Theoretical investigation to thermal equilibrium concentration of point defect through first-principles calculation

Akihide Kuwabara

Japan Fine Ceramics Center, 2-4-1 Matsuno, Atsuta, Nagoya 456-8587, Japan

Received 17 June 2007; received in revised form 10 September 2007; accepted 18 September 2007
Available online 19 November 2007

Abstract

First-principles calculations based on density functional theory enable us to quantify the thermal equilibrium concentration of a point defect through calculating the formation energy of the defect. In order to know the formation energy, we need three parameters: the energy difference between defective and perfect systems, chemical potential of constitutive atoms and Fermi level. The energy difference is obtained from calculations based on “supercell method”. The energy of a supercell with a charged defect should be corrected by an appropriate way such as alignment of electrostatic potential. The chemical potential and Fermi level can be fixed from conditions of phase equilibrium and charge neutrality.

Keywords: Density functional theory; Supercell; Point defect; Thermal equilibrium concentration

1. Introduction

Point defect is one of the key factors to determine properties and characteristics in a crystalline material. As is well known, point defect is the origin of optical property like a color center. Donor-type and acceptor-type doping is necessary for the production of n-type and p-type semiconductors, respectively. In addition to these things, one of the most important phenomena related to defect chemistry is atomistic diffusion. Diffusion is strongly dependent on the formation of point defect because it occurs through movement of point defects. Ionic devices like fuel cells, Li-ion batteries, and gas sensors are recent key technology of power resources. It is significant to understand defect formation behavior in the materials used for ionic devices because these devices actively use the diffusion phenomenon to exhibit their functional property. Diffusion also controls the processing of materials like sintering, solid solution, and precipitation. Thus, a vast array of experiments have been conducted to figure out point defects so far.

One of the major difficulties in an experiment for defect chemistry is that defect concentration is very sensitive to impurities in original materials and atmosphere where specimens are made. These parameters should be strictly controlled. The difficulty causes scattering results among the literatures. This is obstructive to comprehension of diffusion phenomena. Recently, theoretical investigation of atomistic simulations has become powerful tool to study the defect chemistry because enhancement of computing power makes first-principles calculations for defect structures possible in realistic time. Especially, the density functional theory (DFT) calculation [1] is well suited for analysis of physical and chemical properties of condensed matters. This paper shows theoretical methodology of modeling point defect through DFT calculations and also surveys a way of quantitative estimation for thermal equilibrium concentration of a point defect.

2. Formalism of defect formation energy

Formation of a point defect is a thermally activated process. The thermal equilibrium concentration of a point defect, $C_{\text{defect}}$, is expressed as

$$C_{\text{defect}} = Z_{\text{defect}} N_{\text{defect}} \exp \left( - \frac{\Delta G_{\text{form}}}{k_B T} \right),$$

(1)
where \( \Delta G_{\text{form}} \) is Gibbs formation free energy of a point defect. \( Z_{\text{defect}} \) is the number of equivalent configurations for the defect and \( N_{\text{defect}} \) is the total number per unit volume of sites the defect can occupy. \( k_B \) and \( T \) are the Boltzmann constant and absolute temperature, respectively. In order to quantitatively know \( C_{\text{defect}} \), the value of \( \Delta G_{\text{form}} \) should be known.

In thermodynamics, the formation energy of surface or interface is defined as the excess energy to form defect structures into a perfect crystal. This formalism was adopted for calculating formation energy of a point defect by Zhang and Northrup [2]. Many theoretical investigations based on their formalism have been reported from various groups [3–13].

According to the thermodynamical definition, the defect formation energy can be given by

\[
\Delta G_{\text{form}} = G_{\text{defect}} - \sum_i n_i \mu_i + q \mu_e,
\]

where \( G_{\text{defect}} \) is the Gibbs free energy of a defective system and \( n_i \) is the number of atom \( i \) in the defective system. \( \mu_i \) is a chemical potential of an atom \( i \) and \( \mu_e \) is the Fermi energy, namely the chemical potential of an electron. \( q \) is a charge state of the defect.

In the case of a stoichiometric and uncharged defect like Frenkel or Shottky pair, the defect formation energy is independent of chemical potential and Fermi energy. In the case of a Shottky pair, \( V_M^0 V_X^0 \), in a perfect system composed of \( N \) formula units, its formation energy is

\[
\Delta G_{\text{form}}(V_M^0 V_X^0) = G_{\text{defect}}(M_{N-1}X_{N-1}, q = 0) - (N-1)\mu_M - (N-1)\mu_X. \tag{3}
\]

Gibbs free energy of MX in a perfect crystal has a relationship with chemical potentials of M and X

\[
G_{\text{perfect}}(M_X N_Y) = N_M \mu_M + N_X \mu_X. \tag{4}
\]

Thus, Eq. (3) is rewritten using Eq. (4) as

\[
\Delta G_{\text{form}}(V_M^0 V_X^0) = G_{\text{defect}}(M_{N-1}X_{N-1}, q = 0) - \frac{(N-1)}{N} G_{\text{perfect}}(M_X N_Y). \tag{5}
\]

The formation energy of a Frenkel pair of X is

\[
\Delta G_{\text{form}}(X^0_i X^0_i) = G_{\text{defect}}(M_N N_X, q = 0) - G_{\text{perfect}}(M_X N_Y). \tag{6}
\]

\( G_{\text{defect}} \) is obtained from a calculation of a “supercell” model with a defect. The supercell model is an assembly of unit cells and is constructed by expanding a unit cell in directions of lattice vector. Adequate dimensions of a supercell should be chosen to eliminate finite size errors when a supercell including a defect is constructed. This point will be discussed in the latter section of this paper. \( G_{\text{perfect}} \) is free energy of perfect crystal and should be calculated from a perfect supercell model which has similar dimensions to the defective one for keeping numerical errors of the calculations at the same level. At this point, one major problem is that DFT calculation is based on the Born–Oppenheimer approximation [1]. The motion of nucleus is neglected in this approximation, namely the effects of zero-point vibration and finite temperatures are not taken into account. Therefore, basic DFT calculation cannot calculate Gibbs free energy including entropy of vibration without additional methodology like frozen phonon [14], linear response method [15–16], and so on. However, it is reasonable to assume that the difference of entropy and volume between defective and perfect systems is negligible. Based on this assumption, the total energy from DFT calculation is allowed to be substituted for Gibbs free energy in Eqs. (5) or (6). In the following section, defect formation energy is represented by Gibbs free energy in a precise expression. It should be noted that total energy from DFT calculation will be used in an actual estimation of defect concentration.

When a defect causes nonstoichiometry in the system and has a charge state, the defect formation energy depends on the chemical potential of constitutive atoms, \( \mu_M \) and \( \mu_X \). The values of these parameters need to be known for quantifying defect formation energy. First, neutral defects in a simple binary compound of \( M^+X^- \) are described as an example. When an anion vacancy (\( V_X^0 \)) is introduced into a perfect crystal composed of \( N \) formula units, the formation energy of \( V_X^0 \), \( \Delta G_{\text{form}}(V_X^0) \), is given by

\[
\Delta G_{\text{form}}(V_X^0) = G_{\text{defect}}(M_N X_{N-1}, q = 0) - N \mu_M - (N-1)\mu_X. \tag{7}
\]

Eq. (7) is rewritten using Eq. (4) as

\[
\Delta G_{\text{form}}(V_X^0) = G_{\text{defect}}(M_N X_{N-1}, q = 0) - G_{\text{perfect}}(M_X N_Y) + \mu_X. \tag{8}
\]

In the case of interstitial X (\( X_i^0 \)), formation energy is

\[
\Delta G_{\text{form}}(X_i^0) = G_{\text{defect}}(M_N X_{N+1}, q = 0) - G_{\text{perfect}}(M_X N_Y) - \mu_X. \tag{9}
\]

In order to obtain the formation energies of these nonstoichiometric defects, one needs the value of chemical potential of X. Under an equilibrium condition, chemical potential has limited value. In the M-rich limit, MX coexists with the elementary substance of M. Under this condition, \( \mu_M \) reaches a maximum. The M-rich limit directly corresponds to the X-low limit and \( \mu_X \) reaches a minimum. X-rich limit is the phase equilibrium state between MX and elementary substance of X. After that, the region of chemical potential for X is

\[
\mu_{X}^{\text{element}} < \mu_X < \mu_{X}^{\text{element}}. \tag{10}
\]

\( \mu_M \) and \( \mu_X \) are not independent parameters and are in relationship with

\[
\mu_M + \mu_X = \mu_{MX}. \tag{11}
\]

where \( \mu_{MX} \) is total energy of MX per molecule. \( \mu_X \) is uniquely determined if the value of \( \mu_M \) is fixed. Fig. 1 is a schematic image which illustrates the dependence of defect formation energies of \( V_X^0 \) and \( X_i^0 \). The formation energy of
$X^+_i$ becomes lower at the X-rich side than at the M-rich side. This means that $X^+_i$ more easily forms at the X-rich side than at the M-rich side. On the contrary, $V^-_X$ is likely to form at M-rich limit rather than at X-rich side.

3. Equilibrium with gas phase

In materials like oxide and nitride, thermal equilibrium with gas phase influences on defect formation behavior. With regard to oxide, reduction atmosphere, i.e., low oxygen partial pressure ($p_{O_2}$), make oxide ion vacancy increase. On the contrary, oxidization atmosphere, that is high $p_{O_2}$, leads to reduction in the population of oxide vacancy and is likely to form excess oxide ions like interstitial oxide ions or cation vacancies. For these materials, it is significant to know the defect formation behavior depending on ambient temperature and pressure. As mentioned before, basic DFT calculation is based on the Born–Oppenheimer approximation and cannot directly include the influence of temperature and partial pressure of gas phase. Reuter and Scheffler [17] have proposed a solution to this problem and reported the dependence of RuO$_2$(1 1 0) surface structure on the chemical potential of O$_2$ molecule. They take into account the temperature and partial pressure effect of O$_2$ gas molecule on surface stability. This method is easily applied to a point defect in bulk. If a material is in thermal equilibrium with gaseous O$_2$, chemical potential of oxygen can be written by the Gibbs free energy of gas molecule as

$$\mu_{O} = \frac{1}{2} G_{O(g)}.$$  \hspace{1cm} (12) 

Therefore, the defect formation energy of neutral oxide ion vacancy ($V^-_O$) in an MO$_x$ compound is given by

$$\Delta G_{\text{form}}(V^-_O) = G_{\text{defect}}(M_N O_{N_x-1}, q = 0) - G_{\text{perfect}}(M_N O_{N_x}) + \frac{1}{2} G_{O(g)}. \hspace{1cm} (13)$$

Based on the condition that O$_2$ molecule is an ideal gas, the Gibbs free energy of O$_2$ at given temperature $T$ and partial pressure $p$ is expressed by

$$G_{O_2(g)}(T, p) = G^e(T, p^e) + k_B T \ln \frac{p}{p^e}. \hspace{1cm} (14)$$

$G^e(T, p^e)$ is Gibbs free energy at a temperature $T$ and a standard pressure $p^e$. $G^e(T, p^e)$ measured from 0 K is expressed as

$$G^e(T, p^e) = G^e(0 \text{ K}, p^e) + G^e(T, p^e) - G^e(0 \text{ K}, p^e)$$

$$= H^e(0 \text{ K}, p^e) + \{H^e(T, p^e) - H^e(0 \text{ K}, p^e)\}$$

$$- TS(T, p^e). \hspace{1cm} (15)$$

The relation of $G = H - TS$ is applied to the equation above. If 1 atm is chosen as a standard pressure $p^e$, the temperature dependence of enthalpy and entropy can be referred from a thermochemical table [18]. Assuming that the total energy obtained from DFT is equal to the enthalpy at 0 K, we can calculate chemical potential of oxygen atom at arbitrary temperature and pressure using Eqs (14) and (15). Such a methodology has succeeded in revealing nonstoichiometry at surface [17,19] and point defect formation energy in oxides [11,13,20,21]. In the studies reported by Sundell et al., they have clarified the Gibbs free energy of defect formation in BaZrO$_3$ by taking into account vibrational entropy [13,20]. It has been revealed that the temperature dependence of Gibbs energy of defect formation mainly derives from the change in the chemical potential of a gaseous molecule. This result supports usefulness of the analysis explained in this section. However, it has also been reported that the hierarchy of Gibbs formation free energies of point defects upsets along with elevating temperature in Si [22]. It should be noted that it is possible for vibrational entropy to play an important role in the defect chemistry.

4. Charged defect

In addition to the chemical potential, the formation energy of a charged defect is dependent on the Fermi energy as shown by Eq. (2). The formation energy of $V^-_X$ is

$$\Delta G_{\text{form}}(V^-_X) = G_{\text{defect}}(M_N X_{N_x-1}, q = +1) - G_{\text{perfect}}(M_N X_{N_x}) + \mu_X + \mu_e. \hspace{1cm} (16)$$

In addition to $\mu_X$, Fermi level $\mu_e$ should be known. In experimental, Fermi energy is measured from a vacuum level. However, most of the DFT calculations for a crystal system investigate electronic structures of an infinite crystal under a periodic boundary condition. In such case, a vacuum level cannot be uniquely determined [23]. Alternatively, Fermi energy is measured from the energy of valence band top, $E_{\text{VBT}}$. As a result, Fermi energy is expressed as

$$\mu_e = E_{\text{VBT}} + E_f \hspace{1cm} (0 \leq E_f \leq E_g), \hspace{1cm} (17)$$

where $E_g$ is band gap. $E_{\text{VBT}}$ is determined from the energy difference between neutral and positively charged (+1) supercells of the host material without any point defects. Defect formation energy depends on $E_f$. From Eq. (2), it is straightforwardly shown that the formation energy of a positively charged defect, namely a donor type, become higher when Fermi level gets close to the conduction band.
bottom. It is more difficult for a donor-type defect to form. On the other hand, an acceptor-type defect becomes preferable. This behavior is represented in Fig. 2. If the Fermi level is located at the side of the valence band top, this defect formation tendency becomes reversed. This behavior is one of the origins for “doping limit” in semiconductors [10]. Doping of p-type acceptor dopant makes Fermi level shift to the valence band top and also increases the concentration of holes. However, if doping concentration increases to such an extent that the Fermi level becomes closer to the valence band top, native donor-type defects of the host material spontaneously form and compensate for the charge of acceptor-type dopant. Doping is therefore no longer effective to increase the concentration of holes. The same thing must happen for doping of donor-type dopant. Over the limit of doping, native acceptor defects compensate for the charge of donor-type dopant. This compensation will stop the increase of electron concentration.

As shown in Fig. 2, there is a crossing point for the energy dependence of neutral and charged defects on Fermi energy. This point represents a defect level induced by a specific defect in a band gap. For example, the defect formation energy of \( V_X^\ast \) can be expressed by that of charged one, \( V_X^\ast \), and the defect level from neutral (0) to a positive lycharge state (+1), \( \varepsilon(0/1) \), as follows:

\[
\Delta G_{\text{form}}(V_X^\ast) = \Delta G_{\text{form}}(V_X^\ast) + \varepsilon(0/1) - E_T. 
\]  

(18)

The defect level, \( \varepsilon(0/1) \), is located at the position of Fermi level in a band gap where \( \Delta G_{\text{form}}(V_X^\ast) \) becomes equal to \( \Delta G_{\text{form}}(V_X^\ast) \). Generally, the defect level from a charged state to another state is obtained by the equation:

\[
\varepsilon(q/q') = \frac{G_{\text{defect}}(q') - G_{\text{defect}}(q)}{q - q'}. 
\]  

(19)

If we try to evaluate the formation energy of charged defects, how can we determine Fermi energy? A simple way is to assume the material as a simple system where the Fermi energy is located at valence band top. In this condition, \( E_T \) is zero. In n-type limit, the Fermi energy is at conduction band bottom. This situation corresponds to \( E_T = E_g \). If wide band gap materials are objective, the environment of insulator can be assumed by the setting of \( E_T = 1/2E_g \). After fixing Fermi level, defect formation energy is calculated from Eq. (2). However, it should be noted that thermal equilibrium condition and electrical neutrality is not considered in this method. Therefore, one can only quantitatively estimate which defect is preferable rather than others.

In the other way, the Fermi energy can be “self-consistently” determined through the view point of defect chemistry. The charge neutrality condition is described by

\[
\sum_{\text{defect}} q_{\text{defect}} C_{\text{defect}} + n_h - n_e = 0, 
\]  

(20)

where \( n_e \) and \( n_h \) are concentrations of electron and hole, respectively. These are calculated from effective valence band and conduction band density [24]. Fermi energy under the given condition of chemical potentials \( \mu \) and temperature \( T \) can be obtained with fulfilling Eqs. (1), (2) and (20). Once the value of Fermi energy is determined, equilibrium concentration of point defects is deductively evaluated from Eqs. (1) and (2) using the fixed Fermi level.

5. Artificial error in the calculation of charged defect

Most of DFT calculations deal with structure models under a periodic boundary condition. In such a periodic system, charge interaction between images of a supercell should diverge. In the calculation for charged system, a jellium background [25] neutralizing a charged system is introduced to avoid divergence of electrostatic potential. However, the jellium background consequently affects electronic structure and total energy because of the finite size of a supercell.

Makov and Payne [26] have reported the dependence of total energy in charged systems on the size of a cubic supercell

\[
E = E_0 - \frac{q^2\pi}{2L} - \frac{2q\pi Q}{3L^3} + O(L^{-5}), 
\]  

(21)

where \( E_0 \) is the required energy of an isolated defect. The second term is a Madelung potential between the supercell images. \( L \) is a linear dimension of supercell. \( q \) is a charge state of a defect and \( \pi \) is Madelung constant. The third term is interaction between the point charge \( q \) and quadrupole moment of electronic density, \( Q \). The last term gives the higher orders of the interaction. If the second and third terms are calculated, \( E_0 \) is directly obtained. Alternatively, \( E_0 \) is evaluated through total energy calculations for a charged defective system in various supercell sizes and least-square fitting of the total energies to the function of \( 1/L \) and \( 1/L^3 \). It should be reminded that Eq. (21) can be applied only to cubic supercell in a precise sense.

Potential alignment [3,7] is a simple way of error correction. The unfavorable coulomb interaction causes \( E_{\text{VBT}} \) and electrostatic potential in a defective supercell to
shift. In a supercell with a point defect, a region far enough from the point defect can be regarded as a bulk-like one. The energetical shift approximately corresponds to the difference of electrostatic potentials between a bulk-like region in a defective supercell and a perfect crystal. Assumed that local electronic structures of a bulk-like region in a defective supercell are same as that of a perfect crystal, the $E_{VBT}$ position is corrected by the alignment of electrostatic potential by this equation:

$$E_{VBT} = E_{VBT}^{bulk} + \bar{V}_{av}^{defect} - \bar{V}_{av}^{bulk}, \quad (22)$$

where $\bar{V}$ is an average electrostatic potential within a certain area. The advantage of this correction is that the shape of a supercell does not matter.

6. Conclusion

This paper has summarized the theoretical quantification of point defects through first-principles calculations. The defect formation energy is evaluated by a supercell method using DFT calculations. From the thermodynamical definition, defect formation energy is dependent on chemical potential and Fermi energy. The exact value of a chemical potential cannot be determined, but the lowest and highest limits of a chemical potential are fixed from a phase equilibrium condition with elementary substances of constitutive atoms. If the elementary substance is a gaseous molecule, one would like to consider the effect of temperature and partial pressure on defect formation. In this case, the chemical potential of a gas phase can be calculated at arbitrary temperature and partial pressure by the combination of total energy from DFT calculations with temperature dependence of enthalpy and entropy referred from thermochemical tables. If the values of chemical potentials are fixed and electrical neutrality is considered, Fermi level is uniquely determined. Consequently, we can obtain thermal equilibrium concentration of point defects. A correction of total energy is needed for a charged and defective supercell because finite size of the supercell under a periodic boundary condition causes unrealistic Coulomb interaction in the model. In order to suppress this kind of error, suitable dimensions of a supercell should be chosen. If necessary, error corrections should also be done. Alignment of electrostatic potentials is a way of error correction with general versatility.

Though there is still a limitation of computer simulations to handle the system, recent development of computational materials science makes it possible to more quantitatively estimate formation behavior of point defects. Especially, calculation of phonon allows one to do theoretical analysis actually at finite temperatures. Further computations are expected to provide fruitful insight to the defect chemistry.

References

[1] W. Kohn, L.J. Sham, Phys. Rev. A 140 (1965) 1133.
[2] S.B. Zhang, J.E. Northrup, Phys. Rev. Lett. 67 (1991) 2339.
[3] D.B. Laks, C.G. Van de Walle, G.F. Neumark, Phys. Rev. B 45 (1992) 10965.
[4] A. Garcia, J.E. Northrup, Phys. Rev. Lett. 74 (1995) 1131.
[5] S. Pöyskkö, M.J. Puska, R.M. Nieminen, Phys. Rev. B 53 (1996) 3813.
[6] S.B. Zhang, S.-H. Wei, A. Zunger, H.K. Yoshida, Phys. Rev. B 57 (1998) 9642.
[7] T. Mattila, A. Zunger, Phys. Rev. B 58 (1998) 1367.
[8] A.F. Kohan, G. Ceder, D. Morgan, C.G. Van de Walle, Phys. Rev. B 61 (2000) 15019.
[9] C. Khé, A. Zunger, Phys. Rev. Lett. 88 (2002) 095501.
[10] S.B. Zhang, J. Phys.: Condens. Matter 14 (2002) R881.
[11] A. Kuwabara, I. Tanaka, J. Phys. Chem. B 108 (2004) 9168.
[12] K. Matsunaga, T. Tanaka, T. Yamamoto, Y. Ikuhara, Phys. Rev. B 68 (2003) 085110.
[13] P.G. Sundell, M.E. Björketun, G. Wahnström, Phys. Rev. B 73 (2006) 104112.
[14] K. Parlinski, Z.-Q. Li, Y. Kawazoe, Phys. Rev. Lett. 78 (1997) 4063.
[15] X. Gonze, Phys. Rev. A 52 (1995) 1096.
[16] S. Baroni, S. de Gironcoli, A.D. Corso, P. Giannozzi, Rev. Mod. Phys. 73 (2001) 515.
[17] K. Reuter, M. Scheffler, Phys. Rev. B 65 (2001) 034506.
[18] M.W. Chase Jr., NIST-JANAF Thermochemical Tables, forth ed, The American Institute of Physics for The National Institute of Standards and Technology, New York, 1998.
[19] W. Berghmayer, I. Tanaka, Appl. Phys. Lett. 84 (2004) 909.
[20] M.E. Björketun, P.G. Sundell, G. Wahnstrom, Faraday Discuss. 134 (2007) 247.
[21] K. Matsunaga, A. Kuwabara, Phys. Rev. B 75 (2007) 014102.
[22] O.K. Al-Mushadani, R.J. Needs, Phys. Rev. B 68 (2003) 235205.
[23] L. Kleinman, Phys. Rev. B 24 (1981) 7412.
[24] J.S. Blakemore, J. Appl. Phys. 53 (1982) R123.
[25] C.G. Van de Walle, P.J.H. Denteneer, Y. Bar-Yam, S.T. Pantelides, Phys. Rev. B 39 (1989) 10791.
[26] G. Makov, M.C. Payne, Phys. Rev. B 51 (1995) 4014.