free energy of the system at equilibrium are

$$\begin{aligned} \Delta G = & c (\mu_C^0 + RT \ln a_C) + d (\mu_D^0 + RT \ln a_D) \\ & - a (\mu_A^0 + RT \ln a_A) - b (\mu_B^0 + RT \ln a_B) = 0 \end{aligned} \quad [35]$$

The total free energy change is set equal to zero since the system is at equilibrium. This can be rearranged to give

$$a\mu_A^0 + b\mu_B^0 - c\mu_C^0 - d\mu_D^0 = -\Delta G^0 = RT \ln \frac{a_C^c a_D^d}{a_A^a a_B^b} = RT \ln K \quad [36]$$

where ΔG^0 is the standard free energy change for the reaction and K is the mass-action constant. The dilute solution approximation then involves the substitution of concentrations for the activities and partial pressures for the fugacities. It should be noted that according to Eq. (36), K is a dimensionless number, so that unitless, fractional concentrations should be used, unless the reactants and products involve the same total number of species, in which case the concentration units cancel out, or unless another factor is included to take account of the specific concentration units used. Site fractions are commonly used for lattice defects, but electronic defects may be expressed in volume concentration, carriers per cm^3 , or relative to a lattice species such as electrons per cation. The latter may be appropriate for a hopping or polaron conduction mechanism, where the carriers are localized between discrete jumps on an ionic species, but for band conduction the concentrations should be related

to the total number of available states that they can occupy, and this is the effective density of states at the band edges.

Another commonly used approximation is the neglect of the change in the concentration of some species. As an example, consider an oxide MO that has been doped with an acceptor cations A^+ , that are compensated by oxygen vacancies. The material can be oxidized by the partial filling of the compensating vacancies according to the reaction described in Eq. (14), for which the complete mass-action expression is

$$\frac{[O_o] p^2}{[V_o^{\bullet\bullet}]} = K_{ox} P(O_2)^{1/2} \quad [37]$$

If the degree of oxidation is modest, the change in the concentration of occupied oxygen sites can be neglected, and if only an insignificant fraction of the compensating vacancies become occupied, their concentration can also be taken as constant and as half of the net acceptor content. Eq. (37) then reduces to

$$p \approx \frac{[A']}{2} K_{ox}^{1/2} P(O_2)^{1/4} \quad [38]$$

It must always be recognized that this is an approximation, and that its validity must be examined when the defect concentrations become large. The observation of a p-type conductivity that varies as $P(O_2)^{1/4}$, or the analogous situation of an n-type conductivity that varies as $P(O_2)^{-1/4}$, is quite common, and it should be noted that this does not depend on the validity of the dilute solution assumption for the lattice species and lattice defects, since in the approximations used to obtain Eq. (38) from Eq. (37), the activities of both O_o and $V_o^{\bullet\bullet}$ are taken to be constant. Only the hole concentration is required to behave in an ideal manner, and this also implies that a negligible fraction of the available hole states have been filled, since any change in the concentration of unoccupied states has been neglected. It is unfortunate that the traditional pedagogical approach to nonstoichiometry emphasized, for example, the formation of neutral oxygen vacancies, that is vacancies with two associated electrons, and then considered the successive ionization of the electrons. This leads to a dependence of the carrier concentration on the negative fourth root of the $P(O_2)$ for the case of only single ionization of the vacancies if the reduction reaction is the only significant source of defects. As a result, the observation of such a pressure dependence was frequently cited as evidence for partial ionization of the nonstoichiometric defects, when in fact the dependence is due to the type of situation embodied in Eqns. (14) and (38), i.e. there is an extrinsic source of vacancies, such as acceptor impurities or cation nonstoichiometry in complex oxides, that fixes the vacancy

concentration. At least in the case of the reduction of transition metal oxides that contain a reducible cation, the author is not aware of any case of less than full ionization of the ionic defects under equilibrium conditions.

The probability that an electronic state in a solid will be occupied is given rigorously by the Fermi-Dirac expression

$$f(E) = \frac{1}{e^{(E-E_f)/kT} + 1} \quad [39]$$

where the probability is related to the distance of the state at energy E from the Fermi energy E_f . E_f is the energy at which the probability of being filled is $1/2$. Since $f(E) = n/N$, the ratio of occupied states to the total number of states, Eq. (40) can be rearranged to give

$$\frac{n}{N-n} = e^{-(E-E_f)/kT} \quad [40]$$

For $N \gg n$, this reduces to the familiar Boltzmann relationship which is an approximation to the Fermi-Dirac expression for the case that E is sufficiently far removed from E_f so that the exponential term in the denominator of the latter is large compared to unity.

The effective density of states at the edge of the conduction band, as derived from the model of a free electron Fermi gas, is

$$N_c = 2(2\pi m_e kT/h^2)^{3/2} \quad [41]$$

where m_e is the effective mass of the electron. There is a comparable expression for the valence band. Using the rest mass of the electron and a temperature of 1000°C , this amounts to $2.2 \times 10^{20} \text{ cm}^{-3}$, and it would appear that one should begin to worry about significant filling of these states at carrier concentration levels above about 10^{18} cm^{-3} , or even less if the effective mass of the carrier is significantly less than the rest mass. The band gap, the characteristic energy for the intrinsic ionization of electrons, is usually found to be linearly dependent on the temperature, and can be expressed as $E_g = E_g^0 - \alpha T$. This emphasizes the fact that the band gap of the physicist is a free energy, and E_g^0 the so-called band gap at absolute zero, corresponds to the standard enthalpy of the intrinsic ionization reaction, while α , the linear temperature coefficient, corresponds to the entropy. For an intrinsic semiconductor where the effective masses of the electrons and holes are equal, $N_c = N_v$, and E_f is located at the midpoint of the gap so that $E_c - E_f = E_g/2$. The electron concentration is then properly written as

$$n = N_c e^{\alpha/k} e^{-E_g^0/2kT} \quad [42]$$

and the preexponential term may thus be larger than the density of states by as much as a factor of 100 because of the inclusion of this entropy term. It is ironic that the characteristic enthalpy that may be determined from data taken at several hundred degrees is known as the band gap at absolute zero.

From the analysis given above, it is apparent that when an equilibrium system involving electrons or holes at concentration levels of 10^{19} or above can be correlated with a mass-action expression, it is not likely that the conduction is by a simple band mechanism. As the carrier concentration rises, the behavior should first deviate from the uncorrected mass-action expression as a significant fraction of the available states are filled, and then, if the carrier concentration exceeds the effective density of states, the behavior should become metallic. In this case, the conductivity would cease to increase with changing nonmetal activity, because only the uppermost carriers in a partially filled band will contribute to the conductivity, and their concentration will no longer be affected by the generation of additional electronic species.

When the mass-action approach does remain valid for very high carrier concentrations, it is likely that the electrons or holes are localized to some degree and that some form of polaron mechanism is operative. The effective density of states then becomes the concentration of the ion with which the carrier is associated, and that is typically about 10^{22} cm^{-3} , or about two to three orders of magnitude greater than the density of states at the band edge. In many cases, the behavior seems to be in some kind of nebulous intermediate state. The carrier concentrations exceed the expected density of states and the mobilities lie in the range $0.01 - 0.1 \text{ cm}^2/\text{v sec}$, and correspond to mean free paths of less than a lattice spacing, so that a free electron gas picture is not valid. On the other hand, the temperature dependence of the mobility may show no evidence of the thermal activation expected for a hopping mechanism, and may even decrease with increasing temperature. BaTiO_3 is a good example of such behavior (14). The details of the conduction process in such materials does not seem to be thoroughly understood.

6. DEFECT ASSOCIATION

6.1 Specific Equilibrium Interactions

The use of the mass-action approach can be extended to higher concentration levels by taking into account specific types of interactions between the defects. These include partial ionization of defects, as seems to be important in the case of the oxygen-excess, p-type conduction in NiO , for example (15), or to the formation of defect complexes due to the electrostatic attraction of oppositely charged ionic defects. The latter type of interaction has been found

to be of importance in many systems such as the alkali halides, NiO (16), and calcia-doped zirconia (17).

6.2 Extrinsic Defect Complexes

The alkali halides are a convenient model group of compounds in which defect association can be demonstrated. In the case of extrinsic disorder, e.g. the incorporation of substitutional divalent cations that are compensated by cation vacancies, the formation of impurity cation-cation vacancy pairs is observed at sufficiently low temperatures.

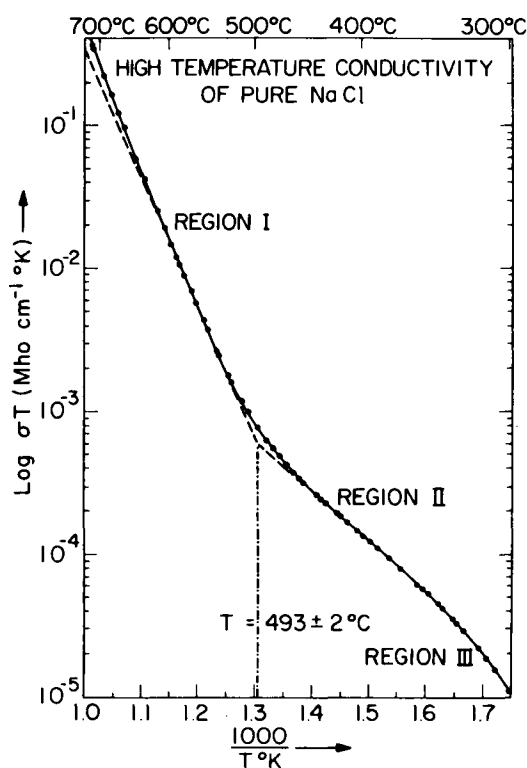


Fig. 5. An Arrhenius plot of the ionic conductivity of "pure" NaCl according to Kirk and Pratt (18), showing regions of intrinsic (I), extrinsic (II), and association (III). (Reprinted from (18) with permission from the authors and the British Ceramics Society.)

The complex is indicated by enclosing the associated species in parentheses. In this case, the complex is a neutral species and will not contribute to ionic conductivity. Its presence can be seen as a fall-off of the conductivity with decreasing temperature below the extrinsic region, as shown in Fig. 5 (18). The complexes can contribute to ionic diffusion, however, and are an essential part of the vacancy diffusion mechanism for the divalent cation which must be adjacent to a vacancy, and hence in the form of the complex, in order to move. They can also be detected by other techniques such as thermally stimulated currents, ac losses, and internal friction. Detailed information on their configuration can be obtained from electron paramagnetic resonance measurements when the divalent cation has an unpaired electron (19).

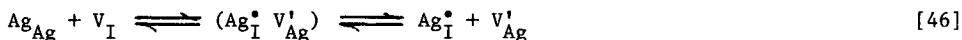
The complexes are bound by electrostatic attraction, and a simplistic picture of the binding energy treats it as a coulombic energy modified by the polarizability of the medium

$$\Delta H_c \sim \frac{z_1 z_2 e^2}{k_s r} \quad [45]$$

where ΔH_c is the enthalpy of association, z_1 and z_2 are the number of electronic charges e on each defect, with their signs, k_s is the static dielectric constant of the material, and r is the distance between the defects in the complex. ΔH_c is negative, since the complex dissociates with increasing temperature. This expression gives a value of about -0.5 eV (-50 kJ/mol) for a typical alkali halide, and this is in surprisingly good agreement with the experimentally determined values.

6.3 Intrinsic Defect Association

Association can also take place between the defects involved in intrinsic disorder. In the case of cation Frenkel disorder in AgBr, the first step toward forming a defect pair corresponds to an associated pair, which may then undergo dissociation into isolated, randomly distributed defects



The concentration of the associates is determined by the mass-action expression

$$[(Ag_I^\bullet V'_{Ag})] = [Ag_{Ag}] [V_I] K_p \quad [47]$$

where the mass-action constant K_p contains an enthalpy for pair formation, ΔH_p . It should be noted that the concentration of bound intrinsic pairs is only a function of temperature. Since the pair is electrically neutral, it is not affected by the presence of aliovalent impurities. The dissociation of the pair is characterized by its own mass-action expression

$$\frac{[Ag_I^\bullet] [V'_{Ag}]}{[(Ag_I^\bullet V'_{Ag})]} = K_D \quad [48]$$

where the mass-action constant K_D contains the enthalpy of dissociation, which is the negative of the enthalpy of association. The sum of the enthalpies of pair formation and of pair dissociation must equal ΔH_s , the enthalpy for formation of isolated, unassociated Frenkel defects

$$\Delta H_p + \Delta H_D = \Delta H_s \quad [49]$$

If ΔH_p is less than $1/2 \Delta H_s$, there will be more interstitials and vacancies in associated pairs than as isolated defects. Also in this case, the associated pairs can contribute to ionic diffusion, but not to ionic conductivity.

However, because of their dipolar nature, they can be detected by electrical relaxation measurements.

An analogous treatment can be given for the association of vacancy pairs in the case of Schottky disorder in the alkali halides. For NaCl the stepwise defect formation reaction can be written as



The simplified expression for the enthalpy of association given as Eq. (45) does not work as well for the association of intrinsic defects. Because of their greater proximity, the static dielectric constant overestimates the effect of the polarizability of the lattice, and the enthalpies are somewhat higher than predicted. In the alkali halides, the enthalpy of association of intrinsic defects should be $\sqrt{2}$ times higher than that for the extrinsic defect pairs, based on the difference in separation distances, but are actually about twice as large.

6.4 Structural Elimination of Defects

In some systems, compositional changes are accommodated by modification of the structure, rather than by the formation of lattice defects. The formation of the well-known Magneli phases in reduced or Cr-doped TiO_2 , for example, can be viewed as the organization of oxygen vacancies in certain equally spaced planes, followed by their elimination by a process of crystallographic shear (20,21). The resulting structures have two dimensional slabs of the rutile structure separated by equally-spaced metal-rich layers. The Ti sites in adjacent slabs are interstitial relative to one another. The perovskite structure, with the nominal composition ABO_3 , has a particularly rich variety of superlattice variations, and can accommodate nonstoichiometry by the incorporation of oxygen vacancies, extra oxygen, or extra amounts of the binary oxide of the larger cation in layers that separate two-dimensional slabs of the perovskite structure of varying incremental thicknesses (22,23). These lead to various series of homologous compositions having generic formulas such as $\text{A}_n\text{B}_n\text{O}_{3n-1}$, $\text{A}_n\text{B}_n\text{O}_{3n+2}$, and $\text{AO}-n\text{ABO}_3$, where n is an integer. It appears that the perovskite pattern is so favorable that these systems retain as much of it as possible by isolating the nonstoichiometric content into discrete layers that separate equally-spaced layers of the perovskite structure.

The question of whether a system will respond to a compositional change by forming a new structure, or by the formation of lattice defects, involves the usual balance between the configurational entropy, which favors disorder, and the lattice enthalpy, which is a function of the chemical bonding and promotes structural order. Lattice energy calculations by Catlow have shown that the superlattice structures are favored by materials that can relax extensively around the shear planes or interposed layers (24). The dielectric constant is a

good indication of the polarizability of the lattice, so it is to be expected that structural accommodation will be favored in those oxides that have unusually high dielectric constants. These include the transition metal oxides where the cations are in their group oxidation state, e.g. TiO_2 , Nb_2O_5 , and WO_3 , and these are in fact among those oxides for which superlattice ordering is frequently observed. On the other hand, it will be recalled that lattice polarizability reduces the electrostatic attraction of oppositely charged defects, and reduces the tendency for the formation of defect complexes. The polarizability of the structure plays quite different roles in these two different ways of organizing defects in a crystalline lattice.

7. SUMMARY

Studies of defect chemistry often emphasize the "defect" part over the "chemistry", and this is unfortunate, because this is really a sub-branch of the chemical behavior of materials. As such, the defect chemistry of a compound is closely coupled to the chemical properties of the constituent elements. Some knowledge of the available oxidation states for the metallic elements can be very useful in interpreting and predicting the behavior of their compounds. This information is most readily organized in the form of the periodic arrangement of the elements, and with the various trends in properties that result from horizontal, vertical, and diagonal movements within the periodic chart. A major goal of this chapter has been to demonstrate how the chemical properties can be related to such phenomena as the mode of compensation of aliovalent impurities, the tendency for a compound to oxidize or reduce, the magnitude of the band gap, and the relative depths of donor and acceptor levels.

Correlation of the chemical properties with the band gap is particularly useful, because this parameter sets the scale of the electronic aspects of nonstoichiometry, and this in turn determines the electrical behavior at the ambient temperatures where electronic and optical devices are used. The band gap determines the concentration of electronic charge carriers at equilibrium in the stoichiometric material. From this reference composition, the electron concentration rises with reduction as some defined function of the nonmetal activity, or the hole concentration rises with oxidation. The functional dependences range typically from the square root to the eighth root of the nonmetal activity, so even large changes in this experimental variable give carrier concentrations that are still strongly influenced by the starting levels determined by the band gap. This is the basis for the enormous difference in behavior between compounds of the main group metals, such as the alkali metals, the alkaline earths, Al, Sc, Y, and La, and those of the transition metals, such as Ti, Zr, Nb, Ta, W, Mn, Fe, Co, Ni, and Cu.

The lesson is: the electronic properties of the solid compounds are closely linked to the chemistry of their constituent atoms.

ACKNOWLEDGEMENT

The long-term support by the Materials Research Division of the National Science Foundation for a research program devoted to the defect chemistry of perovskite-related oxides has provided essential experience in this small corner of science. Collaboration, discussions, and arguments with my many good friends and colleagues in the field have added greatly to my understanding. I am grateful to all of them.

REFERENCES

- 1 R. G. Fuller, in: J. H. Crawford, Jr. and L. M. Slifkin (Eds.), *Point Defects in Solids*, Plenum, New York, 1972, pp. 103-150.
- 2 M.J.L. Sangster and D. K. Rowell, *Phil. Mag.*, A44 (1981) 613-624.
- 3 D.M. Smyth, *J. Electrochem. Soc.*, 106 (1959) 635-639.
- 4 S. C. Abrahams and P. Marsh, *Acta Cryst.*, B42 (1986) 61-68.
- 5 D.M. Smyth, *Ferroelectrics*, 50 (1983) 419-428.
- 6 C.R.A. Catlow, personal communication.
- 7 P. Kofstad, *J. Less-Common Metals*, 13 (1967) 635-638.
- 8 P. Kofstad, *Corrosion*, 24 (1968) 379-388.
- 9 R. Dieckmann and H. Schmalzried, *Ber. Bunsenges. physik. chem.*, 81 (1977) 344-347.
- 10 H.-G. Sockel and H. Schmalzried, *Ber. Bunsenges. physik. chem.*, 72 (1968) 745-754.
- 11 E.K. Chang, D.J.L. Hong, A. Mehta, and D.M. Smyth, *Mat. Letters*, 6 (1988) 251-253.
- 12 A. Mehta and D.M. Smyth, in: J. Nowotny and W. Weppner (Eds.), *Proceedings of the NATO Advanced Research Workshop on Low Dimensional Defects in Non-Stoichiometric Compounds*, Kluwer, Dordrecht, in press.
- 13 D.M. Smyth, *Prog. Solid St. Chem.*, 15 (1984) 145-171.
- 14 A.M.J.H. Seuter, *Philips Res. Repts.*, Suppl. No. 3 (1974) 1-84.
- 15 F. Farhi and G. Petot-Ervas, *J. Phys. Chem. Solids*, 39 (1978) 1169-1173.
- 16 W.J. Minford and V.S. Stubican, *J. Am. Ceram. Soc.*, 57 (1974) 363-367.
- 17 F. A. Kroger, *J. Am. Ceram. Soc.*, 49 (1966) 215-218.
- 18 D.L. Kirk and P.L. Pratt, *Proc. Brit. Ceram. Soc.*, 9 (1967) 215.
- 19 G.D. Watkins, *Phys. Rev.*, 113 (1959) 79-90.
- 20 S. Andersson, B. Collen, U. Kuylenstierna, and A. Magneli, *Acta Chem. Scand.*, 11 (1957) 1641-1652.
- 21 S. Andersson, A. Sundholm, and A. Magneli, *Acta Chem. Scand.*, 13 (1959) 989-997.
- 22 D.M. Smyth, *Ann. Rev. Mater. Sci.*, 15 (1985) 329-357.
- 23 D.M. Smyth, in: *Crystal Lattice Defects and Amorphous Materials*, *Proceedings of the Intern. Conf. on Crystals*, Parma, Italy, 29 August - 2 September, 1988, Gordon and Breach, in press.
- 24 C.R.A. Catlow, in: J.M. Cowley, J.B. Cohen, M.B. Salamon, and B.J. Wuench (Eds.), *AIP Conf. proc. No. 53, Modulated Structures - 1979*, Am. Inst. Phys. New York, 1979, pp. 149-161.