First-principles calculations of thermal electron emission from H⁻ in silicon

Yoshiyuki Yamamoto, Alpin Novianus Tatan, and Osamu Sugino
The Institute for Solid State Physics, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8581, Japan
(Dated: May 29, 2020)

Thermal electron emission process of a hydrogen impurity is an important topic of fundamental semiconductor physics. Despite of decades-long study, theory is not established yet. Here, we study the process of H⁻ in silicon, H⁻ → H⁰ + e⁻, using a first-principles calculation. Our calculation indicates that the process consists of two steps: slow diffusion of H⁻ from a tetrahedral site to a bond-center site, which is the rate-limiting step, and faster nonradiative transition from H⁻ to H⁰ + e⁻ that occurs subsequently at the body-center site. The calculated rate is consistent with a deep level transient spectroscopy experiment.

I. INTRODUCTION

Due to its amphoteric nature and fast diffusivity, hydrogen impurities passivate various defects and help increase the efficiency of various semiconductor devices. Their nonradiative carrier capture/emission capability significantly affects diffusivity and carrier lifetimes. Due to its technological importance, this phenomenon has been studied over the past decades.

In silicon, hydrogen is known to exist as H⁺ or H⁻ depending on doping conditions, while H⁰ is metastable and can exist only at very low temperatures or under illumination. These properties can be consistently explained by first-principles total energy calculations, which indicate that H⁺ or H⁻ is thermodynamically stable while H⁰ is always unstable against the reaction 2H⁰ → H⁺ + H⁻. The transition between different charge states was experimentally studied with a deep level transient spectroscopy (DLTS). The DLTS experiment showed that thermal electron emission from H⁻, i.e., H⁻ → H⁰ + e⁻, has a slow reaction rate, r = 2.8 × 10⁻¹ s⁻¹ at room temperature with activation energy E_a = 0.84 eV. The activation energy is usually thought to correspond to the energy difference between H⁻ and H⁰ + e⁻. However, density functional theory (DFT) calculation of this quantity only amounts to 0.50 eV. The nature of the observed activation energy in the experiment therefore remains to be investigated.

Here, we show our theoretical study on the rate for thermal electron emission by H⁻ in silicon. Although there are many studies treating hydrogen impurity in silicon, as far as we know, there has been no first-principles study on the rate because it requires accurate description of the band gap, the defect thermodynamics, and electron-phonon couplings. Recently, however, Kim et al. demonstrated that first-principles calculation can reproduce the rate for nonradiative carrier capture from the DX center in GaAs despite large anharmonicity of the potential energy surface (PES). Because the PES for hydrogen impurity in silicon is also known to be anharmonic, their approach should help us to elucidate the reason for the high activation energy observed in the DLTS experiment on H⁻.

In this context, we calculated thermodynamic transition level with screened-hybrid DFT and confirmed that the calculated result is different from the experimental activation energy. Then, we constructed a configuration coordinate diagram for the transition between H⁻ and H⁰ + e⁻. The calculated diagram for H⁻ was found to have local minimum not only at its most stable tetrahedral site but also at a bond center site, which is the global minimum for H⁰ + e⁻. The existence of local minimum at the bond-center site is a key to explain the fast thermal electron emission because nuclear wavefunctions overlap considerably at the bond-center site. The calculated rate shows that, with the assumption of thermal equilibrium, thermal electron emission is much faster than the experiment, indicating another step should rate-determine the whole process. Finally, we calculated the migration barrier of H⁻ between tetrahedral and bond-center site, and conclude that the migration is the rate-limiting step for thermal electron emission from H⁻.

II. METHODS

We performed first-principles calculations based on DFT with Vienna Ab initio Simulation Package (VASP), where the Kohn-Sham orbitals are expressed by the plane-wave basis. We used the screened hybrid functional formulated as Heyd-Scuseria-Ernzerhof exchange correlation functional (HSE06 functional). Projector augmented wave (PAW) method was used for the interactions between valence electrons and ions. The cutoff energy for the plane-wave basis was 400 eV, which was found to give sufficient accuracy for defect thermodynamics and electron-phonon coupling in our system. We optimized the lattice constant of silicon with conventional cell calculations using Γ-centered 6 × 6 × 6 k-point mesh for Brillouin-zone integration. The calculated lattice constant is 5.433 Å, in good agreement with experimental one extrapolated to 0 K, 5.430 Å.

For defect calculations, we use 3 × 3 × 3 super cell of conventional cell, including 216 Si atoms, in order to suppress spurious interactions between defects. To reduce the computational cost, Γ-point calculation was performed. The calculated band gap of silicon is 1.31 eV.
which is larger than the experimental one, 1.17 eV, because the conduction band bottom is not correctly sampled with Γ-point in 3 × 3 × 3 supercell. To obtain correct energetics between H\(^-\) and H\(^0\) + e\(^-\), we adjust the Kohn-Sham energies of the conduction bands by shifting them by −0.14 eV in the calculated configuration coordinate diagram. Finite-size correction for charged defect is calculated with the Markov-Payne correction\(^{33}\) by extrapolating the calculated energies with 2 × 2 × 3, 2 × 3 × 3 and 3 × 3 × 3 supercells to infinite cell size. For the extrapolation, we used 2 × 2 × 2 k-point for all supercells: this k-point mesh is found to be necessary for the extrapolation to work well. The correction was calculated for H\(^-\) at a tetrahedral site, and the same correction was used for all other configurations. The finite-size correction changes the calculated thermodynamic transition level, ε(0/−), by +0.06 eV.

The rate for thermal electron emission was calculated with the formalism proposed by Alkauskas et al.\(^{34}\). In the formalism, with the aid of the static approximation for electronic states and one-dimensional approximation for nuclear degrees of freedom, the rate for nonadiabatic transition from electronic state \(i\) to \(j\) with multiphonon absorption is calculated with the Fermi’s golden rule:

\[
r_{ij} = \frac{2\pi}{\hbar} |W_{ij}|^2 \sum_n \omega_n \sum_m |\langle \chi_{jm} | Q - Q_0 | \chi_{in} \rangle|^2 \delta(E_{in} - E_{jm}).
\]

(1)

\(W_{ij}\) is the electron matrix element defined as \(W_{ij} = \langle \psi_j | \partial \psi_i / \partial Q \rangle\), where \(\langle \psi_{i(j)} \rangle\) is the Kohn-Sham orbital for state \(i(j)\). With first-order perturbation theory, \(W_{ij}\) can be expressed as \(W_{ij} = (\varepsilon_j - \varepsilon_i) \langle \psi_j | \partial \psi_i / \partial Q \rangle\), where \(\varepsilon_i\) is the Kohn-Sham energy for a state \(i\). The Kohn-Sham orbitals overlaps, \(\langle \psi_j | \partial \psi_i / \partial Q \rangle\), is calculated with all-electron wavefunctions including PAW augmented core contributions reproduced by PAWPyseed code\(^{35}\). \(\omega_n\) is the thermal weight for the \(n\)-th eigenstate of nuclear wavefunction, \(\langle \chi_{in} \rangle\). The nuclear wave functions and eigenergies are calculated by solving the one-dimensional Schrödinger equation for configuration coordinate diagram to account for large anharmonicity of the potential energy surface\(^{36}\). \(Q\) is the configuration coordinate defined as \(Q = \sqrt{MR}\), where \(M\) is a diagonal matrix with masses in its diagonal element and \(R\) is the 3\(N\)-dimensional coordinate for lattice. We assume a linear reaction pathway between the most stable configurations of two charge states, H\(^0\) and H\(^-\), following the scheme proposed by Alkauskas et al.\(^{34}\) although this treatment is known to underestimate the rate\(^{36}\) because the linear pathway does not necessarily give a major contribution. The electron emission rate from H\(^-\) is calculated by summing up the contributions of the transition from a defect state to conduction band states. The electron emission rate is thus calculated as \(r_i = \sum_{j=CB} r_{ij}\), where \(i\) is the defect state, and \(j\) is the conduction band state. As the summation for the conduction band states, we take the bands within 0.15 eV above the conduction band minimum. We have confirmed that conduction bands above them give only negligible effects to the calculated rate.

The delta function in Eq. (1) is approximated as Gaussian with finite width, \(\sigma\), which corresponds to a lifetime of the vibrational mode considered in the configuration coordinate. Although lifetimes of local vibrational modes at low temperatures are known\(^{37}\), we cannot use those values as \(\sigma\) because the local vibrational modes are different from our reaction pathway. Thus, we tried several values of \(\sigma\) to investigate its effect on the calculated reaction rate. The value of \(\sigma\) is, indeed, found to affect the calculated rate. However, our conclusion is robust against the choice of \(\sigma\) value, and we used \(\sigma = 0.026\) eV throughout this study.

III. RESULTS

A. Thermodynamics

Firstly, we investigate the thermodynamics of hydrogen in silicon. The calculated energies suggest that the most stable site is the bond-center (BC) site for H\(^0\) while the tetrahedral (Td) site for H\(^-\), and the calculated thermodynamic transition level is ε(0/−) = 0.67 eV from the valence band maximum. These results are in good agreement with previous theoretical study using local density functional\(^{38}\). For the electron emission process, the energy difference between H\(^0\) + e\(^-\) and H\(^-\) is calculated to be \(E_g - \varepsilon(0/−)\) = 0.50 eV, where \(E_g\) is the band gap of silicon. Thus, the activation energy for thermal electron emission is expected to be 0.50 eV from first-principles calculations. This value is, however, inconsistent with an experimental activation energy, 0.84 eV, obtained by a DLTS experiment\(^{37}\). Considering the inconsistency, we investigated possible origins of the discrepancy from computational point of views. We calculated the effects of thermal expansion of silicon at room temperature, zero-point energy of hydrogen and the type of pseudopotentials (hard/standard). These effects change the calculated \(\varepsilon(0/−)\) as −0.01 eV, +0.06 eV, and −0.02 eV, respectively. None of these effects is thus confirmed to fill the gap between theoretical and experimental values. In conclusion, our calculation indicates that thermodynamics could not explain the activation energy observed in the DLTS experiment.

B. Configuration Coordinate Diagrams

Figure II shows the configuration coordinate diagram for the transition between H\(^-\) and H\(^0\) + e\(^-\). The configuration coordinate is taken as the linear interpolation between their most stable Td site for H\(^-\) and BC site for H\(^0\). The diagram for H\(^0\) + e\(^-\) is calculated as the summation of total energy of H\(^0\) and the Kohn-Sham energy of the conduction band minimum of pristine silicon. The calculated configuration coordinate diagrams have...
two unique features: firstly, the $H^0$ state shows large anharmonicity around the configurations of the Td site. At the configuration of the Td site, the hydrogen occupies the interstitial void of silicon. Because the radius of $H^0$ is smaller than that of $H^-$, the $H^0$ interacts less strongly with Si. Thus, the flat PES appears around the Td site for $H^0$ while the PES of $H^-$ has a deep minimum at the Td site. The calculated energy difference of $H^0 + e^-$ and $H^- \cdot $ is much smaller than that at the Td site. Therefore, we are likely to be dominated at the BC site where the energy required to activate the emission is much smaller than that at the Td site. Therefore, we can use the Kohn-Sham orbitals around BC site to calculate $W_{ij}$ in Eq. (1) to discuss the electron emission at room temperature. It should be noted that we neglect relative efficiency of the Td-to-BC migration of $H^-$ and the thermal electron emission at the BC site. Hence the migration is assumed to occur quickly toward thermal equilibrium so that the thermal weight, $\omega_n$, is allowed to be used in Eq. (1). Later, we will discuss the validity of this assumption based on the calculated rate from first principles.

C. Nonradiative Electron Emission

The electronic matrix element, $W_{ij}$, is calculated with the finite-displacement method at the configuration of the BC site as discussed in the previous section. In Table 1, we tabulated the values of $W_{ij}$ which give major contributions to the calculated rates. As shown in the table, not only the CBM but also other bands within 0.15 eV from CBM have large values of $W_{ij}$, indicating the im-

Figure 1. (Color online) Configuration coordinate diagrams of $H^0 + e^-$ and $H^- \cdot $ For $H^0 + e^-$, electron is assumed to occupy the conduction band bottom. The dot symbols show the calculated energies from first principles, and the lines show the ones calculated with the spline interpolation between the calculated ones. $Q = 0$ amu$^{1/2}$ Å is the most stable configuration for $H^0$ (BC site) while the configuration for $Q = 4.49$ amu$^{1/2}$ Å is the most stable one for $H^- \cdot $ (Td site).
portance to include those contributions into the calculations. Only eleven conduction bands are included in the calculations because the other conduction bands are well-separated by 0.5 eV and are thus negligible. The nuclear wavefunctions in Eq. (1) are obtained by solving the one-body Schrödinger equation with PES obtained with spline interpolation of the calculated PES. We note that the configuration coordinate diagrams around the BC site is in the Marcus inverted region\textsuperscript{22}, because positions of the local minima are virtually the same within our configuration coordinate, the overlaps of nuclear wavefunctions are expected to be large compared to that in the normal region, thus possibly giving large rate for thermal electron emission.

Figure 3 shows the calculated rate for the thermal electron emission from H\textsuperscript{-} to conduction bands. The calculated rate at room temperature is $5.2 \times 10^7 \text{s}^{-1}$ and the activation energy is estimated to be 0.46 eV from the Arrhenius plot. The calculated activation energy is consistent with the results expected with thermodynamics calculations, $E_g - \varepsilon(0/-) = 0.50 \text{ eV}$. Although the calculated rates depend on the smearing factor $\sigma$ to describe the $\delta$ function, the calculated rates are much larger than the experimental value, $r = 2.8 \times 10^{-1} \text{s}^{-1}$ at room temperature\textsuperscript{8}.

The argument so far assumes thermal equilibrium: i.e. the diffusion of H\textsuperscript{-} from the Td site to the BC site is assumed to be much faster than the thermal electron emission at BC site, thus enabling the use of thermal weight, $\omega_n$ in Eq. (1) as a population of initial state. To test this assumption, we calculate the activation energy for the migration of H\textsuperscript{-} with the climbing image nudged elastic band method\textsuperscript{24,25}. Figure 4 shows the calculated PES for H\textsuperscript{-} migration between Td site and BC site. Here, the configuration of the BC site is optimized for H\textsuperscript{-}, thus slightly different from the one used in the configuration coordinate diagram in Figure 1 which is the configuration optimized for H\textsuperscript{0}. The calculated activation energy is 0.96 eV, which is comparable to the one from the DLTS experiment, 0.84 eV. The calculated activation energy is used to estimate the rate for the migration with a simple formula, \( r = \frac{kT}{h} \exp(-E_a/kT) \). Then, the calculated rate at $T = 300 \text{ K}$ is $4.3 \times 10^{-3} \text{s}^{-1}$, which is much smaller than the calculated rate for thermal electron emission under thermal equilibrium. Thus, the rate-determining step is suggested to be the migration of H\textsuperscript{-} from Td site to BC site. If the experimental activation energy, $E_a = 0.84 \text{ eV}$, is used, the calculated migration rate is $3.0 \times 10^{-1} \text{s}^{-1}$, showing good agreement with the experimental rate, $2.8 \times 10^{-1} \text{s}^{-1}$. This fact strongly suggests that the activation energy observed in the DLTS experiment corresponds to the one for the migration process of H\textsuperscript{-}, and the migration is the rate-limiting step for the thermal electron emission from H\textsuperscript{-}. It should be noted here that the rate for the migration is estimated with treating hydrogen as a classical particle, thus neglecting nuclear quantum effect such as quantum tunneling. Although the quantum tunneling can be dominant as migration mechanism at room temperature, we expect that its effect is small because the distance between the Td site and the BC site is large as shown in Fig. 1.

| $j$ | $\varepsilon_j - \varepsilon_{\text{CBM}}$ (eV) | $W_j$(eV/amu$^{3/2}$ Å) |
|-----|---------------------------------|-------------------------|
| CBM | 0.000                           | 2.34 × 10$^{-2}$        |
| CBM+1 | 0.042                          | 6.47 × 10$^{-2}$        |
| CBM+5 | 0.047                          | 7.66 × 10$^{-2}$        |
| CBM+9 | 0.102                          | 9.80 × 10$^{-2}$        |
| CBM+10 | 0.102                         | 2.39 × 10$^{-2}$        |
IV. CONCLUSION

Here, we study the thermal electron emission process of \( H^- \) in silicon, \( H^- \rightarrow H^0 + e^- \), with density functional calculations. Our calculation indicates that the electron emission process from \( H^- \) in Si consists of two steps. The first step is \( H^- \) migration from its most stable Td site to a metastable BC site. This migration is the rate-limiting step for the whole reaction with calculated activation energy of 0.96 eV. The second step is the electron emission from \( H^- \) in the BC site to a conduction band. This step is much faster than the first step, having the calculated activation energy of 0.46 eV. The activation energy observed in a DLTS experiment, 0.84 eV, is likely to correspond to the one in the first step. Our study reveals that modern first-principles calculation can elucidate nonadiabatic process of hydrogen impurity, which is the most ubiquitous and mysterious impurity in semiconductors, and thus shows the possibility to clarify the degradation mechanism of semiconductor devices by hydrogen impurity from first principles.

ACKNOWLEDGMENTS

The calculations were performed on the supercomputers at the Institute for Solid State Physics, the University of Tokyo. This research was supported by MEXT as Priority Issue on Post-K computer (Development of new fundamental technologies for high-efficiency energy creation, conversion/storage and use) and by JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas Hydrogenomics, No. JP18H05519. We receive further support from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

* Current affiliation: Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan; E-mail: YAMAMOTO.Yoshiyuki@nims.go.jp

1. A. Alkauskas, M. D. McCluskey, and C. G. Van de Walle, Tutorial: Defects in semiconductors Combining experiment and theory, J. Appl. Phys. 119, 181101 (2016).
2. A. J. Tavendale, D. Alexiev, and A. A. Williams, Field drift of the hydrogenrelated, acceptorneutralizing defect in diodes from hydrogenated silicon, Appl. Phys. Lett. 47, 316 (1985).
3. N. M. Johnson, C. Herring, and D. J. Chadi, Interstitial hydrogen and neutralization of shallow-donor impurities in single-crystal silicon, Phys. Rev. Lett. 56, 769 (1986).
4. N. M. Johnson, C. Herring, and C. G. Van de Walle, Inverted order of acceptor and donor levels of monatomic hydrogen in silicon, Phys. Rev. Lett. 73, 130 (1994).
5. K. Bonde Nielsen, B. B. Nielsen, J. Hansen, E. Andersen, and J. U. Andersen, Bond-centered hydrogen in silicon studied by in-situ deep-level transient spectroscopy, Phys. Rev. B 60, 1716 (1999).
6. Y. V. Gorelikinskii and N. N. Nevinnyi, Electron paramagnetic resonance of hydrogen in silicon, Physica B: Condensed Matter 13, 13 (1991).
7. C. G. Van de Walle, P. J. H. Denteneer, Y. Bar-Yam, and S. T. Pantelides, Theory of hydrogen diffusion and reactions in crystalline silicon, Phys. Rev. B 39, 10791 (1989).
8. C. Herring, N. M. Johnson, and C. G. Van de Walle, Energy levels of isolated interstitial hydrogen in silicon, Phys. Rev. B 64, 125209 (2001).
9. S. Kim, S. N. Hood, and A. Walsh, Anharmonic lattice relaxation during nonradiative carrier capture, Phys. Rev. B 100, 041202 (2019).
10. P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, Phys Rev 136, B864 (1964).
11. W. Kohn and L. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, Phys Rev 140, A1133 (1965).
12. Kresse and Furthmuller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set., Phys. Rev., B Condens. Matter 54, 11169 (1996).
13. G. Kresse and J. Furthmuller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. [Comp Mater Sci 6, 15 (1996)]
14. G. Kresse, Ab initio molecular dynamics for liquid metals, J Non-cryst Solids 192, 222 (1995).
15. Kresse and Hafner, Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium., Phys. Rev., B Condens. Matter 49, 14251 (1994).
16. A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, Influence of the exchange screening parameter on the performance of screened hybrid functionals, The Journal of Chemical Physics 125, 224106 (2006).
17. V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, Tests of a ladder of density functionals for bulk solids and surfaces, Phys. Rev. B 69, 075102 (2004).
18. G. Makov and M. C. Payne, Periodic boundary conditions in ab initio calculations, Phys. Rev. B 51, 4014 (1995).
19. A. Alkauskas, Q. Yan, and C. G. Van de Walle, First-principles theory of nonradiative carrier capture via multiphonon emission, Physical Review B 90, 10.1103/PhysRevB.90.075202 (2014).
20. K. Byström, D. Broberg, S. Dwaraknath, K. A. Persson, and M. Asta, Pawpyseed: Perturbation-extrapolation band shifting corrections for point defect calculations, arXiv:1904.11572 [cond-mat] (2019) arXiv: 1904.11572.
21. L. Shi, K. Xu, and L.-W. Wang, Comparative study of ab initio nonradiative recombination rate calculations under different formalisms, Physical Review B 91, 10.1103/PhysRevB.91.205315 (2015).
22. M. Budde, G. Lpke, C. Parks Cheney, N. H. Tolk, and L. C. Feldman, Vibrational Lifetime of Bond-Center Hydrogen in Crystalline Silicon, Phys. Rev. Lett. 85, 1452 (2000).
23. R. A. Marcus, Electron Transfer Reactions in Chemistry: Theory and Experiment (Nobel Lecture), Angew. Chem. Int. Ed. Engl. 32, 1111 (1993).
G. Henkelman and H. Jnsson, Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points, The Journal of Chemical Physics 113, 9978 (2000).

G. Henkelman, B. P. Uberuaga, and H. Jnsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths, The Journal of Chemical Physics 113, 9901 (2000).