Hydrogen in semiconductors and insulators

Chris G. Van de Walle *

Materials Department, University of California, Santa Barbara, CA 93106-5050, USA

Abstract

We describe a first-principles methodology for studying the interactions between hydrogen and materials. In non-metals, isolated hydrogen atoms occur as electrically charged impurities, and this charge has profound effects on the incorporation of hydrogen into the material. We have developed a model that enables us to predict the electrical properties of hydrogen in any material, based on its band alignment on an absolute energy scale. We define the key quantities that describe the interaction between hydrogen and the host materials, and discuss how fundamental knowledge of this type can be applied to hydrogen storage materials.
© 2006 Elsevier B.V. All rights reserved.

1. Introduction

Hydrogen plays a very important role in many semiconductors. It is an indispensable ingredient in integrated-circuit fabrication, since it is used to passivate defects at the Si/SiO_2 interface that is present in every CMOS transistor [1]. While eliminating these detrimental defects is beneficial, introduction of hydrogen has other consequences. Among these is passivation of electrically active impurities (donors and acceptors) through formation of hydrogen-impurity complexes [2]. First-principles calculations have been very helpful in elucidating the fundamental mechanisms of hydrogen incorporation and interactions. Quantities that can be computed include the binding energy of hydrogen-impurity complexes and the diffusion barrier for interstitial hydrogen; based on this information the temperatures at which complex dissociation and hydrogen removal will take place can be estimated. Frequencies of local vibrational modes can also be calculated, and allow a direct identification of complexes responsible for features observed with vibrational spectroscopy [3].

Hydrogen almost never occurs as an electrically neutral impurity inside a semiconductor or insulator. It invariably either gives up its electron, becoming positively charged and acting as a tively charged and acting as an acceptor. In most semiconductors hydrogen therefore acts as an *amphoteric impurity* (an impurity that can be either a donor or an acceptor). Which of these two prevails depends on the Fermi level (E_F), i.e., the chemical potential of electrons. In a semiconductor or insulator, the Fermi level can be located anywhere between the valence-band maximum (VBM) and the conduction-band minimum (CBM). E_F is located near the VBM in p-type material, which is doped with acceptors; E_F is located near the CBM in n-type material, which is doped with donors.

In silicon, we found that hydrogen acts as a donor (H^+) in p-type material and as an acceptor (H^-) in n-type material, thus always *counteracting* the prevailing conductivity [4,5]. First-principles calculations for a number of other semiconductors, including GaAs and wide-band-gap materials such as ZnSe and GaN [6], produced similar results, seemingly confirming this behavior as a general feature of hydrogen's interactions with semiconductors. It therefore came as a surprise when calculations showed that H behaves exclusively as a *donor* in ZnO [7]. Only the positive charge state, H^+, occurs in ZnO, and other charge states cannot be stabilized. Hydrogen can therefore act as a *source* of doping, rather than merely reducing the conductivity introduced by other dopants. These first-principles results for ZnO have triggered dozens of experimental investigations that have confirmed the predictions. A partial overview of the experimental studies can be found in Ref. [8].

---

* Tel.: +1 805 893 7144; fax: +1 805 893 8983.
E-mail address: vandewalle@mrl.ucsb.edu.

The realization that hydrogen acts as an electrically active dopant impurity in ZnO has important consequences for device development, since the fabrication of electronic and optoelectronic devices requires strict control over doping of the layers. The unintentional incorporation of an impurity that acts as a dopant can severely impact the materials properties. Hydrogen is a particularly insidious impurity in this respect, since many growth techniques either intentionally or unintentionally introduce hydrogen: it is often used as a carrier gas, it can be part of the source gases, and it can easily be adsorbed on the walls of growth chambers. In addition, many processing steps (such as annealing in forming gas) may introduce hydrogen into the material [9,10].

The qualitatively different behavior of hydrogen in ZnO, compared to the semiconductors that were previously studied, raised important questions about the underlying physics. We strongly believed that understanding the fundamentals of hydrogen interactions with these semiconductors would lay the foundation for predicting the behavior in a much wider range of materials. These investigations have now shown that a "universal alignment level" exists, which allows predicting the electrical behavior of hydrogen in a wide range of materials. We will discuss the universal alignment in Section 3, but first we will provide a brief review of the framework for addressing the energetics and electronic structure of impurities in Section 2. In Section 4 we will discuss the application of the formalism to hydrogen storage materials.

2. Methodology

The key quantities that determine the properties of hydrogen impurities are: (i) the formation energy, i.e., the energy needed to incorporate the impurity into the host, and (ii) the electronic transition level, which defines the electrical behavior. These quantities can be obtained from first-principles calculations. The approach is based on density-functional theory (DFT) [11] within the local density approximation (LDA) [12] and the pseudopotential-plane-wave method. Our calculations for interstitial hydrogen in semiconductors are typically carried out in supercells containing either 64 atoms for zinc-blende-structure materials or 96 atoms for wurtzite. For hydrogen the Coulomb potential is used. Energy differences for hydrogen-containing systems are well converged at a 40 Ry cutoff in a plane-wave expansion of wave functions and charge densities. Relaxation of the host atoms is always included, and for each charge state of hydrogen many

possible sites in the lattice are explored and the global minimum is identified. Further computational details can be found in Refs. [13,14].

We obtain the formation energy of interstitial hydrogen in charge state q (where $q = -1$, 0, or $+1$) by placing the hydrogen impurity in a volume of host material, calculating the total energy $E_{tot}(H^q)$ of this structure, and subtracting the energy $E_{tot}(bulk)$ of a corresponding volume of pure host material [4,9,13,14]:

$$E^f(H^q) = E_{tot}(H^q) - E_{tot}(bulk) - \tfrac{1}{2}E_{tot}(H_2) + qE_F. \qquad (1)$$

Here the reference for the hydrogen energy is given by an H_2 molecule at $T = 0$. At finite temperatures and pressures, this term should be replaced by the chemical potential of hydrogen; the temperature and pressure dependence of the other terms in Eq. (1) can also be included, but the net effect on the formation energy is typically quite small. The last term in the formation energy accounts for the fact that H^+ donates an electron, and H^- accepts an electron; the energy of the reservoir with which these electrons are exchanged is the electron chemical potential or Fermi level, E_F.

Defects and impurities usually introduce levels in the band structure of the host; these *transition levels* can be obtained based on the total energies that we calculate from first principles. The transition level $\varepsilon(q + 1/q)$ is defined as the Fermi-level position for which the formation energies of charge states $q + 1$ and q are equal. When the Fermi level is below the transition level, charge state $q + 1$ is stable; when the Fermi level moves above the transition level, charge state q is stable. Since the possible charge states for hydrogen are $+1$, 0, and -1, the relevant transition levels are $\varepsilon(+/0)$ and $\varepsilon(0/-)$. However, hydrogen behaves as a negative-U impurity in virtually all semiconductors studied so far, meaning that the donor level $\varepsilon(+/0)$ occurs *above* the acceptor level $\varepsilon(0/-)$. This negative-U character implies that the neutral charge state is always higher in energy than either the positive or the negative charge state. As the Fermi level is swept through the band gap, the stable charge state changes from H^+ for Fermi levels low in the gap to H^- for Fermi levels high in the gap, without H^0 ever being thermodynamically stable.

Negative-U behavior is usually related to the presence of unusual lattice relaxations. This is indeed the case for hydrogen. In the negative charge state, hydrogen prefers to stay as far away as possible from the electronic charge in the host material, and usually ends up sitting in an interstitial site surrounded by cations (in a compound). In the positive charge state, on the other hand, the proton likes to take advantage of Coulomb attraction by sitting closely to an anion. This switch in positions explains why, as more electrons are added to the hydrogen level (going from + to 0 to − charge states), the energy does not necessarily increase, and the $\varepsilon(0/-)$ acceptor level can actually be lower than the $\varepsilon(0/+)$ donor level.

Fig. 1(a) illustrates these concepts with the example of hydrogen in ZnSe, a wide-band-gap semiconductor with a band gap of 2.7 eV. H^0 is never the lowest-energy state (reflecting the *negative-U* property), causing the donor level $\varepsilon(+/0)$ to lie above the acceptor level $\varepsilon(0/-)$. When E_F moves through the band gap, the stable charge state thus changes directly from positive (for E_F below 1.4 eV) to
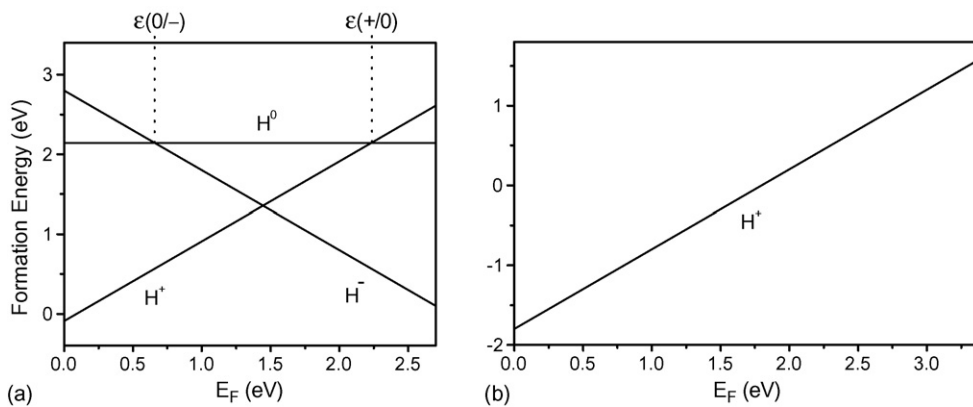

Fig. 1. Calculated formation energy of interstitial hydrogen as a function of Fermi level in two different semiconductors, illustrating the qualitatively different position of the $\varepsilon(+/-)$ level. In ZnSe (a), $\varepsilon(+/-)$ lies within the band gap, at about 1.4 eV above the valence-band maximum (VBM), while in ZnO (b) $\varepsilon(+/-)$ lies above the conduction-band minimum (CBM). Energies are obtained from DFT-LDA calculations, and the range of E_F corresponds to the band gap, with $E_F = 0$ at the VBM.
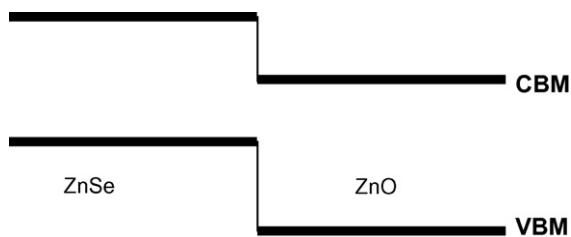
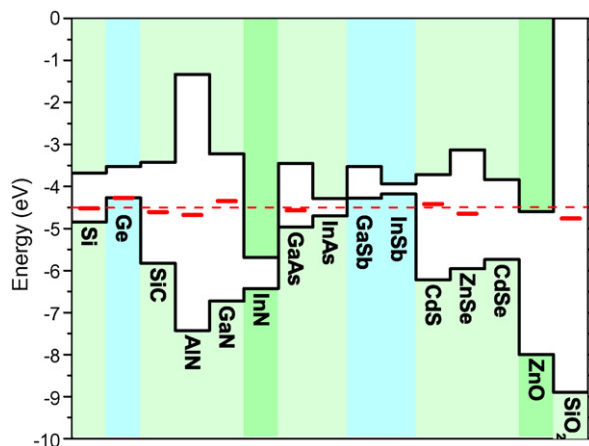Fig. 2. Calculated band alignment between ZnSe and ZnO.

Fig. 3. Band alignments and position of the hydrogen $\varepsilon(+/-)$ level for selected semiconductors and oxides. For each semiconductor, the lower line indicates the position of the valence-band maximum and the upper line the position of the conduction-band minimum. The thick bar shows the position of $\varepsilon(+/-)$ (calculated with density-functional theory) with respect to the valence-band maximum, calculated for selected semiconductors. The dashed line at $-4.5\,eV$ indicates the "universal alignment" level and provides a prediction for the position of the $\varepsilon(+/-)$ level in those semiconductors for which it was not explicitly calculated. Materials in which hydrogen is predicted to act as a shallow dopant are indicated by shading.

negative (for E_F above 1.4 eV). This implies that in p-type ZnSe [E_F close to the valence-band maximum (VBM)] H$^+$ is favored, while in n-type ZnSe [E_F close to the conduction-band minimum (CBM)] H$^-$ is stable, providing the basis for hydrogen's tendency to counteract the prevailing conductivity. The Fermi-level position where the positive and negative charge states are equal in energy is labeled $\varepsilon(+/-)$, and plays a crucial role in our theory. This behavior of hydrogen in ZnSe is qualitatively very similar to what has been observed or calculated for other materials [2,14], although quantitatively the values of formation energies and transition levels vary over a large range.

In semiconductors such as Si, GaAs, ZnSe, or GaN, hydrogen is an amphoteric impurity, implying that the $\varepsilon(+/-)$ transition level occurs within the band gap and both H$^+$ and H$^-$ can, in principle, occur. Fig. 1(b) shows that in ZnO, however, only H$^+$ is stable, meaning that $\varepsilon(+/-)$ occurs above the conduction-band minimum (CBM). We will elaborate on the cause of this behavior in Section 3.

3. Universal alignment

In Fig. 1 we showed two examples of calculated formation energies and transition levels for hydrogen in semiconductors. In ZnSe, the $\varepsilon(+/-)$ occurs roughly in the middle of the band gap, and hydrogen can act as either a donor or an acceptor, depending on the Fermi-level position. In ZnO, on the other hand, $\varepsilon(+/-)$ is above the CBM and hydrogen can only act as a donor. At first sight it is very puzzling that in ZnO the $\varepsilon(+/-)$ level occurs at a much higher energy (above the CBM) than it does in ZnSe. The puzzle is resolved, however, by taking the band alignment of these materials into account. Indeed the valence-band offset between ZnSe and ZnO is very large, more than 2.0 eV, with the valence band of ZnO occurring at a much lower energy, as illustrated in Fig. 2. Therefore the $\varepsilon(+/-)$ level can be considered as essentially constant (on an "absolute energy scale") when going from ZnSe to ZnO.

These band alignments between semiconductors can be calculated using first-principles calculations, using techniques similar to those described in Section 2. The methodology is addressed in Refs. [15–17]. We have performed such calculations for a large number of interfaces. The resulting alignments can be combined with our calculated results for the hydrogen $\varepsilon(+/-)$ level in each of the materials. The results [18] show that the $\varepsilon(+/-)$ values exhibit remarkable consistency in terms of their absolute energy across a wide range of materials. An example of the resulting diagram is shown in Fig. 3. The spread in the values around this "universal alignment level" is very narrow in light of the fact that the band edges of the materials (including oxides such as SiO$_2$) span a range of more than 9 eV.

A number of interesting predictions have already emerged from Fig. 3. For instance, it shows that the $\varepsilon(+/-)$ level for InN lies above the CBM, and therefore hydrogen in InN should

behave exclusively as a donor, a prediction that has already been confirmed [19]. Fig. 2 also indicates that it is possible for hydrogen to act as a source of p-type doping: in Ge and GaSb hydrogen should act as a shallow acceptor.

4. Hydrogen storage materials

The methodology outlined above can be productively applied to materials for hydrogen storage. Indeed, a fundamental understanding of the hydrogen uptake and release processes is still lacking. First-principles calculations have already made important contributions to the understanding of hydrogen storage materials [20–26], but to date they have focused almost exclusively on *bulk* properties. Although such calculations have produced valuable insights relating to quantities such as heats of formation, a detailed treatment covering the behavior of hydrogen atoms as *impurities* in the host material is essential for developing a fuller understanding of the processes of hydrogen uptake and release that are key to the performance of storage materials.

Starting from the fully hydrogenated material, the first step of a hydrogen release process consists of a hydrogen atom leaving its substitutional lattice site. This creates a hydrogen interstitial, which can be studied with the methodology outlined in Section 2. In addition, a vacancy is now present at the site where the hydrogen used to sit, and such vacancies can be studied with similar first-principles techniques. If the host material is an insulator (i.e., if it has a finite band gap), various possible charge states need to be considered for both the interstitial and the vacancy. The formation energy of the interstitial and the vacancy determines the likelihood of forming the defect. If this energy is high, then the host material will be very stable and decomposition will

be very slow; if the formation energy is low, then it is likely that dehydrogenation will occur. As illustrated in Fig. 2, the formation energy of a charged defect depends sensitively on the Fermi level, so consideration of the likely position of the Fermi level is important.

In addition to studying the formation energy, first-principles calculations can also be used for investigation the diffusion of the hydrogen interstitial (or the vacancy) through the lattice. The energy barriers along the path determine how fast the hydrogen can move, and hence play a decisive role in reaction kinetics. Migration paths of hydrogen interstitials have been extensively studied in our work on semiconductors [27].

We have recently performed such investigations of hydrogen defects and their migration in the case of complex hydrides, specifically for sodium alanate. The fact that hydrogen itself is a constituent of this host material does not really matter in terms of the methodology: it is straightforward to distinguish between hydrogen host atoms residing in regular lattice sites and hydrogen "impurities" which are additional interstitial hydrogen atoms. A report on the application of this formalism to hydrogen in sodium alanate can be found in the paper by Peles and Van de Walle elsewhere in this volume [28].

Metal hydrides form another class of materials that have been widely investigated for hydrogen storage. Most metal hydrides studied to date have difficulty meeting the weight-percent criterion that applies to hydrogen storage for mobile applications; however, for other applications different criteria (such as volume or kinetics of hydrogen uptake and release) may be more important. The formalism for studying interactions of hydrogen impurities with metals of course differs in some ways from the formalism outlined above for insulators. The definition of formation energy still applies, but in a metal the charge of the impurity has to be neutral, and the Fermi level does not vary. In spite of that, we think it will be productive to investigate hydrogen in metals along the same lines as outlined above; specifically, an analysis of the local bonding character may also indicate whether the hydrogen impurity is more "donor-like" or "acceptor-like" in a specific configuration within the metal. If this turns out to be true, then it may be possible to extend the "universal alignment" model (Fig. 3) to metals as well.

5. Summary

We have discussed the behavior of hydrogen in semiconductors and insulators in general, and shown how first-principles calculations can provide qualitative and quantitative information about this behavior. A "universal alignment" model allows predicting the electronic behavior of hydrogen in a wide range of materials. Specific examples and predictions were discussed. The approach described here to treat hydrogen as an impurity

in a semiconductor or insulator can be productively applied to hydrogen storage materials as well. Bulk quantities such as formation enthalpies are of course also important and informative, but ultimately the crucial aspects of the functionality as a storage material revolve around the removal of hydrogen atoms (or their addition, in the case of regeneration or charging). This requires describing the behavior of hydrogen atoms as impurities inside the host material, and such calculations are now in progress.

Acknowledgements

References

[1] C.G. Van de Walle, B. Tuttle, IEEE Trans. Electron. Dev. 47 (2000) 1779.
[2] C.G. Van de Walle, Physica B 170 (1991) 21.
[3] S. Limpijumnong, J.E. Northrup, C.G. Van de Walle, Phys. Rev. B 68 (2003) 075206.
[4] C.G. Van de Walle, Y. Bar-Yam, S.T. Pantelides, Phys. Rev. Lett. 60 (1988) 2761.
[5] C. Herring, N.M. Johnson, C.G. Van de Walle, Phys. Rev. B 64 (2001) 125209.
[6] S. Limpijumnong, C.G. Van de Walle, Phys. Stat. Sol. 228 (2001) 303.
[7] C.G. Van de Walle, Phys. Rev. Lett. 85 (2000) 1012.
[8] L.E. Halliburton, L. Wang, L. Bai, N.Y. Garces, N.C. Giles, M.J. Callahan, B. Wang, J. Appl. Phys. 96 (2004) 7168.
[9] C.G. Van de Walle, Physica B 308–310 (2001) 899.
[10] C.G. Van de Walle, Phys. Stat. Sol. (b) 229 (2002) 221.
[11] P. Hohenberg, W. Kohn, Phys. Rev. 136 (1964) B864.
[12] W. Kohn, L.J. Sham, Phys. Rev. 140 (1965) A1133.
[13] C.G. Van de Walle, J. Neugebauer, J. Appl. Phys. 95 (2004) 3851.
[14] C.G. Van de Walle, J. Neugebauer, Annu. Rev. Mater. Res. 36 (2006) 179.
[15] C.G. Van de Walle, R.M. Martin, Phys. Rev. B 35 (1987) 8154.
[16] C.G. Van de Walle, Phys. Rev. B 39 (1989) 1871.
[17] C.G. Van de Walle, J. Neugebauer, Appl. Phys. Lett. 70 (1997) 2577.
[18] C.G. Van de Walle, J. Neugebauer, Nature 423 (2003) 626.
[19] E.A. Davis, S.F.J. Cox, R.L. Lichti, C.G. Van de Walle, Appl. Phys. Lett. 82 (2003) 592.
[20] M.E. Arroyo y de Dompablo, G. Ceder, J. Alloys Compd. 364 (2004) 6.
[21] O.M. Løvvik, O. Swang, Europhys. Lett. 67 (2004) 607.
[22] O.M. Løvvik, J. Alloys Compd. 373 (2004) 28.
[23] J. Íñiguez, T. Yildirim, T.J. Udovic, M. Sulic, C.M. Jensen, Phys. Rev. B 70 (2004) 060101.
[24] A. Peles, J.A. Alford, Z. Ma, L. Yang, M.Y. Chou, Phys. Rev. B 70 (2004) 165105.
[25] H. Smithson, C.A. Marianetti, D. Morgan, A. Van der Ven, A. Predith, G. Ceder, Phys. Rev. B 66 (2002) 144107.
[26] K. Miwa, A. Fukumoto, Phys. Rev. B 65 (2002) 155114.
[27] S. Limpijumnong, C.G. Van de Walle, Phys. Rev. B 68 (2003) 235203.
[28] A. Peles, C.G. Van de Walle, J. Alloys Compd. 446–447 (2007) 459.