2009 Master Thesis

Theoretical Study on Ballistic Transport Characteristics of Silicon Nanowire FETs

Supervisor

Professor Hiroshi Iwai

Department of Electronics and Applied Physics

Interdisciplinary Graduate School of Science and Engineering

Tokyo Institute of Technology

08M36474

Yeonghun Lee

Abstract

Silicon nanowire field-effect transistors (SiNW FETs) have been focused as one of the new FET structures to obtain good gate controllability owing to the immunity against short channel effects. In this thesis, we have investigated size dependent potential performances of SiNW FETs based on a combination of first-principle calculations and a ballistic transport model from Landauer's formula.

Size dependent band structures of SiNWs aligned to [100] direction, ranging from 0.77 to 2.69 nm in thickness, has been investigated by the first-principles calculation. Our first-principles calculation is based on the density functional theory (DFT) with the local density approximation (LDA). Based on the obtained band structures, we have extracted effective masses and eigenvalues of subbands in the SiNWs with various sizes, which strongly affect charge density and carrier velocity. In addition, the four unprimed minima have been split while they are 4-fold degenerate in sufficiently large wire, which is not involved in the effective mass approximation. As the size increases, bandgap has approached close to that of the bulk silicon as we expected. The electron effective masses have been increased as shrinking size in both unprimed and primed subbands. With the increase in the size, the effective masses close to 0.19 in the unprimed subband and 0.916 in the primed subband, which correspond transverse and longitudinal electron effective masses of the bulk silicon, respectively.

Combined with Landauer's formula, electrical characteristics of ballistic SiNW FET have been estimated and an assessment of size dependent performances has been conducted. By using the calculated subband structure, size-dependent charge density and saturation injection velocity of each subband have been calculated by the self-consistent calculation of Schrödinger and Poisson equations, which is determined

by the balance between density of states and gate capacitance, both of which increase as size increases. As a result, large SiNW FETs have showed large on-current owing to steadily increasing gate capacitance by longer periphery. Energy gap between the lowest unprimed subband minimum and the source Fermi level has been decreased as shrinking size because decreasing the gate capacitance is more drastic than decreasing the density of states. It would also be supposed that the highest injection velocity is obtained at a certain width larger than 2.69 nm. Therefore, an assessment of ballistic drain current for practical multi-channel SiNW FET would show a trade-off between saturation injection velocity and the number of wires per unit width, determined by geometrical parameters. In conclusion, we have revealed that size-dependent band structures have substantially affected modulation of ballistic transport characteristics, and those effects of each subband have been changed as size modulates.

Contents

4. Conclusions

Acknowledgements

Reference

1. Introduction

1.1 Request of nanowire FETs

The scaling issues in planar metal-oxide field effect transistor (MOSFET) requires a better electrostatic control of the channel to reduce the off-state leakage current. Three dimensional (3D) MOSFETs, including Fin FETs and nanowire FETs, have been extensively studied to surpass the bulk or silicon-on-insulator (SOI) FETs.[1] Generally, suppression of off-current enables to lower the threshold voltage so that large on-current can be obtained. However, one of the concerns for 3D MOSFET is the reduction in net-current as the cross section of the channel becomes small. One way to overcome this problem is to achieve a ballistic transport within the channel, so that large on-current can be obtained without degrading the electrostatic control of the channel. In order to achieve a ballistic transport, a short channel length with reduced scattering event is necessary, and it has already been reported that nanowire FETs have excellent short channel effect immunity.[2] Recently, silicon nanowires (SiNWs) have been focused as one of the new FET structures to obtain a large on-current with high on-current/off-current ratio.[3] Therefore, SiNW FETs can be regarded as one of the extremely scaled 3D MOSFET for future large scale integrated (LSI) devices instead of the planar FET (Fig. 1.1).
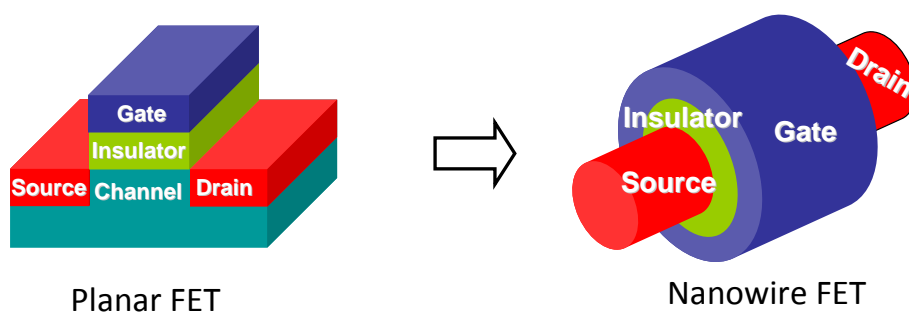

Fig. 1.1 Schematic structures of planar FET and nanowire FET.

6

1.2 Purpose of this work

As SiNWs have different physical properties from the bilk silicon, their band structures have been calculated by first-principles calculation[4-7] or tight-binding method[8-15]. Their carrier transport have also been modeled by several methods.[8-19] In this thesis, we investigate the size-dependent potential performance of SiNW-FETs based on a combination of the first-principles calculations with a ballistic FET model from a paper of Natori.[18,20] First part of this thesis discusses the size-dependent band structures and electronic properties are addressed. Second part of this thesis assesses the on-current of the SiNW-FETs under ballistic transport based on the obtained band structures of SiNWs. In addition, linear charge density and saturation velocity of each subband are discussed to support the assessment.

2. Band Structures of Silicon Nanowires by First-principles Calculation

2.1 First-principles calculation

The band structures of SiNWs were calculated by first-principles calculation based on density functional theory (DFT) with local density approximation (LDA) using pseudo-potential.[21-25] The pseudo-potential is used for saving calculation cost. All the band calculations are performed with Tokyo Ab-initio Program Package (TAPP).[26] DFT, LDA, pseudo-potential, el al. are explained in following subsections referring to ref 27 and 28.

2.1.1 Density functional theory (DFT)

Density functional theory (DFT) is one of methods to calculate total energy and wave functions of ground state. It is based by that ground state energy E_{GS} can be yielded with a function of one electron density $n(r)$. If it is satisfied, a problem to get wave function of many-body can be changed to the simple problem related to one electron density.

Minimum energy of one electron density becomes ground state energy,

$$E[n] = \int v_{ext}(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r} + F[n] \geq E_{GS} \quad \text{(theorem 1)} \qquad (2.1)$$

is proven by Hogenberg, Kohn, Levy, et al.[21,22] Where $E[n]$ denotes total energy, which is a function of one electron density, and $v_{ext}(\boldsymbol{r})$ denotes external potential of position $\boldsymbol{r}$ described as

$$v_{ext}(r) \equiv \sum_{\boldsymbol{R}_I} v_I(\boldsymbol{r} - \boldsymbol{R}_I),$$

where v_I and $\boldsymbol{R}_I$ denote potential from nuclei I and nuclei position, respectively. Because wave function corresponding to one electron density is not unique, the $E[n]$ is

impossible to be directly derived by arbitrary n. Ground state energy E_{GS}, however, can be obtained from one electron density n. Here, $F[n]$ described as

$$F[n] = \left\langle \psi^n_{\min} \left| \hat{T} + \hat{V}_{ee} \right| \psi^n_{\min} \right\rangle$$

is installed into eq. (2.1), where the $F[n]$ denotes expectation value of electron kinetic energy $\hat{T}$ + interaction between electron $\hat{V}_{ee}$ which has a wave function minimizing that expectation $\psi^n_{\min}$ among anitisymmetric wave functions with the $n(r)$. It is also proven that E_{GS} is same with $E[n_{GS}]$, where n_{GS} denotes one electron density of ground state (theorem 2). A problem to get wave function of many-body is eventually changed to the problem related to one electron density. So that, using a function of one electron density $n(r)$, ground state energy E_{GS} is obtained. Theorem 1 and 2 is proved in appendix A.

2.1.2 Kohn-Sham equations

To calculate the ground state energy on DFT, Kohm-Sham imports independent particle system which yields one electron density n, where particles are non-interacted.[23] Hereby, many-body problem can be changed to effective one electron problem described as

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + v(r) \right] \psi_i(r) = \varepsilon_i \psi_i(r), \tag{2.2}$$

$$n(r) = \sum_i^N \left| \psi_i(r) \right|^2, \tag{2.3}$$

where $\psi_i(r)$ called Kohn-Sham orbit denotes wave function of one particle, and i summation upto N is carried out by order of small ε_i where i has to involve spin degree of freedom. In this system, $F[n]$ can be devided by three components shown as

$$F[n] = T_s[n] + U[n] + E_{XC}[n],\qquad(2.4)$$

where the first term of eq. (2.4) indicates kinetic energy in virtual non-interacted system described as

$$T_s[n] = \sum_i^N \int \psi_i^*(\boldsymbol{r})\left(-\frac{\hbar^2}{2m}\nabla^2\right)\psi_i(\boldsymbol{r})d\boldsymbol{r},\qquad(2.5)$$

From (2.2), eq. (2.5) can be described as

$$T_s[n] = \sum_i^N \varepsilon_i - \int v(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r}.\qquad(2.6)$$

Second term of eq. (2.4) indicates potential of coulomb interaction between electrons described as

$$U[n] = \frac{e^2}{2}\iint d\boldsymbol{r}d\boldsymbol{r}'\frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}.\qquad(2.7)$$

Third term of (2.4) denotes exchange-correlation energy involving entire many-body effects. Substituting eqs. (2.6) and (2.7) in eq. (2.4), $E[n]$ is changed to

$$E[n] = \sum_i^N \varepsilon_i - \int v(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r} + \int v_{ext}(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r} + \frac{e^2}{2}\iint d\boldsymbol{r}d\boldsymbol{r}'\frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} + E_{XC}[n].\quad(2.8)$$

It is variation principle to determine the $v(\boldsymbol{r})$ involved in eq. (2.2). In other words, variation equation has to be yielded with one electron density of ground state. Firstly, from eq. (2.2),

$$\sum_i^N \delta\varepsilon_i = \int \delta v(\boldsymbol{r})n_{GS}(\boldsymbol{r})d\boldsymbol{r}\qquad(2.9)$$

is yielded. Using eq. (2.9), variation of eq. (2.8) is yielded described as

$$\delta E[n_{GS}] = \int \delta v(\boldsymbol{r})n_{GS}(\boldsymbol{r})d\boldsymbol{r} - \int \delta n_{GS}(\boldsymbol{r})v(\boldsymbol{r})d\boldsymbol{r} - \int \delta v(\boldsymbol{r})n_{GS}(\boldsymbol{r})d\boldsymbol{r}$$

$$+ \int \delta n_{GS}(\boldsymbol{r})\left(v_{ext}(\boldsymbol{r}) + e^2\int \frac{n_{GS}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r}' + \frac{\delta E_{XC}[n_{GS}]}{\delta n_{GS}(\boldsymbol{r})}\right)d\boldsymbol{r}.\qquad(2.10)$$

Here, based on condition of an electron number invariability,

$$\int \delta n_{GS}(\boldsymbol{r})d\boldsymbol{r} = 0 \,,$$

in order that $\delta E[n_{GS}]$ of eq. (2.10) becomes zero, $v(\boldsymbol{r})$ is yielded as

$$v(\boldsymbol{r}) = v_{ext}(\boldsymbol{r}) + e^2 \int \frac{n_{GS}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r}' + \frac{\delta E_{XC}[n_{GS}]}{\delta n_{GS}(\boldsymbol{r})} \,, \qquad (2.11)$$

where independence constants on $\boldsymbol{r}$ are neglected. In eq. (2.11),

$$\mu_{XC}(\boldsymbol{r}) = \frac{\delta E_{XC}[n_{GS}]}{\delta n_{GS}(\boldsymbol{r})} \qquad (2.12)$$

is called exchange-correlation potential. Equations (2.2), (2.3) and (2.11) are called Kohn-Sham equations and they have to be solved by self-consistence. Substituting arbitrary wave function into eq. (2.3), $n(\boldsymbol{r})$ is yielded. And substituting the yielded n for n_{GS} in eq. (2.11), $v(\boldsymbol{r})$ is obtained. In this turn, substituting the obtained $v(\boldsymbol{r})$ in eq. (2.2), new wave function is yielded. This flow iterates until when yielded $v(\boldsymbol{r})$ is same with $v(\boldsymbol{r})$ before one step. If we know the E_{XC}, we can exactly calculate one electron density and total energy of ground state.

2.1.3 Local Density Approximation (LDA)

Exchange-correlation potential $\mu_{XC}(\boldsymbol{r})$ in eq. (2.12) is necessary for solving Kohn-Sham equation. However, correctly calculating exchange-correlation energy $E_{XC}[n]$ is very difficult. Thus, $E_{XC}[n]$ is calculated approximately. In this section, local density approximation, which is one of the approximate methods, is briefly explained.

Electron density in matter has spatial variation. When the spatial variation of electron density is gradual, we can consider as homogeneous electron gas. $E_{XC}[n]$ with electron density $n(\boldsymbol{r})$ each point in space is described as

$$E_{XC}[n] = \int \varepsilon_{XC}(n(\boldsymbol{r}))n(\boldsymbol{r})d\boldsymbol{r} \,, \qquad (2.13)$$

where $\varepsilon_{XC}(n)$ denotes exchange-correlation energy of homogeneous electron gas. From eq. (2.13), $\mu_{XC}(r)$ in eq. (2.12) is transformed as

$$\mu_{XC}(r) = \varepsilon_{XC}(n(r)) + n(r)\frac{d\varepsilon_{XC}(n)}{dn}.$$

$\varepsilon_{XC}(n)$ can not be derived analytically, but it has numerically been derived by quantum Monte Carlo method.[24] The results are imported in LDA.[25]

2.1.4 Pseudo-potential

To solve the Kohn-Sham equation is same as to solve a 3D differential equation. Wave functions can be expended by basis functions $\{\chi_\mu(r)\}$ as shown by

$$\psi_i(r) = \sum_\mu \chi_\mu(r)C_{i\mu} . \tag{2.14}$$

Kohn-Sham equation is solved by plane wave basis functions $\{\exp(i(k+G_\mu)\cdot r)\}$ where G_μ denotes a reciprocal lattice vector. Then the plane wave basis functions are orthogonal set. In the case of one-dimensional lattice, the smallest G_μ is $2\pi/R$, and the other G_μ is a multiple of the smallest G_μ where R denotes period of lattice. A number of necessary plane wave basis functions affect to calculation cost, and the number of necessary plane wave basis are determined by a maximum G_μ. Because wave function of core electron (e.g. 1s of silicon atoms) is strongly localized, plane wave basis functions with sufficiently large wave number are required for an expression of the core electron. The large wave numbers make a number of necessary plane wave basis functions increase. They hardly affect solid sate properties because the core electron has much lower energy state then valence electron (e.g. 3s of silicon atoms) and is strongly bound to atoms. So that, we can neglect the core electrons and cut calculation cost because expression is done by only small wave numbers. Eliminating nodes of valence

wave function in the core area, the calculation cost can be also cut. These are concept of pseudo-potential.

In periodic corollary as crystal, wave function $\psi_i(r)$ of eq. (2.14) is expended by plane wave basis functions as

$$\psi_{kn}(r) = \sum_{\mu} \exp\left(i\left(k + G_{\mu}\right) \cdot r\right) C_{i\mu} \,, \tag{2.15}$$

where a suffix letter i is consist of wave number k and the other quantum numbers n: $i = (k,n)$. Here, orthogonalized plane wave (OPW) is induced where core electrons independent of concerning solid state properties are neglected. Using Bloch's condition and isolated atom's wave functions with orbits c, $b_c(r)$, a wave function in crystal, $b_{kc}(r)$, is described as

$$b_{kc}(r) = \sum_{R} e^{ik \cdot R} b_c(r - R) \,.$$

where R denotes lattice vectors and this express is called linear combination of atomic orbital (LCAO). In this case, there is an atom in each unit cell. Here, valence states have to be orthogonalizied with the core states. So that, the OPW denotes

$$\zeta_k(r) = e^{ik \cdot r} - \sum_{c} b_{kc}(r) \left\langle b_{kc}(r) \middle| e^{ik \cdot r} \right\rangle \,,$$

as shown in fig. 2.1. Then the wave function of (2.15) can be altered by

$$\psi_i(r) = \sum_{\mu} \zeta_{k+G_{\mu}} C_{i\mu} \,. \tag{2.16}$$

When the wave function of (2.16) is exactly derived, a new wave function ϕ_{kn} with the same coefficient $C_{i\mu}$ can be considered by

$$\phi_{kn}(r) = \sum_{\mu} \exp\left(i\left(k + G_{\mu}\right) \cdot r\right) C_{i\mu} \,.$$

Then, the Kohn-Sham equation

$$H\psi_{kn}(r) = \varepsilon_{kn} \psi_{kn}(r)$$

is transformed to

$$(H + V_{ad})\phi_{kn}(r) = \varepsilon_{kn}\phi_{kn}(r),$$

where a new potential including with an added potential V_{ad} is called pseudo-potential, and ϕ_{kn} is called pseudo wave function. However, the pseudo-potential does not assure exact wave functions on even valence states. Physics properties can not be derived exactly owing to incorrect wave function.
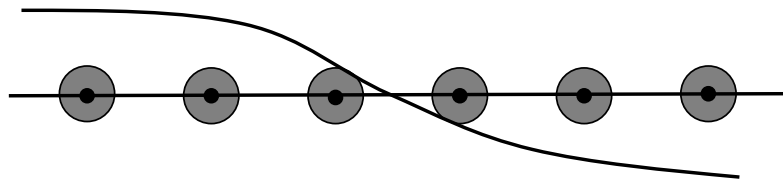

Norm conserving pseudo-potential

Norm conserving pseudo-potential is induced for derivation of true wave function of valence states. Here, a radius of core area is denoted by r_c. A pseudo-potential which yields that a wave function in $r > r_c$ is a real wave function and a wave function in $r < r_c$ is a pseudo wave function as shown in fig. 2.2 is assumed. A nodeless wave function as the pseudo wave function in fig. 2.2 has the smallest wave number and lowest energy state, so that pseudo-potential is more shallow and softer then real potential. Then wave function can be expended by fewer plane wave basis functions because a necessary maximum reciprocal lattice vector G_μ for expression by plane wave basis functions becomes small. When electron density in the core area derived from pseudo-potential is same as that derived by real potential, a potential of valence area is same as real potential of valence area. From above, the pseudo-potential has to satisfy following three conditions:
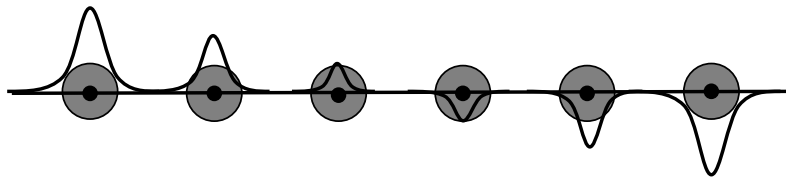
1. Pseudo wave functions does not have any node in $r < r_c$;

2. Pseudo-potentials $\phi_{ps}(r)$ are same as real potentials $\phi_t(r)$ in $r > r_c$;

3. Norm conserving is satisfied:

$$\int\limits_{r<r_c} dr |\phi_{ps}(r)|^2 = \int\limits_{r<r_c} dr |\phi_t(r)|^2.$$

(a) Plane wave function

(b) Core electron wave function

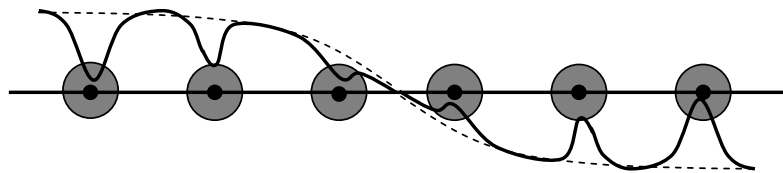(c) OPW = plane wave function - core electron wave function

Figure 2.1 Core electron wave function with Bloch's condition and OPW orthogonalized with that.
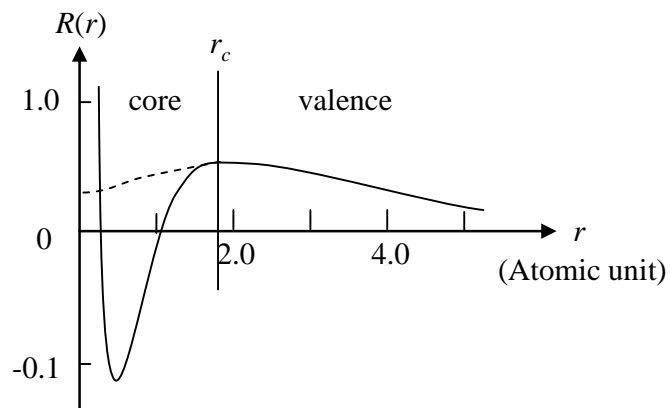
Figure 2.2 Concept of norm conserving pseudo-potential. Solid line indicates a true wave function of a radius vector for 3s state of Na. Dotted line indicates a pseudo wave function. Norm of pseudo wave function is same as that of true wave function in $r < r_c$.

2.1.5 Preparation for calculation

2.1.5.1 Silicon nanowire models

Calculated models of SiNWs were aligned to [100] direction ([100] SiNWs) with a square cross sectional shape. Figure 2.3 shows the cross sections of modeled SiNWs with thicknesses t_{SiNW} of 0.77 and 2.69 nm. The dangling bonds of the atoms at the periphery were passivated by hydrogen atoms.

2.1.5.2 Calculation parameters

As a periodic boundary condition, large supercells with neighboring wires separated by 0.7 nm are adopted. It has been confirmed that wires with 0.7-nm separation are enough to eliminate the interaction between the neighboring wires. Brillouin zone integration was performed by two k sampling points. The cutoff energy was set to 12.25 Ry, so as to reproduce the bandgap of bulk silicon to converge within 4 %.

The evolution of the bandgaps and the electron effective masses m^* were estimated using approximation of eigenvalues E at band edges based on the following equation.

$$E = \frac{\hbar^2 k^2}{2m^*},$$

where k and $\hbar$ are the wave vector and reduced Plank's constant, respectively. Wave vectors which have been used for the regression are within 4 pm^{-1} from the center wave vector with minimum energy and with a resolution of 2 pm^{-1}.
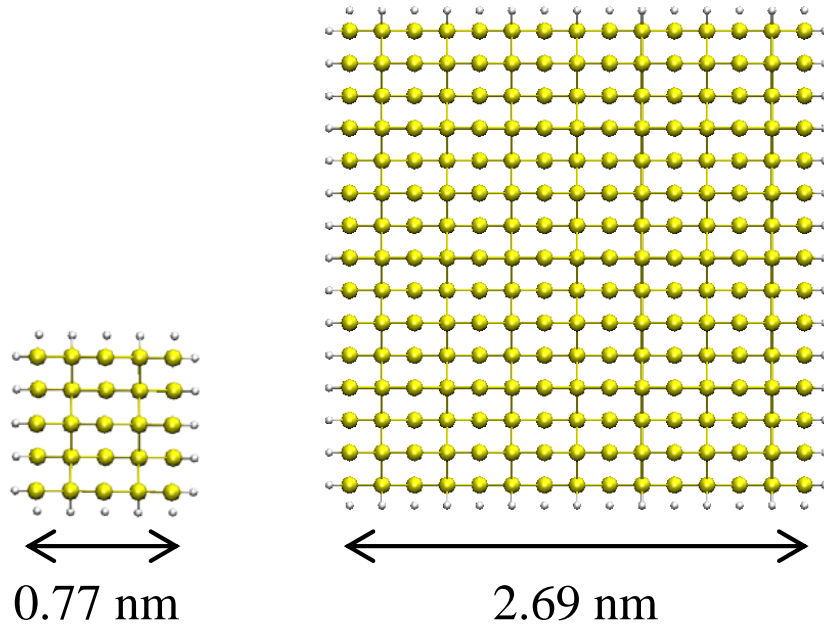
Fig. 2.3 Models of [100] SiNWs with a t_{SiNW} of 0.77 and 2.69 nm. Cross sections are square. Inside big circles and outside small circles represent silicon and hydrogen atoms, respectively.

2.2 Size-dependent band structures

The band structures of SiNWs were calculated by the first-principles calculation. In the next section, the band structures are also used to estimate transport parameters; linear charge density, carrier velocity, and current. Bandgaps, effective masses, and subband minima, which are purely extracted from the band structures, are shown in following subsection.

2.2.1 Band structure

Figure 2.4 shows the band structures of [100] SiNWs with w_{SiNW} ranging from 0.77 to 2.69 nm. Direct bandgaps at Γ point are obtained for both SiNWs, which is in contrast to the bandgap for bulk silicon. While increasing the w_{SiNW}, minima of four unprimed subbands at Γ point move toward lower energy due to relaxation of the quantum confinement, and eventually they will coincide with the minima of primed bands of the bulk silicon. Valley splitting of the four unprimed minima occurs while they are in 4-fold degeneracy in sufficiently large wire, which is in good agreement with previous reports.[4,9-11,14,15] A SiNW with w_{SiNW} of 0.77 nm has two 1-fold and one 2-fold degenerate unprimed band, whereas a SiNW with w_{SiNW} of 2.69 nm has 1-fold and 3-fold degenerate unprimed bands. One of the noticeable evolutions with w_{SiNW} is that a primed subband moves toward lower unprimed subband minima, which have strong impact on ballistic on-current, which will be discussed in section 3. These subband minima evolution is numerically shown in section 2.2.4.
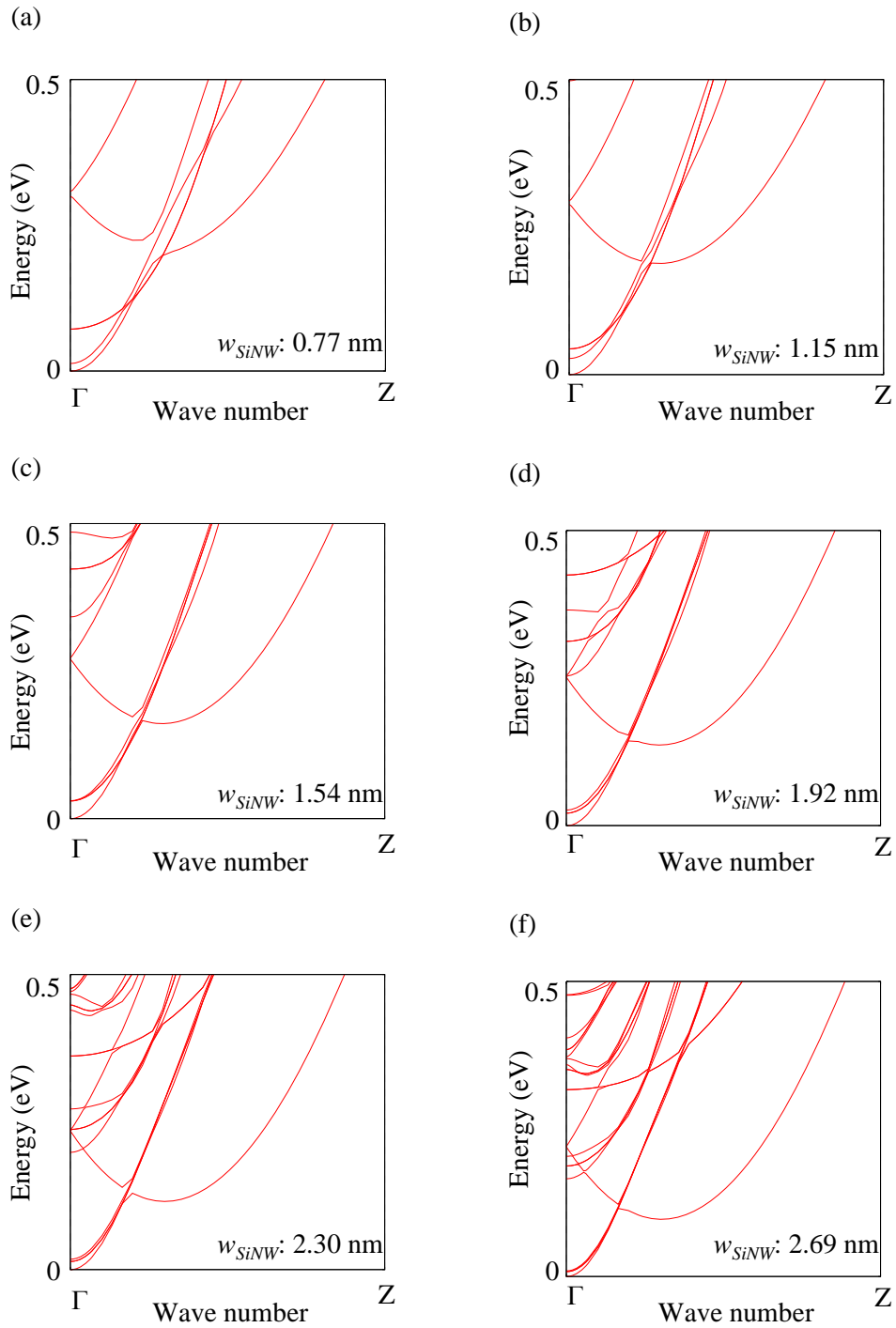
Fig. 2.4 Band structures of [100] SiNWs with w_{SiNW} ranging from (a) 0.77 to (f) 2.69 nm. The longitudinal axes indicate relative values from each band edge. Small SiNW has few subbands for strong quantum confinement.

2.2.2 Bandgap

Our calculation also showed that a bulk silicon have an indirect bandgap of 0.58 eV, and conduction band minimum was located at a point 84 % on the way from Γ to X. Although bandgaps estimated by DFT usually underestimate the experimental values, the tendencies on different size of the wire sample can be discussed. Figure 2.5 shows the w_{SiNW} dependence of bandgaps on [100] SiNWs. The bandgaps of the SiNWs become wide as the w_{SiNW} decreases, and a large bandgap of 2.55 eV was obtained with a 0.77-nm thick SiNW. As the w_{SiNW} increases, the primed subband moves toward unprimed subband minima and bandgaps approaches close to that of the bulk silicon, which is calculated by DFT in this case. The bandgaps estimated by DFT are about 0.5 eV smaller than values of other works with tight-binding method. [9,10,13] However, the tendency shows a good agreement.
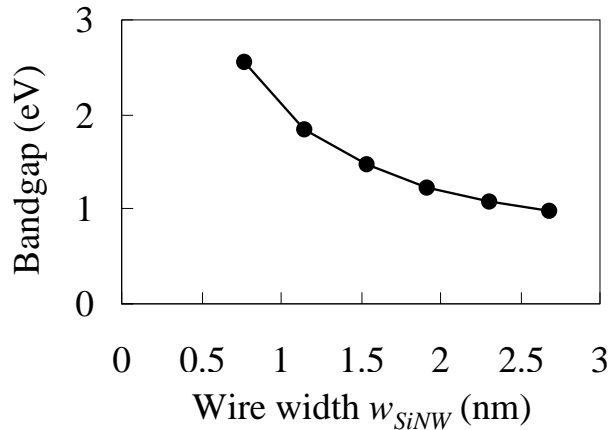

Fig. 2.5 w_{SiNW} dependence of bandgaps. The obtained bandgaps are underestimated because of DFT calculation. As w_{SiNW} increases, the bandgap closes to that of bulk silicon.

2.2.3 Electron effective mass

Figure 2.6 shows w_{SiNW} dependences of electron effective mass m^* of the lowest unprimed and primed subband in [100] SiNWs. The m^* becomes light as the w_{SiNW} increases in both unprimed and primed subband. With the increase in the size, one can expect that the m^* moves to 0.19 m_0 in unprimed subbands and 0.916 m_0 in primed subband, which corresponds to a transverse and longitudinal m^* of primed band minima of the bulk silicon respectively, where the m_0 denotes the electron mass. These wire size-dependent parameters are in good agreement with previous works too. [9]
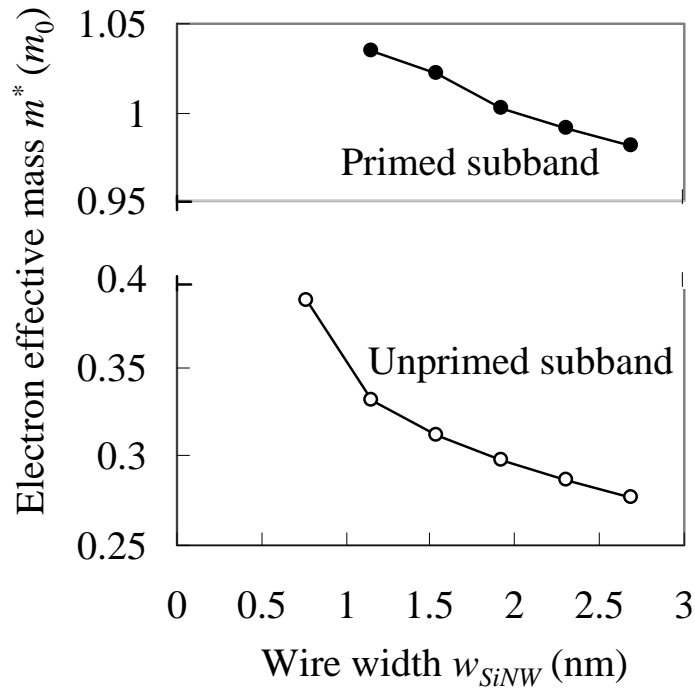

Fig. 2.6 w_{SiNW} dependence of m^*. Upper solid circles indicate effective masses of primed subband and lower open circles indicate them of unprimed subband. Both of them decrease as the w_{SiNW} increases.

2.2.4 Subband minima

Size-dependent i-th subband minima $E_{i\text{min}}$ are shown in fig. 2.7. $E_{0\text{min}}$ indicates the lowest subband minimum, which is the conduction band minimum. As w_{SiNW} increases, subband minima close to each other, and second groups of unprimed subbands below ($E_{0\text{min}} + 0.25$ eV) are shown in 2.30- and 2.69-nm width. Evolution of primed subband minima, which are connected by solid line, is important to discuss transport parameters in the section 3.2.
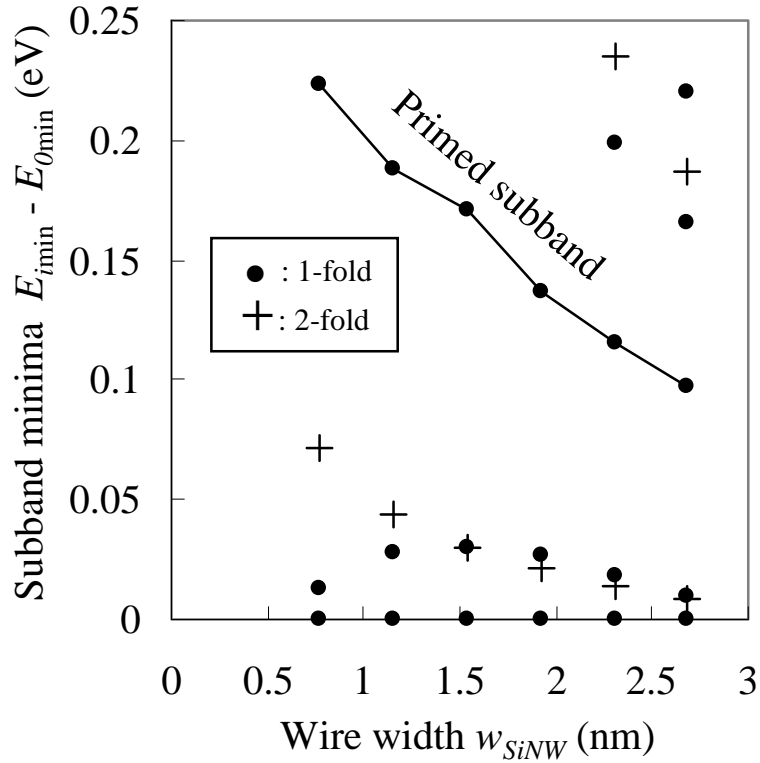

Fig. 2.7 w_{SiNW} dependences of the subband minima $E_{i\text{min}}$ based on $E_{0\text{min}}$. Solid circles (cross point) indicate 1-fold (2-fold) degenerate subband minima. Primed subband minima are connected by solid line.

3. On-currents for Silicon Nanowire FETs under Ballistic Transport

3.1 One-dimensional ballistic transport model (Natori model)

The ballistic transport characteristics of SiNW-FETs were derived from the derived band structures. An estimation of ballistic transport characteristics can be performed based on the Natori model.[18,20] The model is based on one-dimensional ballistic current. Calculated drain currents indicate theoretical maximum values obtained in same conditions. To investigate performance of each SiNW FET, we also compare transport parameters; Fermi level of source μ_s, linear charge density $|Q|$, saturation velocity v_{sat}, and a single wire on-current I_{ON}. In addition to account about those parameters, potential model for a nanowire FET, equations, quantum capacitance, example calculation, et al. are explained in following subsections.

3.1.1 Basic concept for the ballistic nanowire FET model

Potential model for a nanowire FET is briefly explained in the followings. Figure 3.1 (a) shows a potential profile from a source to a drain, where potential energy along the channel is the highest and not gradient at the bottleneck. The charge of forward and backward current at bottleneck of the channel is used to estimate the ballistic drain current. Figure 3.1 (b) shows an *E-k* dispersion of a channel and formed Fermi levels at the bottleneck, when a gate voltage V_g above threshold voltage V_{th} is applied. The μ_s and μ_d indicate Fermi levels of source and drain, respectively. The μ_s and μ_d have a relation of

$$\mu_d = \mu_s - qV_d,$$

where V_d denotes drain voltage. States of $dE/dk \geq 0$ and $dE/dk \leq 0$ fead forward and backward currents, respectively. The charges of each current are shown in fig. 3.1 (c)

which shows density of states (DOS) versus energy. Charge difference between the forward and the backward current, which are integral of the DOS over energy range from μ_d to μ_s at $T = 0$ K, determines the drain current per a single wire I_d.

Quantum capacitance

A transverse band diagram at the bottleneck is shown in fig. 3.2. A Gate overdrive $(V_g - V_{th})$ is applied for variation within the insulator ϕ_i and a band drop ($\mu_s - E_{0\min}$ or $\mu_d - E_{0\min}$) within the channel. Multiplying gate capacitance C_g and ϕ_i is same with the charge, stored carriers. Unless there is the band drop, the enough carriers can not be stored in the channel. The band drop causes that all of the gate overdrive $(V_g - V_{th})$ is not applied for gate capacitance C_g, and it seems like a decrease in gate capacitance for the same $(V_g - V_{th})$. The decrease in capacitance is represented as an addition of series connected capacitance, and the series connected capacitance is called the quantum capacitance C_q. An effective capacitance C_{eff} including the quantum capacitance is also adopted.
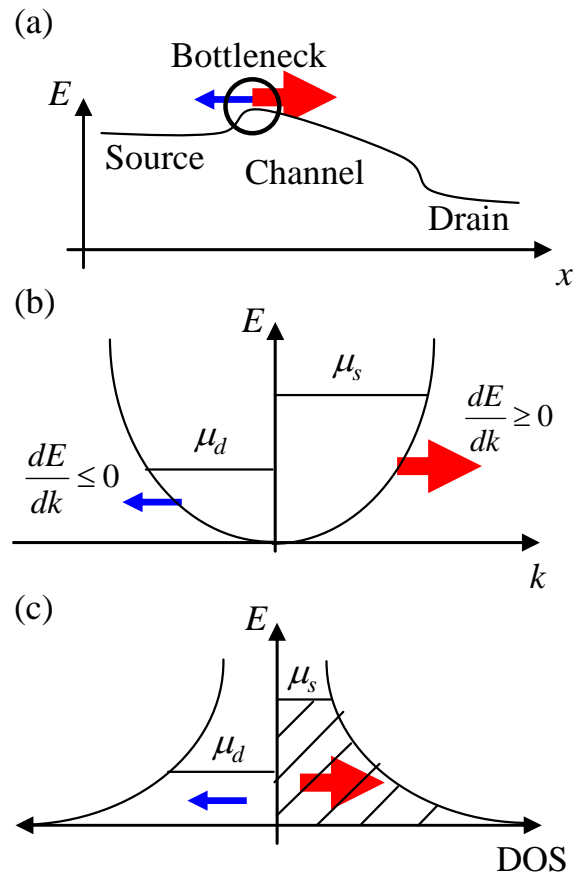
Fig. 3.1 (a) Potential profile from a source to a drain of nanowire FET. There is bottleneck point where forward and backward current are calculated. (b) Parabolic E-k dispersion, and formed Fermi levels at the bottleneck when V_g and V_d are applied. States of $dE/dk \geq 0$ ($dE/dk \leq 0$) become forward (backward) current. (c) DOS versus energy. A right slashed (left open) area indicates an amount of charge for forward (backward) current at $T = 0$ K.

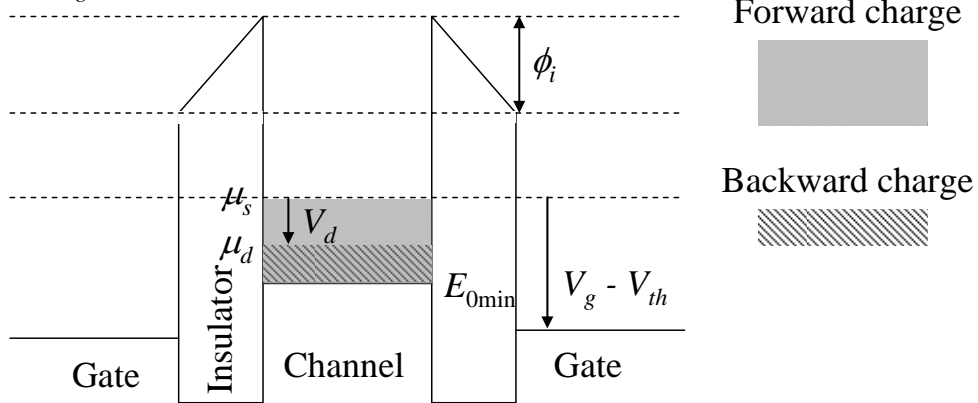a) V_g - V_{th} = 0
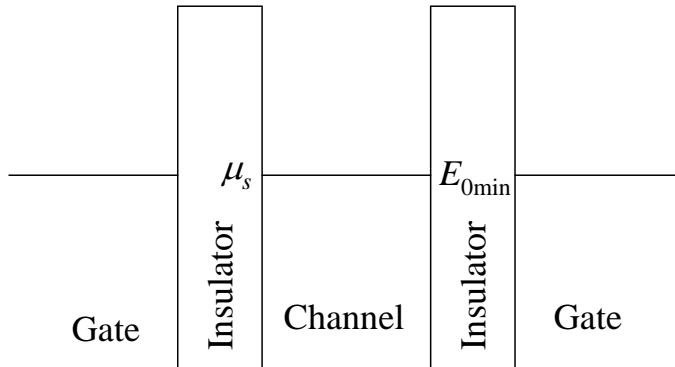
b) V_g - V_{th} > 0, V_d > 0

Fig. 3.2 a) A transverse band diagram when a gate overdrive ($V_g - V_{th}$) is zero. A value of 0-th subband minimum E_{0min} is same with a value of Fermi level of source μ_s. b) A transverse band diagram when ($V_g - V_{th}$) and a drain-voltage V_d are larger than zero. Grey and slashed area indicates charge of forward and backward carriers, respectively. The ϕ_i denotes variation within the insulator.

3.1.2 Equations

From one-dimensional tunnel current simplified by neglected transmit of states

below potential bottleneck, one-dimensional drain current is expressed by[1]

$$I_d = \frac{q}{\pi\hbar}\sum_i \int_{Ei\min}^{\infty}(f(E,\mu_s) - f(E,\mu_d))dE,\qquad(3.1)$$

where q and $\hbar$ denotes elementary charge and reduced Plank's constant, respectively.

In section 2, elementary charge is denoted by e, and e is usually used as elementary

charge in physics. The $E_{i\min}$ denotes minimum of i-th subband, where subband

minimum belong to conduction band minimum (CBM) is $E_{0\min}$, and $f(E, \mu_s)$ denotes

Fermi distribution function described as

$$f(E,\mu_s) = \frac{1}{1 + \exp\left(\dfrac{E - \mu_s}{k_B T}\right)}.$$

Importing band structures which have band maxima $E_{i\max}$ or branches, integration of eq.

(3.1) can be carried out. So that, I_d can be derived as

$$I_d = \frac{qk_B T}{\pi\hbar}\sum_i g_i \left\{ \sum_{dE(k)/dk \geq 0 branch} \ln\left(\frac{1 + \exp[(\mu_s - E_{i\min}^+)/k_B T]}{1 + \exp[(\mu_d - E_{i\max}^+)/k_B T]}\right)\right.$$

$$\left. - \sum_{dE(k)/dk \leq 0 branch} \ln\left(\frac{1 + \exp[(\mu_s - E_{i\min}^-)/k_B T]}{1 + \exp[(\mu_d - E_{i\max}^-)/k_B T]}\right)\right\},\qquad(3.2)$$

where g_i denotes degeneracy of i-th subband, and $E_{i\min}$ and $E_{i\max}$ with suffix plus

(minus) indicates maximum and minimum energy in the positive (negative) velocity

branches of the i-th suband as shown in fig. 3.3. However, $E_{i\max}$ can be neglected when

they are much higher than Fermi level of source μ_s. In our calculation, $E_{i\max}$ is neglected

because they are much higher than Fermi level of source μ_s as shown in sections 2.2.1

---

[1] Appendix B

and 3.2.1. Thus, the drain current I_d can be reduced as

$$I_d = G_0 \left(\frac{k_B T}{q} \right) \sum_i \sum_{branch} g_i \ln \left(\frac{1 + \exp[(\mu_s - E_{i\min})/k_B T]}{1 + \exp[(\mu_d - E_{i\min})/k_B T]} \right), \qquad (3.3)$$

where the total current is the summation of carrier flows in each subband and branch. The G_0 (= $q^2/\pi\hbar$) denotes the quantum conductance of 77.8 µS. Equation (3.3) is a simplified equation by neglecting maxima of subbands which are much higher than the μ_s. In this equation, tunnel current is also neglected. To calculate eq. (3.3), we need the μ_s in addition to the $E_{i\min}$. When the gate overdrive ($V_g - V_{th}$) and the linear gate capacitance C_g are given, the μ_s can be calculated by

$$\frac{|Q|}{C_g} = (V_g - V_{th}) - \frac{\mu_s - E_{0\min}}{q}, \qquad (3.4)$$

where $|Q|$ denotes linear density of carriers along channel, and quantum capacitance C_q is derived from the second term of right hand side of eq. (3.4) as shown in fig. 3.2 (b). The $|Q|$ can also described as

$$|Q| = 2q \sum_i g_i \left(\sum_{dE(k)/dk \geq 0 branch} \int_{E_{i\min}^+}^{E_{i\max}^+} D_{i+}(E) f(E, \mu_s) dE \right.$$

$$\left. + \sum_{dE(k)/dk \leq 0 branch} \int_{E_{i\min}^-}^{E_{i\max}^-} D_{i-}(E) f(E, \mu_d) dE \right), \qquad (3.5)$$

where $D_{i+}(E)$ and $D_{i-}(E)$ denote density of state of the positive and negative velocity branches of i-th subband. In this case, $E_{i\max}$ can be neglected too. Integration over E in eq. (3.5) is also changed to integration over k in order to easier calculation using derived E-k dispersion. So that, the $|Q|$ is reduced as

$$|Q| = \frac{q}{\pi} \sum_i g_i \int_\Gamma^Z \left\{ \frac{1}{1 + \exp\left(\dfrac{E_i(k) - \mu_s}{k_B T}\right)} + \frac{1}{1 + \exp\left(\dfrac{E_i(k) - \mu_d}{k_B T}\right)} \right\} dk \quad . \qquad (3.6)$$

Fermi distribution functions are integrated within the Brillouin zone. The C_g of gate-all-around structures of the square cross section can not be estimated analytically, so that we approximately adopt a cylindrical model as shown in fig. 3.4. The C_g of the cylindrical model can be derived as

$$C_g = \frac{2\pi\varepsilon}{\ln\left(\dfrac{r + t_{ox}}{r}\right)} , \qquad (3.7)$$

where r indicates a radius of circle having the same cross-sectional area as square cross-sectional SiNW adopted for band structure calculation, and ε indicates dielectric constant of the insulator. The μ_s is derived by solving eqs. (3.4) and (3.6) simultaneously. Finally, if μ_s is substituted to eq. (3.3), we can obtain the I_d.
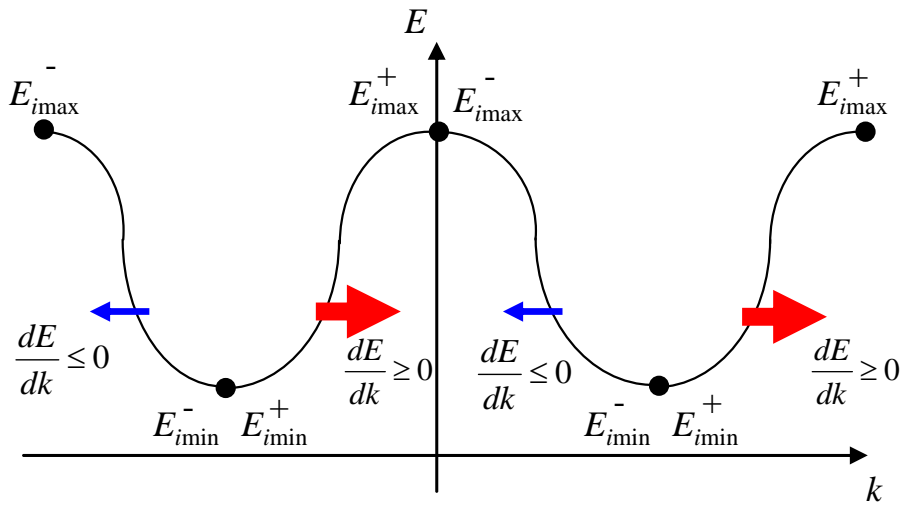
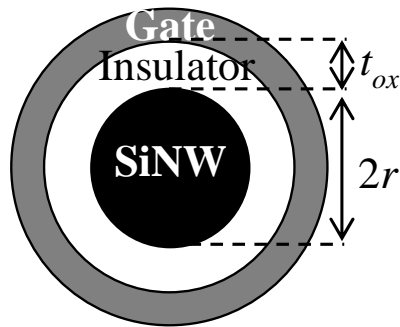Fig. 3.3 $E_{i\min}$ and $E_{i\max}$ with suffix plus (minus) in branches of i-th subband are represented.

Fig. 3.4 Cross-sectional view of FET in gate-all-around structure. We assumed SiO_2 with t_{ox} of 1 nm as an insulator in our calculations. Cylindrical model with same cross-sectional area is adopted for estimation of C_g.

3.1.3 Equation transformation for easy discussion

Conventionally, ballistic I_d can be expressed by the following equation,

$$I_d = C_{eff}(V_g - V_{th})v_{inj}, \tag{3.8}$$

where the v_{inj} denotes a mean velocity of the charge at the bottleneck point, which is called mean injection velocity. C_{eff} denotes linear effective capacitance considering the quantum capacitance C_q which is derived as

$$C_{eff}(V_g - V_{th}) = C_g \left\{ (V_g - V_{th}) - \frac{\mu_s - E_{0\min}}{q} \right\} .$$

Replacing v_{inj} by saturation mean injection velocity v_{sat} in eq. (3.8), which consist of only forward current except backward current, the saturation I_d is obtained. Here, on-current is defined as the saturation I_d. Also replacing $C_{eff}(V_g - V_{th})$ by $|Q|$ in eq. (3.8), single-wire on-current I_{ON} can be described as

$$I_{ON} = |Q|v_{sat}. \tag{3.9}$$

Considering results with eq. (3.9) helps understanding related to parameters to determine the I_{ON}.

3.1.4 Calculation example

Figure 3.5 shows an example for I_d-V_d characteristic of [100] SiNW with w_{SiNW} of

2.69 nm which was calculated in various gate overdrives ($V_g - V_{th} = 0.1, 0.4, 0.7$ and 1.0

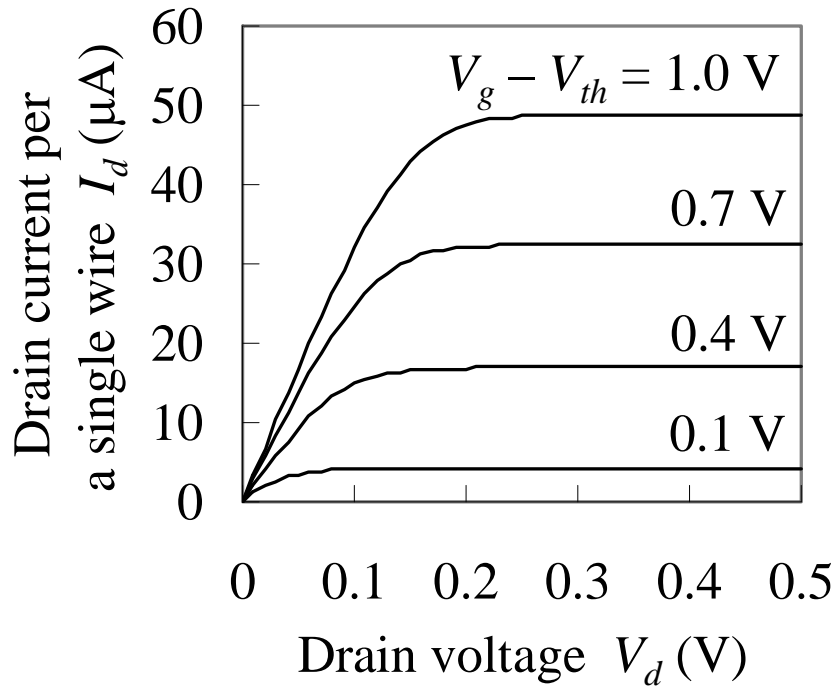V) at room temperature ($T = 300$ K). SiO$_2$ with t_{ox} of 1 nm is adopted as the gate

insulator.

Fig. 3.5 An example for I_d-V_d characteristic at $T = 300$ K. The model is square SiNW
with w_{SiNW} of 2.69 nm. Applied gate overdrives range from 0.1 V to 1.0 V.

3.2 Size-dependent transport characteristics

The transport parameters are derived under gate overdrive $V_g - V_{th} = 1.0$ V at 300 K. The cylindrical wire model with same cross-sectional area is approximately assumed, and 1-nm-thick SiO_2 is adopted as a gate insulator in the same way as the example calculation in previous section 3.1.5. Calculated Fermi level of source μ_s, linear charge density $|Q|$ and effective capacitance C_{eff}, saturation mean velocity v_{sat}, and on-current I_{ON} are shown in following subsections, where those parameters are divided into each subband. Finally, in the case of parallel multi-channel FET, on-current and capacitance evolution is also shown.

3.2.1 Fermi level of source

Based on a w_{SiNW} dependence of the subband minima as shown in fig. 2.5, a size-dependent μ_s is shown in fig. 3.6. The μ_s increases as w_{SiNW} increases, and gradually saturates in large wires. Here, the μ_s is determined by the balance between DOS and effective capacitance, both of which increase as wire size increases. The μ_s of further larger wire is expected to decrease due to sinking of the upper subbands, which have the larger degeneracy and DOS. Subbands in the energy range in this figure mainly determine transport characteristics because the states higher than about 0.1 eV above the μ_s are almost unoccupied based on Fermi distribution, resulting in little contribution to on-current.
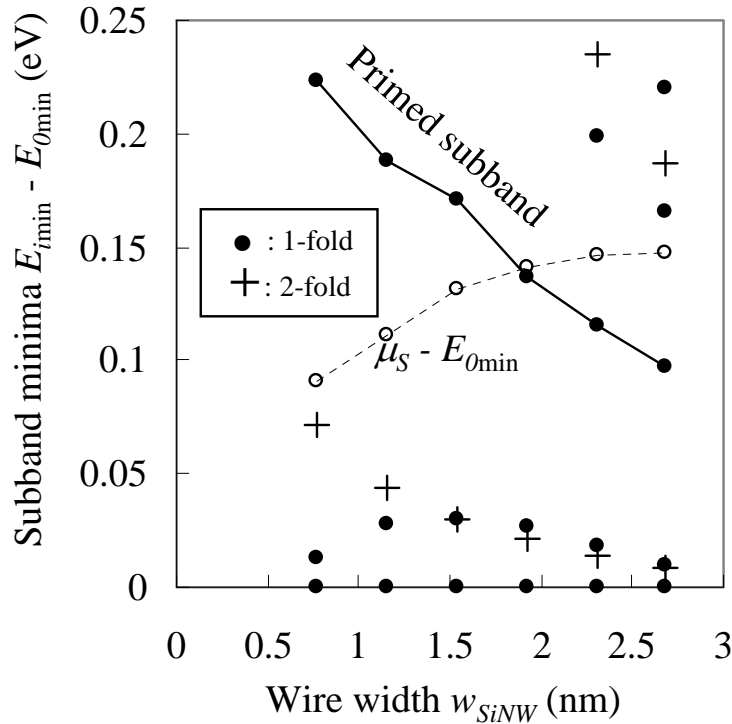

Fig. 3.6 w_{SiNW} dependences of the subband minima and μ_s based on $E_{0\mathrm{min}}$. Solid circles (cross point) indicate 1-fold (2-fold) degenerate subband minima. Primed subband minima are connected by solid line. Open circles with dotted line indicate the μ_s. Evolution of primed subband minima and the μ_s are important to investigation of on-current.

3.2.2 Linear charge density and effective capacitance

Figure 3.7 shows linear charge density $|Q|$ of each subband versus the w_{SiNW}. The Linear charge density is the charge density per unit wire length. Top of the bars indicates $C_g(V_g - V_{th})$, where C_g derived by eq. (3.7), which has the same numerical value as the C_g because of gate overdrive $V_g - V_{th} = 1.0$ V. This capacitance is almost proportional to the w_{SiNW} as excepted. The capacitance, however, is not exactly proportional to the wire periphery in relatively small wires as compared with gate insulator thickness t_{ox}, and C_g-intercept of approximately linear slope is not zero in that case. In fact, all of the $(V_g - V_{th})$ is not applied to the nanowire channel due to series connection of the quantum capacitance C_q. On this account, the effective capacitance C_{eff}, which denotes effective capacitance for the same $(V_g - V_{th})$, smaller than the C_g, and the obtained μ_s are $0.09 \sim 0.15$ eV for $V_g - V_{th} = 1.0$ V, so that the C_{eff} is $9 \sim 15$ % smaller than the C_g. The charge in unprimed subband is almost constant the w_{SiNW} larger than 1.54 nm, and excess charge is stored in primed and other higher subbands. Charge in other higher subbands shows up only in the 2.69-nm w_{SiNW}, where the higher subbands stay close to μ_s.
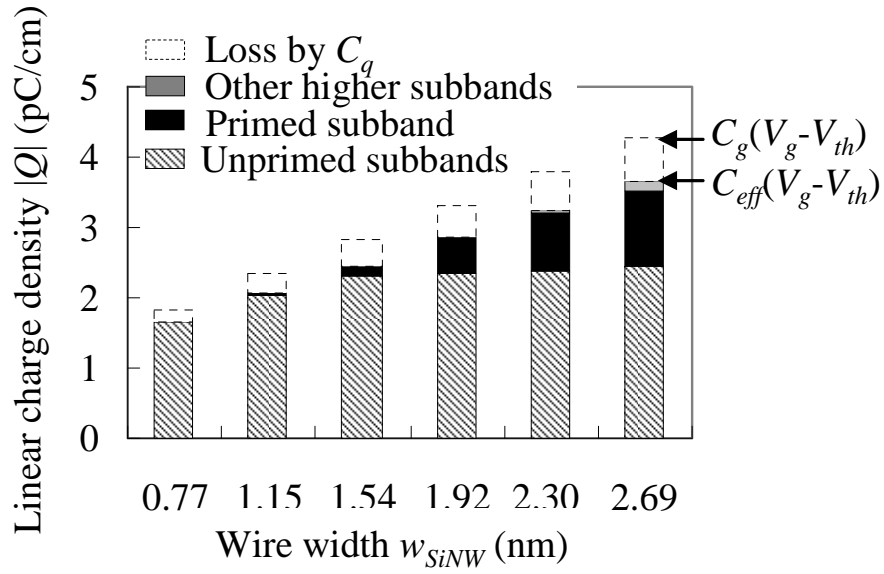
Fig. 3.7 w_{SiNW} dependences of $|Q|$ of each subband. Top of the bars indicates the $C_g(V_g - V_{th})$. The $C_{eff}(V_g - V_{th})$ is represented by rest bars except open dotted square, loss by C_q. Slashed, black and grey square area indicate linear charge density in unprimed, primed, other higher subbands, respectively.

3.2.3 Injection velocity

Not only m^* but also the position of subband minima affects v_{sat}, which denotes the injection velocity for saturation current. When μ_s is large, the v_{sat} becomes large unless there is subband minimum near the μ_s, because occupied states with the small kinetic energy in higher subbands make the v_{sat} to decrease. Basically, subbands at a high energy are disadvantageous for v_{sat} enhancement. The subband minima depend on the geometric structures of SiNWs, and the subband separation decreases as the SiNWs size increases. Therefore, the v_{sat} decreases in large wires unless the μ_s becomes sufficiently large.

Figure 3.8 shows v_{sat} of each subbnad versus the w_{SiNW}. The decrease of m^* as well as the increase of μ_s contributes increasing of the v_{sat} of entire subbands. v_{sat} of the primed subband is much smaller than v_{sat} of the unprimed subbands, and the both of them increase as the w_{SiNW} increases. It is also caused by the both subband minima sinking below μ_s and the m^* decreasing. Despite the increase of μ_s and decrease of m^*, the v_{sat} of entire subbands gradually saturates and peaks out at 2.30-nm w_{SiNW}, where most of the excess charge distributes in the primed subband with small v_{sat} as shown in fig. 3.7.
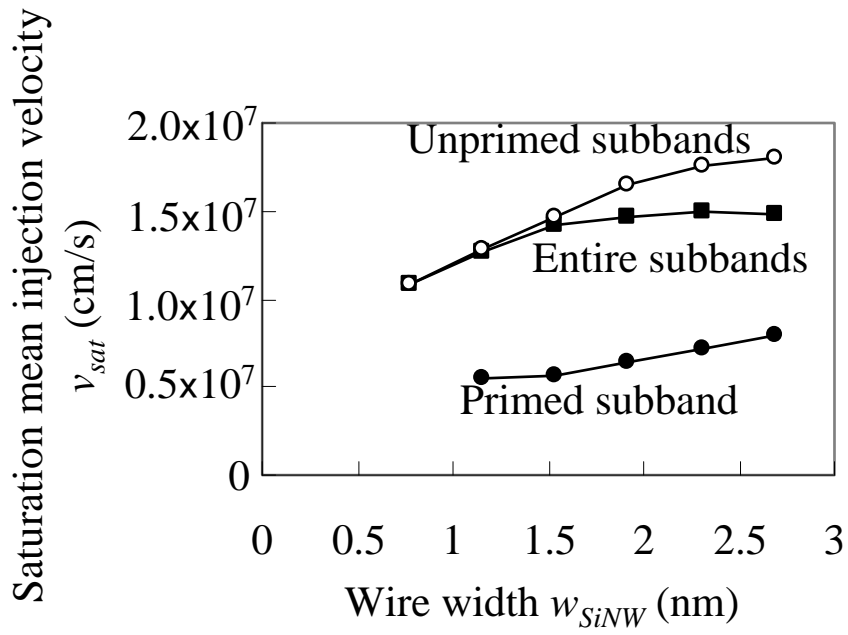
Fig. 3.8 w_{SiNW} dependences of v_{sat} of each subband. Solid square, open circle and solid circle indicate v_{sat} of entire subbands, unprimed subbands and unprimed subband, respectively. The v_{sat} of entire subbands peaks out at w_{SiNW} of 2.30 nm.

3.2.4 On-current

Multiplying the extracted |Q| and v_{sat} of entire subbands, the I_{ON} is obtained as shown in fig. 3.9. Although linear charge density |Q| of unprimed subbands is almost constant in a large wire, the on-current of unprimed subbands steadily increases due to the v_{sat} increasing. A large I_{ON} is obtained from large SiNWs because the increase of linear capacitance is more dominant for current increase than variation of entire subbands' v_{sat}. In addition, it can cause the rapid decrease of the μ_s and the total v_{sat} that subbands in sufficiently large degeneracy move down in large wire, so that the I_{ON} would decrease.
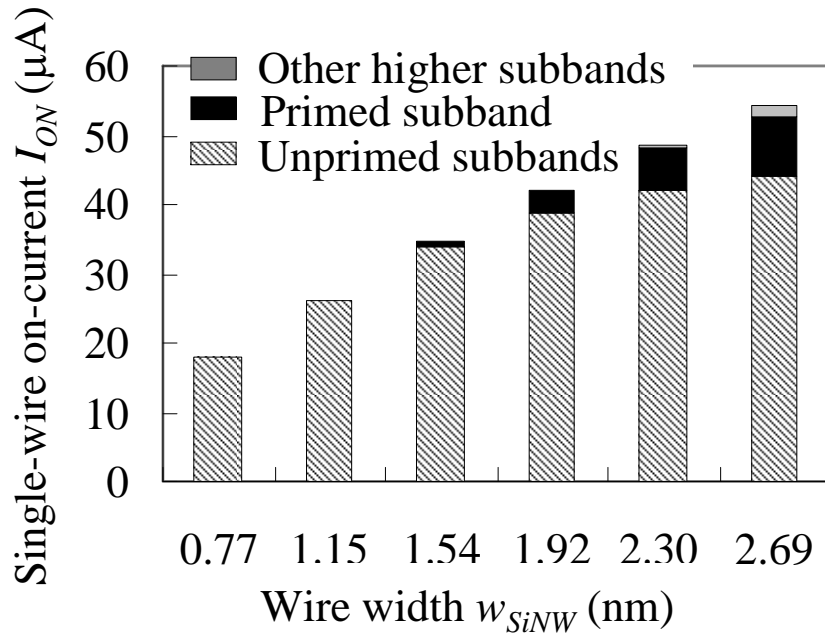

Fig. 3.9 w_{SiNW} dependences of I_{ON} of each subband. Top of the bars indicates the total I_{ON}. Slashed, black and grey square area indicate on-current of unprimed, primed and other higher subbands among total I_{ON}, respectively.

3.2.5 Multi-channel FET

To discuss the performance of a multi-channel FET with parallel SiNWs, two factors have to be considered; how much I_{ON} can be obtained, and how many wires can be aligned in the same width. We assumed the multi-channel SiNW-FET model shown in fig. 3.10. Cylindrical wires, which have the same cross-sectional area as the square cross-sectional wires whose band structures are calculated, are adopted as an approximation. The influences of the neighboring wires are neglected for simplicity. The number of wires per 1-μm width N is calculated by 1-μm/($2r + 2t_{ox} +$ spacing). w_{SiNW} dependences of $N|Q|$ with a spacing of 0 and 5 nm are shown in fig. 3.11 (a). It is expected that the $N|Q|$ becomes constant when w_{SiNW} is large enough, because C_{eff} is proportional to w_{SiNW} in sufficiently large wires, and N tends to be inversely proportional to w_{SiNW} when ($2t_{ox} +$ spacing)/w_{SiNW} is sufficiently small. In the case of nano-sized wires, the $N|Q|$ increases as the w_{SiNW} increases. Fig. 3.11 (b) shows w_{SiNW} shows w_{SiNW} dependences of NI_{ON}. The NI_{ON} evolution shows a gentler slope in large wires because of the nonincreasing v_{sat} of entire subbands in the large wires.

Weakly confined wires will have sufficiently large number of degenerate subbands sinking below μ_s, so that v_{sat} will decrease. Then the NI_{ON} will decrease due to constant $N|Q|$ of large wires. In addtion, a qualitatively similar size-dependent NI_{ON} can be shown in other one-dimensional ballistic FETs as well as in [100] SiNW FET. It is a matter of course that the size-dependent linear charge density is almost the same because the C_q can be neglected on account of sufficiently small μ_s. No matter what kind of semiconductor is imported, well-confined nanowires have faster carriers due to their few subbands in opposition to weakly confined wires, and over-confined nanowires have slower carriers with purely small μ_s due to small linear capacitances, so

that it can also be expected that qualitative analyses of size-dependent v_{sat} are similar.
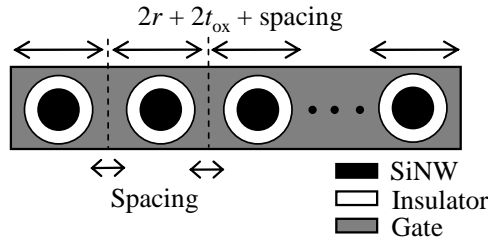

Fig. 3.10 Model of the multi-channel SiNW-FET. There is one SiNW-FET per ($2r + 2t_{ox}$ + spacing). A number of SiNWs per 1-μm width, N, is calculated by 1-μm/($2r + 2t_{ox}$ + spacing).
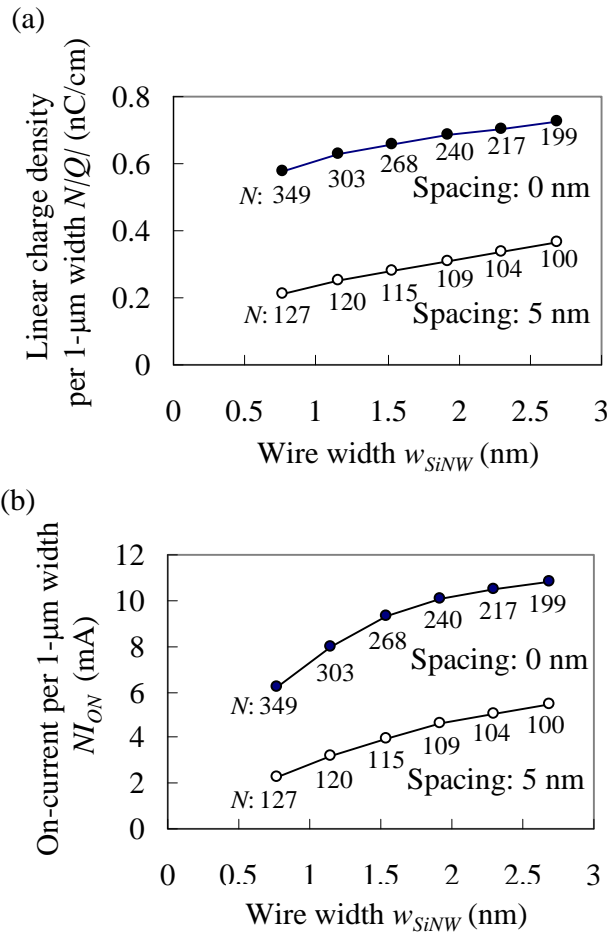

Fig. 3.11 (a) w_{SiNW} dependences of $N|Q|$ when spacings are 0 and 5 nm. Numbers under lines indicate N corresponding to the w_{SiNW}. (b) w_{SiNW} dependences of NI_{ON} when spacings are 0 and 5 nm. Numbers under lines indicate N corresponding to the w_{SiNW}. NI_{ON} evolution of SiNW with large w_{SiNW} has gradual slope.

4. Conclusions

Size-dependent bandgaps and electron effective masses of the [100]-directed SiNWs with the width ranging from 0.77 to 2.69 nm have been calculated by first-principles calculation. The prominent features of SiNWs are the increase in bandgaps, the band splitting, and the increase in effective masses due to quantum confinement. Ballistic drain-current estimations from the calculated band structures have revealed the size dependence of the Fermi level of source μ_s, the linear charge density $|Q|$, the saturation mean injection velocity v_{sat}, which determines the single-wire on-current I_{ON}. Subband structure and parameters which contribute to the on-current evolution have been clarified by systematical analyses. The I_{ON} has steadily increased with large wire size owing to the increase in $|Q|$ despite peaked v_{sat} of the entire subbands. In addition, size-dependent band structures have much effect on modulation of ballistic transport characteristics, and those effects of each subband have also been changed as size modulates. Finally, an assessment of the ballistic drain-current for practical multi-channel SiNW FETs has revealed a trade-off between v_{sat} and the number of wires N, determined by geometrical parameters, spacing and wire-size, which is in contrast to the case of a single wire.

Acknowledgements

Appendix A Proof of theorem 1 and 2 of density functional theory

Theorem 1

Minimum of $E[n]$, which is a function of one electron density, becomes ground state energy,

$$E[n] = \int v_{ext}(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r} + F[n] \geq E_{GS} \text{ . same with eq. (2.1) (A·1)}$$

Proof

In eq. (A·1), $n(\boldsymbol{r})$ have to satisfy with N-representability. N-representability is that the one electron density $n(\boldsymbol{r})$ for N electron is obtained from anitisymmetric wave function ψ^n described as

$$n(\boldsymbol{r}) = N \int \left| \psi^n(x_1, x_2, \cdots, x_N) \right|^2 d\xi_1 dx_2 \cdots dx_N, \qquad x \equiv (\boldsymbol{r}, \xi),$$

where x consist of position coordinate $\boldsymbol{r}$ and spin coordinate ξ.

Let's prove theorem 1 upon this N-representability. Hamiltonian $\hat{H}$ of interact N electron is described as

$$\hat{H} = \hat{V}_{ext} + \hat{T} + \hat{V}_{ee}$$

$$= \sum_i^N v_{ext}(\boldsymbol{r}_i) - \frac{\hbar^2}{2} \sum_i^N \nabla_i^2 + \frac{1}{2} \sum_{i \neq j}^N \frac{e^2}{\left| \boldsymbol{r}_i - \boldsymbol{r}_j \right|},$$

where $\hat{V}_{ext}$ denotes Hamiltonian of external potential, and e and $\hbar$ denotes elementary charge and reduced Plank's constant respectively. Using wave function of ground state ψ_{GS}, ground state energy is described as

$$E_{GS} = \left\langle \psi_{GS} \left| \hat{V}_{ext} + \hat{T} + \hat{V}_{ee} \right| \psi_{GS} \right\rangle. \tag{A·2}$$

From N-representability, expectation value of external potential $\hat{V}_{ext}$ with $\psi_{\min}$ is transformed as

$$\left\langle \psi^n \left| \hat{V}_{ext} \right| \psi^n \right\rangle = \left\langle \psi^n \left| \sum_i v_{ext} \right| \psi^n \right\rangle$$

$$= \sum_i \int \psi^{n^*}(x_1, x_2, \cdots, x_N) v_{ext}(r_i) \psi^n(x_1, x_2, \cdots, x_N) dx_1 \cdots dx_N$$

$$= \frac{1}{N} \sum_i \int v_{ext}(\boldsymbol{r}_i) n(\boldsymbol{r}_i) d\boldsymbol{r}_i$$

$$= \int v_{ext}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} . \qquad (A\cdot 3)$$

Expectation value of $\hat{V}_{ext}$ also satisfies

$$\left\langle \psi^n \left| \hat{V}_{ext} \right| \psi^n \right\rangle = \left\langle \psi^n_{min} \left| \hat{V}_{ext} \right| \psi^n_{min} \right\rangle, \qquad (A\cdot 4)$$

because eq. (A·3) shows that the expectation value is determined by only $n(\boldsymbol{r})$ and ψ^n_{min} is one of the ψ^n. Finally, Using eqs. (A·3) and (A·4),

$$\left\langle \psi^n_{min} \left| \hat{V}_{ext} \right| \psi^n_{min} \right\rangle = \int v_{ext}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r}$$

is satified. Using this relationship, $E[n]$ satisfies

$$E[n] = \int v_{ext}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} + F[n]$$

$$= \left\langle \psi^n_{min} \left| \hat{V}_{ext} \right| \psi^n_{min} \right\rangle + \left\langle \psi^n_{min} \left| \hat{T} + \hat{V}_{ee} \right| \psi^n_{min} \right\rangle$$

$$= \left\langle \psi^n_{min} \left| \hat{V}_{ext} + \hat{T} + \hat{V}_{ee} \right| \psi^n_{min} \right\rangle. \qquad (A\cdot 5)$$

Be careful ψ^n_{min} is wave function which minimizes the $F[n]$. We can see that $E[n]$ has same Hamiltonian with E_{GS} from eqs. (A·2) and (A·5). Therefore, the E_{GS} equals minimum of the $E[n]$.

Theorem 2

E_{GS} is same with $E[n_{GS}]$, where n_{GS} denotes one electron density of ground state, described as

$$E_{GS} = E[n_{GS}] = \int v_{ext}(\boldsymbol{r})n_{GS}(\boldsymbol{r})d\boldsymbol{r} + F[n_{GS}].$$

Proof

$$\left\langle \psi_{GS} \left| \hat{V}_{ext} + \hat{T} + \hat{V}_{ee} \right| \psi_{GS} \right\rangle \le \left\langle \psi_{\min}^{n_{GS}} \left| \hat{V}_{ext} + \hat{T} + \hat{V}_{ee} \right| \psi_{\min}^{n_{GS}} \right\rangle \qquad (A\cdot 6)$$

is shown from theorem 1, since $\psi_{\min}^{n_{GS}}$ minimizes expectation value of only $\hat{T} + \hat{V}_{ee}$ and left part is ground sate energy. Here, it is assumed that ground state is not degenerate. Because the ψ_{GS} yields same $n_{GS}(\boldsymbol{r})$ with the $\psi_{\min}^{n_{GS}}$ although ψ_{GS} can be different wave function with $\psi_{\min}^{n_{GS}}$,

$$\left\langle \psi_{GS} \left| \hat{V}_{ext} \right| \psi_{GS} \right\rangle = \left\langle \psi_{\min}^{n_{GS}} \left| \hat{V}_{ext} \right| \psi_{\min}^{n_{GS}} \right\rangle \qquad (A\cdot 7)$$

stands up. Substituting (A·7) into (A·6),

$$\left\langle \psi_{GS} \left| \hat{T} + \hat{V}_{ee} \right| \psi_{GS} \right\rangle \le \left\langle \psi_{\min}^{n_{GS}} \left| \hat{T} + \hat{V}_{ee} \right| \psi_{\min}^{n_{GS}} \right\rangle \qquad (A\cdot 8)$$

is derived. Here, from definition of $F[n]$,

$$\left\langle \psi_{GS} \left| \hat{T} + \hat{V}_{ee} \right| \psi_{GS} \right\rangle \ge \left\langle \psi_{\min}^{n_{GS}} \left| \hat{T} + \hat{V}_{ee} \right| \psi_{\min}^{n_{GS}} \right\rangle \qquad (A\cdot 9)$$

is also obtained. From inequality sign of eqs. (A·8) and (A·9), it is shown that left part and right part of those equations have same expectation value. Therefore,

$$E_{GS} = \left\langle \psi_{GS} \left| \hat{V}_{ext} + \hat{T} + \hat{V}_{ee} \right| \psi_{GS} \right\rangle = \left\langle \psi_{\min}^{n_{GS}} \left| \hat{V}_{ext} + \hat{T} + \hat{V}_{ee} \right| \psi_{\min}^{n_{GS}} \right\rangle = \int v_{ext}(r)n_{GS}(r)dr + F[n_{GS}] = E[n_{GS}]$$

is derived.

Appendix B Landauer's formula for one-dimensional current

Figure B・1 (a) shows an one-dimensional tunnel current. Using Fermi distribution function f, the one-dimensional tunnel current flowing from left to right I_L and from right to left I_R are described as

$$I_L = q \int_{E_L}^{\infty} f(E, \mu_L) D(E) v(E) T(E) dE \qquad (B・1)$$

and

$$I_R = -q \int_{E_R}^{\infty} f(E, \mu_R) D(E) v(E) T(E) dE , \qquad (B・2)$$

where q denotes elementary charge. E_L (E_R) and μ_L (μ_R) denote a band minimum and Fermi level on the left (right), respectively. $D(E)$, $v(E)$ and $T(E)$ also denote density of states, velocity and transmission coefficient, respectively. The transmission coefficient T is not dealt strictly in this discussion. In the case of one-dimension, a number of states in dE over energy, $D(E)dE$ are described as

$$D(E)dE = 2\frac{dk}{2\pi}$$

and

$$D(E) = \frac{2}{2\pi} \frac{dk}{dE} , \qquad (B・3)$$

where first coefficient 2 on the left hand side indicates two spin states. One-dimensional $v(E)$ can be also described as

$$v(E) = \frac{\hbar k}{m} = \frac{1}{\hbar} \frac{dE}{dk} , \qquad (B・4)$$

where $\hbar$ and m denotes reduced Plank's constant and mass, respectively. Substituting eqs. (B・3) and (B・4) in eqs. (B・1) and (B・2) and eliminating transmission coefficient of eq. (B・2) between E_L and E_R, currents are described as

$$I_L = \frac{q}{\pi\hbar} \int_{E_L}^{\infty} f(E, \mu_L) T(E) dE$$

and

$$I_R = -\frac{q}{\pi\hbar} \int_{E_L}^{\infty} f(E, \mu_R) T(E) dE .$$

Therefore, total current is described as

$$I = I_L + I_R = \frac{q}{\pi\hbar} \int_{E_L}^{\infty} \left(f(E, \mu_R) - f(E, \mu_R) \right) T(E) dE . \qquad (B \cdot 5)$$

At temperature of 0 K, eq. (B·5) is transformed into simpler equation as

$$I = \frac{q}{\pi\hbar} (\mu_R - \mu_L) T$$

$$= \frac{q^2}{\pi\hbar} V_d T ,$$

and conductance G is described as

$$G = \frac{q^2}{\pi\hbar} T . \qquad (B \cdot 6)$$

Equation (B·6) is the simplest Landauer's formula.

In our one-dimensional ballistic current model, there is not potential barrier along the channel, and current at the bottleneck is considered as shown in fig. A·1 (b). Replacing E_L, μ_L and μ_R by i-th subband minimum $E_{i\min}$, Fermi level μ_s of source and Fermi level of drain μ_d in eq (B·5), drain-current I_d of i-th subband is described as

$$I_d = \frac{q}{\pi\hbar} \int_{E_{i\min}}^{\infty} \left(f(E, \mu_s) - f(E, \mu_d) \right) dE . \qquad (B \cdot 7)$$

Considering charge in this discussion, we have to pay attention to the subband minimum. Although there is not charge at the potential barrier in the case of tunnel current, the sum of charge $|Q|$ into right and left side is calculated by

$$|Q| = q\left(\int_{E_R}^{\infty} D(E)f(E,\mu_R)dE + \int_{E_L}^{\infty} D(E)f(E,\mu_L)dE \right).$$

In the case of ballistic current, the sum of *i*-th subband charge $|Q_i|$ flowing forward and backward at the bottleneck can also be calculated by

$$|Q_i| = q\left(\int_{E_{i\,\min}}^{\infty} D(E)f(E,\mu_s)dE + \int_{E_{i\,\min}}^{\infty} D(E)f(E,\mu_d)dE \right), \quad\quad (B \cdot 8)$$

where minima of band of each side are same in contrast to the case of tunnel current because forward and backward current have same states at the bottleneck. Besides, the reason why Fermi level of source and drain are shown at the bottleneck is carriers of both far sides, the source and the drain, don't get or loose energy. Finally, eqs. (B·7) and (B·8) are used for estimation of ballistic current in this thesis.
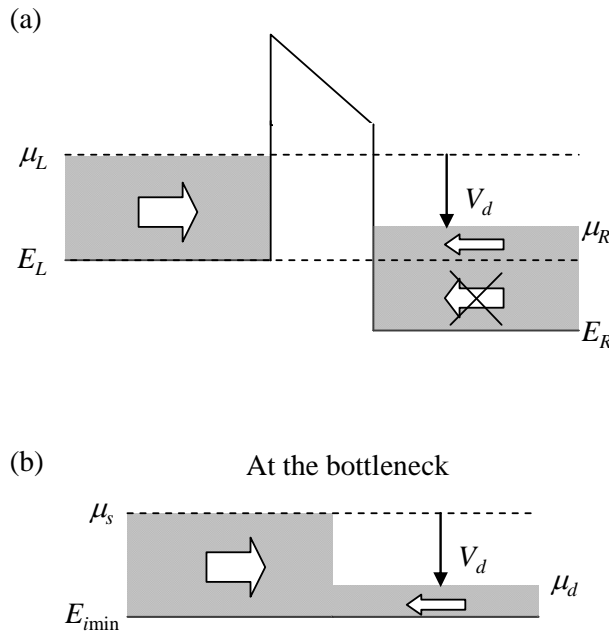

Fig. B·1 (a) A band diagram under one-dimensional tunnel current, when drain-voltage V_d is applied. A value of difference between current into right side and into left side becomes total current. On the right hand side, states below E_L can not transmit. (b) A band diagram under at the bottleneck of one-dimensional ballistic current (see section 3.1.1). In contrast to (a), band minimum of current into left side is same with that into right side at the bottleneck.

References

1) A. Kaneko, A. Yagishita, K. Yahashi, T. Kubota, K. matsuo, I. Mizushima, et al.: IEDM Tech. Dig. 2006, p. 893.

2) S. D. Suk, K. H. Yeo, K. H. Cho, M. Li, Y. Y. Yeoh, et al.: IEEE Trans. Nanotech. **7** (2008), 181.

3) Y. Tian, R. Huang, Y. Wang., J. Zhuge, R. Wang, J. Liu, X. Zhang and Y. Wang: IEDM Tech. Dig., 2007, p. 869.

4) T. Ohno, K. Shiraishi and T. Ogawa: Phys. Rev. Lett. **69** (1992), 2400.

5) T. Vo, A. J. Williamson and G. Galli: Phy. Rev. B **74** (2006), 045116.

6) Jin-An Yan, Li Yang and M. Y. Chou: Phy. Rev. B **76** (2007), 115319.

7) Y. Lee, T. Nagata, K. Kakushima, K. Shiraishi and H. Iwai: ECS Trans. **16** (2009), 1.

8) Y.–J. Ko, M. Shin, S. Lee and K. park, J. Appl. Phys. **89** (2001), 374.

9) Y. Zheng, C. Rivas, R. Lake, K. Alam, T. Boykin and G. Klimeck: IEEE Trans. Electron Devices **52** (2005), 1097.

10) J. Wang, A. Rahman, G. Klimeck and M. Lundstrom: IEDM Tech. Dig., 2005, p. 533.

11) J. Wang, A. Rahman, A. Ghosh, G. Klimeck and M. Lundstrom: IEEE Trans. Electron Devices **52** (2005), 1589

12) M. Luisier, A. Schenk, W. Fichtner and G. Klimeck: Phys. Rev. B **74** (2006), 205323.

13) K. Nehari, N. Cavassilas, J. L. Autran, M. Bescond, D. Munteanu and M. Lannoo: Solid-State Electronics **50** (2006), 716.

14) E. Gnani, A. Gnudi, P. Parruccini, R. Colle, M. Rudan and G. Baccarani: IEEE Trans. Electron Devices **54** (2007), 2243.

15) N. Neophytou, A. Paul, M. S. Lundstrom and G. Klimeck: IEEE Trans. Electron Devices **55** (2008), 1286.

16) E. Ramayya, D. Vasileska, S. M. Goodnick and I. Knezevic: J. Phys. Conference Series **38** (2006), 126.

17) E. Gnani, A. Marchi, S. Reggiani, M. Rudan and G. Baccarani: Solid-State Electronics **50** (2006), 709.

18) K. Natori: IEEE Trans. Electron Devices **55** (2008), 2877.

19) E. Gnani, A. Gnudi, S. Reggiani and G. Baccarani: IEEE Trans. Electron Devices **55** (2008), 2918.

20) K. Natori, Y. Kimura and T. Shimizu: J. Appl. Phys. 97 (2005), 034306.

21) P. Hohenberg and W. Kohn: Phys. Rev. **136** (1964), B864.

22) M. Levy: Proc. Natl. Acad. Sci. (USA) **76** (1979), 6062.

23) W. Kohn and L. J. Sham: Phys. Rev. **140** (1965), A1133.

24) D. M. Ceperley and B. J. Alder: Phys. Rev. Lett. **45** (1980), 566.

25) J. P. Perdew and Y. Wang: Phys. Rev. B 45 (1992), 13244.

26) J. Yamauchi, M. Tsukada, S. Watanabe and O. Sugino: Phys. Rev. B **54** (1996), 5586.

27) A. Oshiyama: Lecture Note of Computational Materials Science Course, Graduate School of Pure and Applied Sciences, University of Tsukuba, Tsukuba, 2004.

28) T. Kurita: Master Thesis, Graduate School of Pure and Applied Sciences, University of Tsukuba, Tsukuba, 2006.

Fundamental constants[1]

Quantity	MKS (SI)
Electron charge (e)	1.60219×10^{-19} coulomb
Electron volt (eV)	1.60219×10^{-19} J·eV^{-1}
Electron mass (m_0)	9.1095×10^{-31} kg
Reduced Planck's constant ($\hbar$)	1.05459×10^{-34} J·s
Reduced Planck's constant ($\hbar$)	6.5822×10^{-16} eV·s
Rydberg (Ry = $\hbar^2/2m_0a_0^2$)	13.6058 eV
Boltzmann's constant (k_B)	1.3807×10^{-23} J·K^{-1}
Boltzmann's constant (k_B)	8.617×10^{-5} eV·K^{-1}

Other constants

Quantity	MKS (SI)
Lattice constant of bulk silicon[2] (Length between silicon and silicon atom)	5.430940×10^{-10} m (2.351666×10^{-10} m)
Length between silicon and H atom of silane (SiH$_4$)[3]	1.4798×10^{-10} m
Atomic unit of length (a.u.)[4]	$5.291772108(18) \times 10^{-11}$ m
SiO$_2$ permittivity (e_{ox})[5]	3.45×10^{-11} F/m
Silicon permittivity (e_{si})[6]	1.04×10^{-10} F/m

[1] E.R. Cohen and B.N. Taylor, Journal of Physical and Chemical Reference Data **2** (4), 663 (1973)
[2] National Astronomical Observatory of Japan: 2003 Chronological Scientific Tables (Maruzen, Tokyo, 2003), p.439 [in Japanese]
[3] National Astronomical Observatory of Japan: 2003 Chronological Scientific Tables (Maruzen, Tokyo, 2003), p.485 [in Japanese]
[4] H. Shull and G. G. Hall, Atomic Units, Nature, volume 184, no. 4698, page 1559 (Nov. 14, 1959)
[5] Y. Tauer and T. H. Ning: *Fundamentals of Modern VLSI Devices* (Cambridge, New York, 1998), p. xi
[6] Y. Tauer and T. H. Ning: *Fundamentals of Modern VLSI Devices* (Cambridge, New York, 1998), p. xi