Compact Modeling of Ballistic Nanowire MOSFETs

Kenji Natori, *Member, IEEE*

(Invited Paper)

Abstract—**Nanowire MOSFETs attract attention due to the probable high performance and the excellent controllability of device current. We present a compact model of ballistic nanowire MOSFET that aids our understanding of physics and the overall properties of the device. The relationship between the gate overdrive and the carrier density is derived and combined with the current expression to yield the current–voltage (I–V) characteristics. The subthreshold characteristics and the short channel effect are also discussed. The effects of the quantum capacitance on device characteristics are analyzed. The low-temperature expression is also derived, and the relation to quantum conductance is discussed. The I–V characteristics are numerically evaluated and examined, employing a reported subband model. The drain- and gate-bias dependences of device current are shown, and the effects of the quantum capacitance and conductance on these characteristics are indicated.**

Index Terms—**Ballistic transport, compact model, nanowire FET, silicon.**

I. INTRODUCTION

RECENTLY, the nanowire transistor has begun to attract a great deal of attention. If the downsizing of the conventional device ever promotes the quasi- or near-ballistic transport of carriers, the potential for performance improvement by further downscaling seems unpromising or, rather, the resultant shorter channel length may lead to the degradation of the off current. As a limiting structure of the advanced Fin-FET approach, the nanowire FET, where a small number of conducting channels within the device can be closely controlled by the gate electrode, may be promising as a next generation device structure. In particular, the carbon nanotube FET seems advantageous [1] due to its large carrier velocity and the smooth channel surface that release from the interface scattering. On the other hand, the advantage for silicon nanowire MOSFET lies in the availability of the many silicon technologies that have accumulated so far. Many theoretical or computational studies of the nanowire MOSFET [2]–[9], as well the analyses of the electronic structure in nanowires, have already been reported [10]–[14]. They include the nonequilibrium Green function approach, the Monte Carlo approach, compact models, etc.
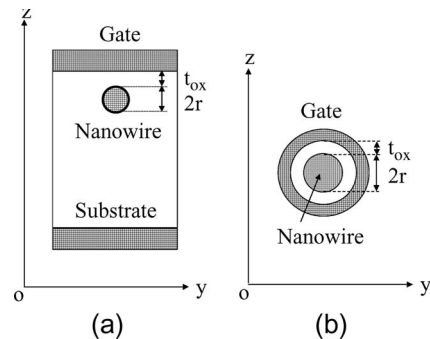
Fig. 1. Cross section perpendicular to the electric current direction of the analyzed nanowire MOSFET. (a) Planar gate structure, where a planar gate is placed close to the nanowire channel and a planar substrate electrode is also placed apart from the wire. (b) Gate-all-around structure without the additional electrodes that cause parasitic capacitance.

However, the physics of the operation of nanowire MOSFETs is not necessarily well discussed in most of the analyses. Reference [4] presents a plain and simple picture of some aspects of the nanowire MOSFET.

In the analysis of nanoscale devices, an approach that first develops the ballistic transport model [15] for the device, and then introduces a limited number of scattering events in the transport model [16], [17], is expected to be fruitful. This paper describes compact modeling of the ballistic operation of nanowire MOSFET and discusses some aspects of the device properties. It is intended to serve as a starting point for analysis of the nanowire MOSFET along the procedure mentioned earlier.

In the second section, the I–V characteristics of the ballistic nanowire MOSFET are derived, and some properties of the device are discussed in an analytical approach. The third section presents a numerical example of the I–V characteristics, as well as other device properties of the ballistic silicon nanowire MOSFET based on the realistic subband parameters. The fourth section is the conclusion.

II. ANALYSIS

A. Electric Current of a Ballistic Nanowire MOSFET

The derivation of current expression has already been discussed elsewhere [1], [2], and in this paper, we consider only the essentials. We assume an n-channel cylindrical nanowire MOSFET with the current directed along the x-axis, and the cross section of the wire is in the y–z plane. The cross section of two types of MOSFET considered in this paper is

schematically shown in Fig. 1. One is the planar gate structure, as shown in Fig. 1(a). The substrate electrode is also placed apart from the wire. The other is the gate-all-around structure shown in Fig. 1(b). The I–V characteristics of the device originate from the potential energy distribution of carriers $U(x, y, z)$. In order to discuss the electronic state in the nanowire, we need to solve the Schrodinger equation with the potential energy $U(x, y, z)$ in the effective mass approximation [18]. The validity of the approach with the use of appropriately defined effective masses of the subband is discussed, and parameters are offered [3], [14]. Within the thin wire region, $U(x, y, z)$ can be approximately decomposed to

$$U(x, y, z) \approx U_C(y, z) + U_L(x) \qquad (1)$$

without serious errors. Here, $U_C(y, z)$ is the confining potential that confines carriers within the wire and is of short range, and the longitudinal potential $U_L(x)$ induces carrier motion along the wire and varies relatively slowly. The Hamiltonian of the wire is then separable into two parts, one that depends on x and another that depends on y and z. The wave function is decomposed into the product of two parts, one that depends on y and z and describes the discrete subband state confined in the wire and another that depends on x and represents the transport along the wire. In our case, the Landauer formalism [19] provides the electric current I_D of the device under the drain bias V_D [1], [2]

$$I_D = \frac{q}{\pi \hbar} \sum_i \int [f(E, \mu_S) - f(E, \mu_D)] \, T_i(E) \, dE \qquad (2)$$

where q and $\hbar$ are the carrier charge and the reduced Planck constant, respectively, $f(E, \mu)$ is the Fermi distribution function

$$f(E, \mu) = \frac{1}{1 + \exp\left(\frac{E - \mu}{k_B T}\right)} \qquad (3)$$

E is energy, and μ_S and $\mu_D = \mu_S - qV_D$, where V_D is the drain bias, are the Fermi levels associated with the source and the drain electrode, respectively. Summation over i shows the summation of contribution from the ith subband. $T_i(E)$ represents the transmission coefficient of the carrier in the ith subband from source to drain and is evaluated with the use of the longitudinal potential $U_L(x)$. k_B and T are the Boltzmann constant and temperature, respectively. The intervalley carrier mixing due to the effective mass potential is neglected due to the slow variation of the potential. The intersubband mixing caused by $U_L(x)$ is also absent due to the decomposition of wave function based on the approximation (1). I_D is then expressed as a sum of independent contributions from each subband as in (2). The source and the drain are assumed to be ideal electrodes that provide sufficient carriers and accept carriers without reflection. $U_L(x)$ typically shows a profile as shown in Fig. 2 and has a maximum point at $x_{\max}$ close to the source edge. The maximum point constitutes the bottleneck of current flow in the device. The electric current in (2) is evaluated, if the subband $E_i(k)$, whose examples are illustrated in Fig. 3, is provided.
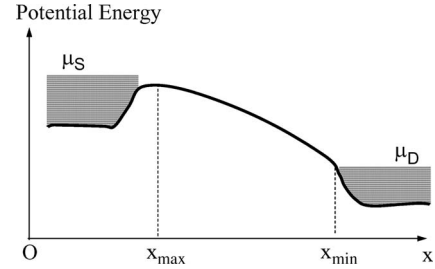

Fig. 2. Rough sketch of the potential energy profile along the channel in the x-direction. μ_S and μ_D are the Fermi levels of the source and drain, respectively. $x_{\max}$ and $x_{\min}$ are the positions where the potential energy takes maximum and minimum values within the channel, respectively.
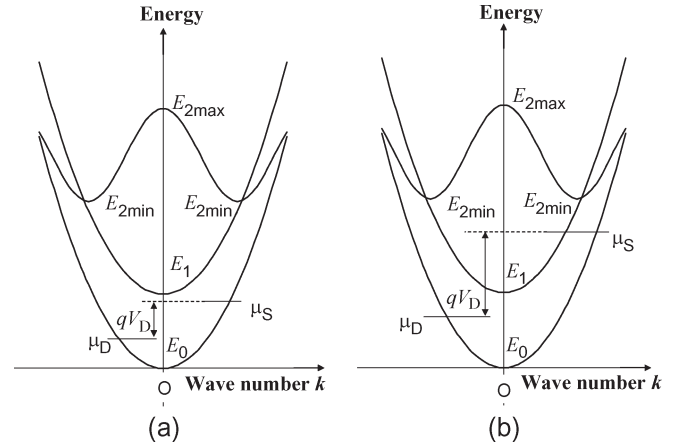

Fig. 3. Example of the E–k relation of the nanowire's subband. (a) Case where the lowest subband is mainly populated. (b) Case where the lowest two subbands are mainly populated.

Here, k denotes the wavenumber in the longitudinal direction, and i is the subband index. Within the vicinity of $x = x_{\max}$, we assume $dU_L(x)/dx \approx 0$, and the carrier state is approximated by the plane wave with the energy $E_i(k)$ correlated with the wavenumber k at the point. When the channel length is not very short, the quantum tunneling can be neglected, and we can assume $T_i(E) = 0$ for the values of E less than the subband minimum at the $x = x_{\max}$ point. Above the subband minimum, we expect that $0 < T_i(E) < 1$. $T_i(E)$ is less than unity due to the carrier reflection caused by the variation of $U_L(x)$ along the wire. However, we can assume that $T_i(E) \approx 1$, if the variation of $U_L(x)$ is sufficiently gentle. On the other hand, we need to be cautious when the subband has a finite width between the maximum and minimum inside the Brilouin zone, as in the highest subband in Fig. 3. The window of the allowed energy value of a carrier extends from the minimum to the maximum, and this window shifts along the wire as the potential energy varies from the source to the drain. As the carrier runs along the channel with the incident energy, the carrier is immediately reflected back to the source once the carrier energy gets outside of this changing window. Therefore, $T_i(E) \approx 1$ holds good only for the values of E that remain within the window at every point of the channel. We can conclude that $T_i(E) \approx 1$ within the energy region between the subband minimum at $x = x_{\max}$ and the subband maximum at the drain edge, $x = x_{\min}$ in Fig. 2, and that $T_i(E) = 0$

otherwise. Thus, (2) is evaluated with the use of (3) [1], such that we have (4), shown at the bottom of the page, where $G_0 \equiv 2q^2/h = 77.8$ μS is the quantum conductance, holds. The degeneracy of the subband is explicitly considered, and g_i $(i = 0, 1, \ldots)$ is the degeneracy of the ith subband. $E_{i\min}(x_{\max})(E_{i\max}(x_{\min}))$ in the summation $dE_i(k)/dk \geq 0$ shows the minimum (maximum) energy within the piecewise part of the ith subband that satisfies $dE_i(k)/dk \geq 0$ when evaluated at $x = x_{\max}$ $(x = x_{\min})$. The similar expression in the summation $dE_i(k)/dk \leq 0$ denotes the similar quantity. The finite bandwidth due to the presence of $E_{i\min}$ and $E_{i\max}$ may have a serious effect on the current–voltage characteristics, depending on the location of μ_S and μ_D. The increase of drain bias causes narrowing of the conduction window in the band diagram and yields a current decrease, leading to the negative differential conductance. Such an effect is well analyzed in [8], and Fig. 7 therein shows some specific dips in I–V_D characteristics.

When the maximum energy of the subband is much larger than both of μ_S and μ_D, just as in the lowest two subbands in Fig. 3, $E_{i\max}$ is effectively regarded as infinity, and (4) is reduced to a simple form [2]

$$I_D = G_0 \left(\frac{k_B T}{q} \right) \sum_i g_i \ln \left\{ \frac{1 + \exp\left[(\mu_S - E_{i0})/k_B T\right]}{1 + \exp\left[(\mu_D - E_{i0})/k_B T\right]} \right\}. \tag{5}$$

Here, E_{i0} stands for the minimum of the ith subband at $x_{\max}$. The case represented by (5) is analyzed in the following.

B. Electrostatics

For the evaluation of the device current I_D in (5), we need the value of $(\mu_S - E_{io})$. The value is evaluated by the electrostatics controlled by the gate bias and the channel carrier charge. We assume an intrinsic channel, where the channel charge consists only of the induced carriers. The electrostatics of a wire MOSFET is basically characterized by the 3-D Poisson equation. However, we see that $dU_L(x)/dx \approx 0$ around $x = x_{\max}$, and the longitudinal electric field almost vanishes. The Poisson equation in the cross section of the device at this point is reduced to a Laplace equation within the insulator region, which is rewritten as

$$\frac{\partial^2 U(x,y,z)}{\partial y^2} + \frac{\partial^2 U(x,y,z)}{\partial z^2} = -\frac{\partial^2 U(x,y,z)}{\partial x^2}. \tag{6}$$

The variation of $U(x,y,z)$ in the x-direction along the long wire is gentle compared to that in the $y-z$ plane defining the thin nanowire structure, and thus, the right-hand side of

(6) is much smaller compared to each term in the left-hand side of the equation. If we neglect the right-hand side as $\partial^2 U(x,y,z)/\partial x^2 \approx 0$, $U(x,y,z)$ becomes uniform along x-axis around there. The x-dependence is neglected, and the 2-D Laplace equation in the $y-z$ plane results. In the analysis of electrostatics in the cross section of devices in Fig. 1 with the boundary condition that the electrodes and the wire part are conductors, this equation yields a solution described by the channel-gate capacitance C_G and also the parasitic capacitance, e.g., the channel-substrate capacitance C_p, both in per unit wire length. The situation is similar to the gradual channel approximation. The gate capacitances of these structures are evaluated by approximating the nanowire channel with the cylindrical conductor. The insulator of SiO$_2$ is assumed. For the planar gate structure in Fig. 1(a), we have [1]

$$C_G = \frac{2\pi \varepsilon_{\mathrm{ox}}}{\ln \left\{ \frac{\sqrt{2r+t_{\mathrm{ox}}}+\sqrt{t_{\mathrm{ox}}}}{\sqrt{2r+t_{\mathrm{ox}}}-\sqrt{t_{\mathrm{ox}}}} \right\}}. \tag{7}$$

The value of C_p is also given by a similar expression. In the gate-all-around structure in Fig. 1(b), the wire channel is covered by the surrounding gate electrode, and therefore, the parasitic capacitance vanishes. The gate capacitance is given by

$$C_G = \frac{2\pi \varepsilon_{\mathrm{ox}}}{\ln \left(\frac{r+t_{\mathrm{ox}}}{r} \right)}. \tag{8}$$

Here, $\varepsilon_{\mathrm{ox}}$ is the dielectric constant of SiO$_2$, r is the radius of the cylindrical wire, and t_{ox} is the thickness of the insulator. Note that the estimation of the EOT for high-k insulators is a little complicated in contrast to the normal MOSFET.

The potential energy distribution in the cross section $x = x_{\max}$ of an n-channel device is schematically shown in Fig. 4. The energy separation between various levels and the barrier heights are defined, as shown in the figure. As for the nanowire channel, the positive velocity carriers with the charge Q_f and running toward the drain (denoted as the Fore channel) and the negative velocity carriers with the charge Q_b and running toward the source (denoted as the Back channel) are separately demonstrated. Fore channel carriers are fed from the source and are controlled by μ_S, while Back channel carriers are fed from the drain and are controlled by μ_D. μ_0 is the band bottom of the lowest subband $(= E_{00})$. The total charge density Q in the wire is stored in parts, in C_G and C_p, respectively. When the positive V_G and negative V_{sub} are applied with respect to the source level μ_S

$$Q = Q_f + Q_b = -C_G \phi_G + C_p \phi_p. \tag{9}$$

$$I_D = G_0 \left(\frac{k_B T}{q} \right) \sum_i g_i \left(\sum_{dE_i(k)/dk \geq 0 \, \mathrm{part}} \ln \left\{ \frac{1 + \exp\left[(\mu_S - E_{i\min}(x_{\max}))/k_B T\right]}{1 + \exp\left[(\mu_S - E_{i\max}(x_{\min}))/k_B T\right]} \right\} \right.$$

$$\left. - \sum_{dE_i(k)/dk \leq 0 \, \mathrm{part}} \ln \left\{ \frac{1 + \exp\left[(\mu_D - E_{i\min}(x_{\max}))/k_B T\right]}{1 + \exp\left[(\mu_D - E_{i\max}(x_{\min}))/k_B T\right]} \right\} \right) \tag{4}$$
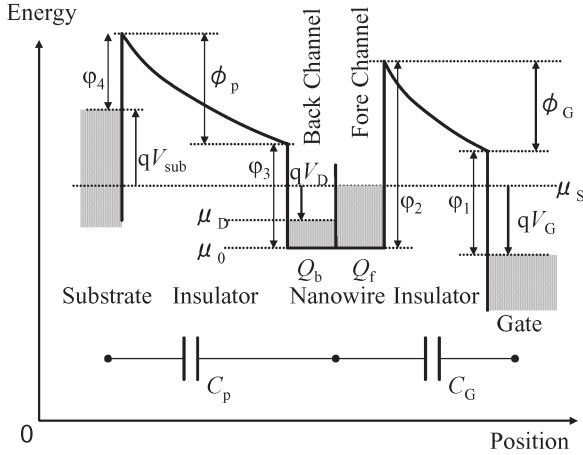
Fig. 4. Potential energy diagram associated with the nanowire channel and the nearby electrodes. μ_S and μ_D are the Fermi levels of the source and drain, respectively, and μ_0 is the subband bottom in the channel. The fore and back channels include carriers with the positive velocity and regulated by the source Fermi level and with the negative velocity and regulated by the drain Fermi level, respectively.

ϕ_G and ϕ_p are the bias voltages applied across the gate and substrate insulators, respectively. According to Fig. 4, ϕ_G and ϕ_p are correlated to V_G and V_{sub}, respectively, as

$$q\phi_G = qV_G - (\mu_S - \mu_0) - (\varphi_1 - \varphi_2) \qquad (10)$$

$$q\phi_p = -qV_{sub} + (\mu_S - \mu_0) - (\varphi_3 - \varphi_4). \qquad (11)$$

Here, φ_1 and φ_4 are the barrier heights from the gate and the substrate Fermi level to the conduction band of the insulators, respectively, and φ_2 and φ_3 are similar energy barriers from the subband bottom of the nanowire to the conduction band of the insulators, as shown in Fig. 4. Substituting (10) and (11) into (9), we obtain [1] the following (a similar expression is discussed in [4])

$$(V_G - V_t) - \alpha\frac{\mu_s - \mu_0}{q} = \frac{|Q|}{C_G} \qquad (12)$$

in view of $Q \leq 0$. Here

$$V_t \equiv \frac{\varphi_1 - \varphi_2}{q} - \frac{C_p}{C_G}\left(V_{sub} + \frac{\varphi_3 - \varphi_4}{q}\right) \qquad (13)$$

$$\alpha = 1 + \frac{C_p}{C_G}. \qquad (14)$$

On the other hand, $|Q|$ is specifically evaluated, considering the carrier distribution within the subband of Fig. 3 at $x = x_{max}$. Since the density of states in the k-space is $1/\pi$, including the spin degeneracy, and carriers are distributed according to (3), we have

$$|Q| = \frac{q}{\pi}\sum_i g_i \left[\int_{k_{i\,min}}^{\infty} \frac{dk}{1 + \exp\left\{\frac{E_i(k) - \mu_S}{k_B T}\right\}} \right.$$
$$\left. + \int_{-\infty}^{k_{i\,min}} \frac{dk}{1 + \exp\left\{\frac{E_i(k) - \mu_D}{k_B T}\right\}}\right]. \qquad (15)$$

Every $E_i(k)$ is assumed to have its minimum at $k_{i\,min}$ and $dE_i(k)/dk \leq 0$ for $k < k_{i\,min}$ ($dE_i(k)/dk \geq 0$ for $k > k_{i\,min}$.) The former (latter) integral in the right-hand side of (15) corresponds to the Q_f (Q_b) charge.

C. Compact Modeling of I–V Characteristics

At the gate voltage $V_G = V_t$, (12) suggests that $\mu_S \leq \mu_0$. At $T = 0$, we actually have $\mu_S = \mu_0$ and $Q = 0$. For $T \neq 0$, we have $\mu_S < \mu_0$ and Q is very small, because (15) implies that $|Q|$ is approximately proportional to $\exp[-(\mu_0 - \mu_S)/k_B T]$. In the region of $V_G > V_t$, $|Q|$ increases, corresponding to the increase of V_G, and we can say that V_t is the effective threshold voltage of the device. According to (13), the threshold voltage V_t depends on the substrate bias V_{sub}, and a deep negative V_{sub} yields a proportionate increase in V_t. V_t also depends on φ_2 and φ_3, which are the energy barriers measured from the subband bottom. As the nanowire thins, the confinement energy of the carrier increases, and these barrier heights are reduced, which, in turn, increases the value of V_t. In MOSFETs with very thin nanowire, the control of the threshold voltage may constitute a serious problem.

If the subband structure $E_i(k)$ and the threshold voltage V_t are given, and the external biases V_G and V_D are identified, we can solve the coupled equation of (15) and (12) and evaluate the value μ_S measured from the energy level μ_0 ($= E_{00}$). We can then evaluate the device current I_D by substituting the value into (5). This procedure constitutes a compact model for the evaluation of the I–V characteristics of a nanowire MOSFET. The short channel effect is discussed later.

Let us examine the subthreshold characteristics. Suppose that the device is in the so-called "weak inversion" region with nondegenerate carriers, and we assume that $\mu_S < \mu_0$. Fermi statistics are replaced by Boltzmann statistics, and we can assume that $\exp[(\mu - E_{i0})k_B T] \ll 1$ in either case of $\mu = \mu_S$ and $\mu = \mu_D$. We can then expand the logarithmic function in (5) to obtain an approximate expression

$$I_D \cong G_0\left(\frac{k_B T}{q}\right)\exp\left(\frac{\mu_S - \mu_0}{k_B T}\right)\sum_i g_i$$
$$\times\left[\exp\left(\frac{E_{00} - E_{i0}}{k_B T}\right) - \exp\left(\frac{E_{00} - E_{i0} - qV_D}{k_B T}\right)\right]. \qquad (16)$$

On the other hand, the charge density $|Q|$ in (12) for the case $V_G < V_t$ is smaller than that for the case of $V_G > V_t$ by some orders of magnitude, and the right-hand side of (12) is negligibly small compared to each term in the left-hand side of (12). Thus, we have approximately

$$(V_G - V_T) - \alpha\frac{\mu_s - \mu_0}{q} \cong 0. \qquad (17)$$

We then derive the subthreshold index S from (16) and (17)

$$S = \left[\frac{d\log I_D}{dV_G}\right]^{-1} = \left[\frac{1}{k_B T \ln 10}\frac{d(\mu_s - \mu_0)}{dV_G}\right]^{-1}$$
$$= \left[\frac{q}{k_B T \ln 10}\frac{C_G}{C_G + C_P}\right]^{-1}. \qquad (18)$$

If the parasitic capacitance C_p can be neglected, (18) is reduced to the ideal S value of 58 meV/dec at 300 K.

D. Short Channel Effect

The right-hand side of (6) is small but has a nonzero value in real cases. The potential profile in Fig. 2 implies that $\partial^2 U/\partial x^2 < 0$ in the vicinity of $x = x_{\max}$. The negative $\partial^2 U/\partial x^2$ can be regarded as equivalent to the presence of effectively positive charge within the insulator in the 2-D Poisson equation (6) associated with the cross section at $x = x_{\max}$. This effective charge yields an increase in magnitude of the channel charge $|Q|$ to a value $|Q'| = |Q| + \Delta Q$, where ΔQ is positive, when compared to the case that the small right-hand side of (6) was neglected. The magnitude of the effective charge within the insulator is approximately given by $\varepsilon_{\text{ox}} \int (\partial^2 U/\partial x^2)\, dy\, dz$, where the integration is over the cross section at $x = x_{\max}$ within the insulator. The effective charge is thought to play a similar role as the oxide charge in the ordinary MOS. If it is concentrated in the region close to the surface of the semiconductor, the aforementioned ΔQ is roughly equal to the magnitude. However, in general, ΔQ is smaller due to sharing with the gate charge. In any case, (12) is modified by ΔQ to

$$\left\{ V_G - \left(V_t - \frac{\Delta Q}{C_G} \right) \right\} - \alpha \frac{\mu_S - \mu_0}{q} = \frac{|Q'|}{C_G} \qquad (19)$$

where $|Q'|$ has the same expression as $|Q|$ in the right-hand side of (15). The modified (19) suggests that the longitudinal potential variation due to the source/drain field causes a reduction of the effective threshold voltage by the amount $\Delta Q/C_G$. This is the V_t shift due to the short channel effect. If the value of ΔQ is identified by some way or another, the coupled equation of (19) and (15), where Q is replaced by Q', allows the evaluation of the renormalized value of $(\mu_S - \mu_0)$ in consideration of the short channel effect, and the electric current is yielded by (5). In the weak inversion region, where the carrier charge within the channel is very small, the potential profile $U_L(x)$ in the channel is controlled by μ_S, μ_D, and V_G. As V_G is increased, the dominance of the gate field over the potential energy profile in the channel strengthens and yields a decrease in the magnitude of $|\partial^2 U/\partial x^2|$ around the point $x = x_{\max}$, which, in turn, causes a decrease in the effective charge ΔQ. The replacement of (12) by (19) induces modification of subthreshold index S to

$$S = \left[\frac{q}{k_B T \ln 10} \frac{C_G}{C_G + C_p} \left(1 + \frac{1}{C_G} \frac{d\Delta Q}{dV_G} \right) \right]^{-1}. \qquad (20)$$

Since $d\Delta Q/dV_G < 0$ is inferred from the discussion earlier, we can see that the short channel effect degrades the value of the subthreshold index S, even if the channel is surrounded by gate electrode and the parasitic capacitance is neglected.

E. Quantum Capacitance

In contrast to the ordinary 2-D MOSFETs, the 1-D wire MOSFETs have a small density of states for channel carriers. The resultant small quantum capacitance [20] due to the density of states may have a serious influence on the device properties. We discuss herein the relation between the quantum capacitance and the I–V characteristics. For simplicity, we assume the case where the parasitic C_p is absent ($\alpha = 1$) and $V_D = 0$ ($\mu_S = \mu_D \equiv \mu$). Then, (15) is reduced to

$$|Q| = \frac{q}{\pi} \sum_i g_i \int_{-\infty}^{\infty} \frac{dk}{1 + \exp\left\{ \frac{E_i(k) - \mu}{k_B T} \right\}}$$

$$= q \sum_i g_i \int_{-\infty}^{\infty} \{ D_{i+}(E - \mu_0) + D_{i-}(E - \mu_0) \}$$

$$\times f(E - \mu_0, \mu - \mu_0)\, dE \qquad (21)$$

where $D_{i\pm}(E - \mu)$ denotes the densities of states of the positive and negative velocity branches of the ith subband. $f(E - \mu_0, \mu - \mu_0)$ is the Fermi distribution function, where the energy E and the Fermi potential μ are explicitly shown to be referred from the energy level μ_0. On the other hand, the differentiation of (12) by V_G yields

$$1 - \frac{1}{q} \frac{d(\mu - \mu_0)}{dV_G} = \frac{1}{C_G} \frac{d|Q|}{dV_G}. \qquad (22)$$

In view of (3), the differentiation of the Fermi distribution function by V_G is

$$\frac{df(E - \mu_0, \mu - \mu_0)}{d(\mu - \mu_0)} \frac{d(\mu - \mu_0)}{dV_G}$$

$$= \left\{ -\frac{df(E - \mu_0, \mu - \mu_0)}{d(E - \mu_0)} \right\} \frac{d(\mu - \mu_0)}{dV_G}. \qquad (23)$$

With the use of (23), as well as the property of Fermi function (3), the differentiation of $|Q|$ in (21) results in

$$\frac{d|Q|}{dV_G} \cong q \sum_i g_i \left\{ D_{i+}(\mu - \mu_0) + D_{i-}(\mu - \mu_0) \right\}$$

$$\times \left\{ -\int_{-\infty}^{\infty} \frac{df(E - \mu_0, \mu - \mu_0)}{d(E - \mu_0)} dE \right\} \frac{d(\mu - \mu_0)}{dV_G}$$

$$= q \sum_i g_i \{ D_{i+}(\mu - \mu_0) + D_{i-}(\mu - \mu_0) \} \frac{d(\mu - \mu_0)}{dV_G}. \qquad (24)$$

Denoting the quantum capacitance C_Q as

$$C_Q = q^2 \sum_i g_i \left\{ D_{i+}(\mu - \mu_0) + D_{i-}(\mu - \mu_0) \right\}. \qquad (25)$$

(22) and (24) yield that

$$\frac{d|Q|}{dV_G} = \left[\frac{1}{C_Q} + \frac{1}{C_G} \right]^{-1}. \qquad (26)$$

We can see that the term $\alpha(\mu_S - \mu_0)/q$ in (12) represents the contribution from the quantum capacitance. Equation (12) suggests that, when the gate bias increases, the bias increase applied to the gate insulator is reduced by an amount equal to the increase in the electrochemical potential caused by the

increase in carriers. The effect is also represented by a series connection of capacitances C_G and C_Q. In order to examine the effect of quantum capacitance on the calculated result, you can simply compare the ordinary result with a realistic value of α to the expedient case with $\alpha = 0$. Notice that we also have another type of quantum capacitance than C_Q. Due to the quantum effect, the probability density of carriers within the wire tends to be depleted off the wire surface and to be concentrated in the core region. This effect also induces a sort of capacitance that is serially connected to the electrostatic capacitance C_G [20]. The inversion layer capacitance in the ordinary MOSFET is of this type. This type of quantum capacitance is not included in the present compact model described in (5), (7) or (8), (12), and (15).

F. Low-Temperature Limits

As will be shown in the following, the effect of temperature on the current magnitude was not so remarkable. A careful examination of the case $T = 0$ is helpful for the assessment of the overall behavior and the rough magnitude of I–V characteristics. For simplicity, we assume that the subband energy is expressed in the form

$$E_i(k) = E_{i0} + \frac{\hbar^2 k^2}{2m_i} \tag{27}$$

with the parabolic effective mass m_i and $k_{i\,\min} = 0$. The Fermi distribution function is reduced to the step function for $T = 0$, and the carrier density in (15) is explicitly evaluated as

$$|Q| = \frac{\sqrt{2}q}{\pi\hbar} \sum_i g_i \Big[\sqrt{m_i(\mu_S - E_{i0})} H(\mu_S - E_{i0}) $$
$$+ \sqrt{m_i(\mu_S - qV_D - E_{i0})} H(\mu_S - qV_D - E_{i0}) \Big] \tag{28}$$

and the current expression in (5) is reduced to

$$I_D = \frac{G_0}{q} \sum_i g_i \left[(\mu_S - E_{i0})H(\mu_S - E_{i0}) \right.$$
$$\left. - (\mu_S - E_{i0} - qV_D)H(\mu_S - E_{i0} - qV_D) \right]. \tag{29}$$

Here, $H(E)$ is the Heaviside function; $H(E) = 1$ for $E > 0$, and $H(E) = 0$ otherwise. The specific current expression is derived by solving (12), (28), and (29). In the case that μ_S is within the lowest subband, as in Fig. 3(a), these equations are analytically solved to give

$$I_D = g_0 G_0 \left[V_D H(V_D^* - V_D) + V_0^* H(V_D - V_D^*) \right]$$
$$\text{for } E_{10} - E_{00} > qV_D^* \tag{30}$$

where

$$V_D^* \equiv \left(\frac{g_0 A}{\alpha} \right)^2 \left[\sqrt{1 + \frac{\alpha(V_G - V_t)}{(g_0 A)^2}} - 1 \right]^2 \tag{31}$$
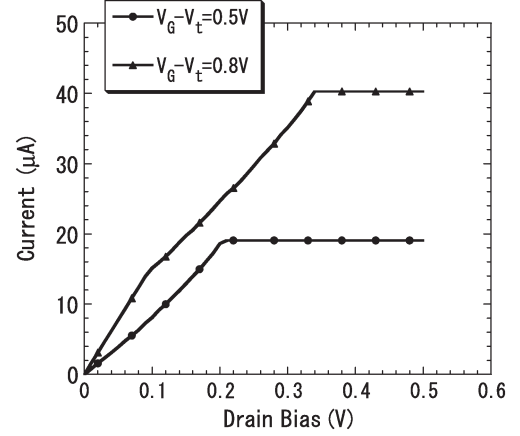
$$A = \frac{\sqrt{2m_0 q}q}{h C_G}. \tag{32}$$

Fig. 5. I–V_D characteristics for a tentative simple subband structure at $T = 0$. Parameters are $m_0 = m_1 = 0.19$ m, $g_0 = g_1 = 1$, $E_{00} = 0$, $E_{10} = 0.16$ eV, and $C_G = 2.5$ pF/cm. The solid circle (triangle) curve shows the case with the source Fermi level arrangement in Fig. 3(a) and (b).

The I–V characteristics for the two tentative cases of carrier population are shown in Fig. 5. The curve with the gate overdrive $V_G - V_t = 0.5$ V ($V_G - V_t = 0.8$ V) shows the case in which only the lowest subband contributes (the lowest two subbands contribute) to the current, as in Fig. 3(a) [Fig. 3(b)]. The assumed parameters are $m_0 = m_1 = 0.19$ m, where m is the electron mass, $g_0 = g_1 = 1$, $E_{00} = 0$, $E_{10} = 0.16$ eV, and $C_G = 2.5$ pF/cm. The I–V characteristics show kinks when μ_S and μ_D cross the subband minimum as the applied bias is increased. For the case that carriers are distributed in the lowest n-subbands, we have

$$I_D = \left(\sum_i g_i \right) G_0 V_D \tag{33}$$

for $V_D \to 0$.

III. NUMERICAL RESULTS

The compact model allows the computation of the I–V characteristics of ballistic nanowire MOSFETs, if the E–k relation of the subband $E_i(k)$ is provided. In a thick wire structure with a comparatively large diameter, the effective mass approximation of the bulk silicon is thought to be valid. In a nanowire MOSFET extended in the $\langle 100 \rangle$ direction, for example, the lowest subband is quadruply degenerated and has an effective mass of $0.19m$. As the wire thins and the cross section falls below several square nanometers, the energy separations between the lowest subband and the upper subbands increase. The influence of the higher subband, with only a small population of carriers, on the transport properties of the wire becomes increasingly small. The tight binding approach and the density functional theory calculations are applied to these ultrathin nanowires, and the subband structures are reported [11], [13], [14]. These calculations give a larger effective mass of $\sim 0.3m$ compared to that in the ordinary effective mass approximation and show a tendency of further increase in the downscaled geometry. The degeneracy of the lowest subband also tends to be removed. In the present analysis, we try to employ the subband parameter of the 7×7-atom [110] square
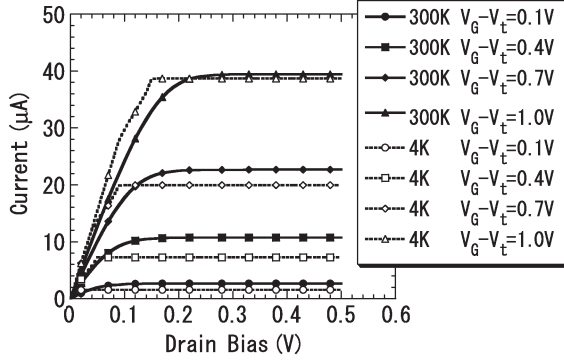
Fig. 6. I–V_D characteristics of the Si nanowire MOSFET with the realistic subband structure in (34) and (35).
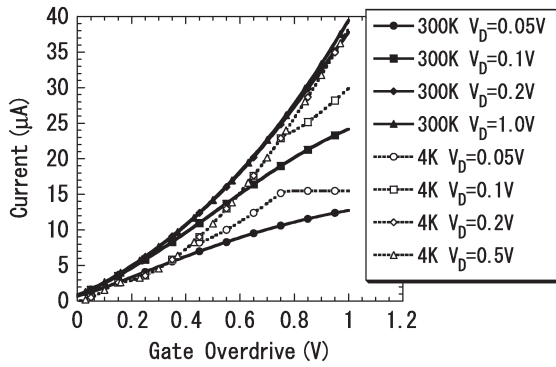

Fig. 7. I–V_G characteristics of the Si nanowire MOSFET with the realistic subband structure in (34) and (35).

Fig. 8. I–V_G characteristics of the Si nanowire MOSFET in logarithmic scale. The slope at the low gate overdrive reproduces the ideal subthreshold index of $S = 58$ meV/dec.

Fig. 9. Effect of quantum capacitance on I–V_D characteristics. Curves with $\alpha = 0$ ($\alpha = 1$) show a case in which the quantum capacitance is neglected (counted).

nanowire extended in the $\langle 001 \rangle$ direction, derived by the density functional calculation and reported by Gnani's group [14], as shown in the following. Denoting that $\gamma_0 \equiv (\hbar k)^2 / 2\,$m in electronvolts, the lowest subband without degeneracy described by the effective mass $0.31m$ is effectively expressed as

$$E_0(k) = \left(\frac{\gamma_0}{0.31}\right) - 0.3\left(\frac{\gamma_0}{0.31}\right)^2 + 0.0375\left(\frac{\gamma_0}{0.31}\right)^3. \quad (34)$$

The threefold-degenerate excited subband with the effective mass of $0.36m$ is expressed as

$$E_1(k) = 0.0327 + \left(\frac{\gamma_0}{0.36}\right) - 0.258\left(\frac{\gamma_0}{0.36}\right)^2 + 0.025\left(\frac{\gamma_0}{0.36}\right)^3. \quad (35)$$

The original nanowire has a square cross section with a 1.34-nm side, but we try to employ an approximate gate capacitance value evaluated by (8) for the cylindrical nanowire with the same cross-sectional area. We assume SiO$_2$ for the gate insulator and $t_{ox} = 1$ nm. We then have $C_G = 2.57$ pF/cm and $C_p = 0$.

Fig. 6 shows the I–V_D characteristics evaluated for a room temperature of 300 K, as well as for a low temperature of 4 K. The parameter is the gate overdrive. We see that the room-temperature I–V_D characteristics are similar to those of an ordinary MOSFET, and the electric current magnitude is 20 to 40 μA in this bias range. At low temperatures, the curve shows kink structures when the Fermi level crosses the subband minimum and displays a similar magnitude to that of
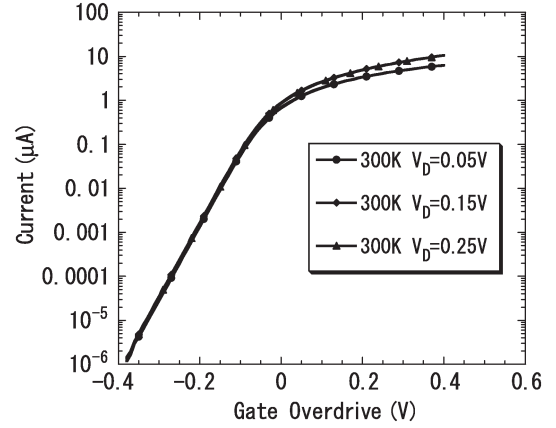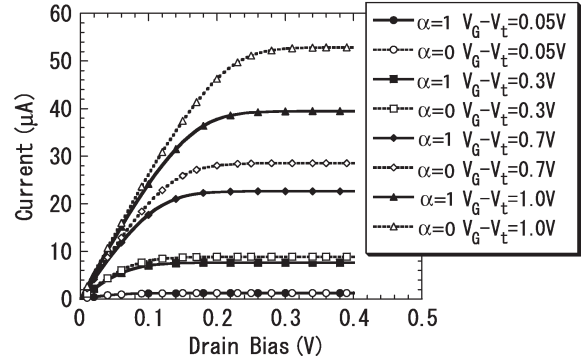
room-temperature curves. Fig. 7 shows the I–V_G characteristics with the drain bias as the parameter. The low-temperature curve shows characteristic kinks or branching behaviors due to the same origin as that in the I–V_D characteristics. Fig. 8 shows the I–V_G characteristics plotted in a logarithmic scale. The normal subthreshold characteristics are reproduced. The subthreshold parameter S is 58 meV/dec in accordance with the prediction (18) with $C_p = 0$. Fig. 9 shows the effect of quantum capacitance C_Q on the I–V_D characteristics. The curves with $\alpha = 1$ ($\alpha = 0$) show the cases where the effect of C_Q is considered (neglected). We can see that the effect is considerable in the current-saturation region, and the neglect of the effect brings about a serious overestimation of the current value. Fig. 10 shows how the conductance dI_D/dV_D varies as a function of the gate overdrive. At room temperature, the conductance is an increasing monotonous function of the gate overdrive. However, at the low temperature, it shows a characteristic behavior reflecting the subband structure. The curve for $T = 4$ K and $V_D = 0$ V shows a step structure reflecting the quantized conductance of the 1-D subband, as suggested by (33). The lower step of $G_0 = 77.8$ μS represents the quantized conductance of the single lowest subband in Fig. 3. As the gate overdrive increases and the Fermi level goes into the threefold second-lowest subband, the conductance goes up to the value of total $4G_0. = 310$ μS. The characteristic structure
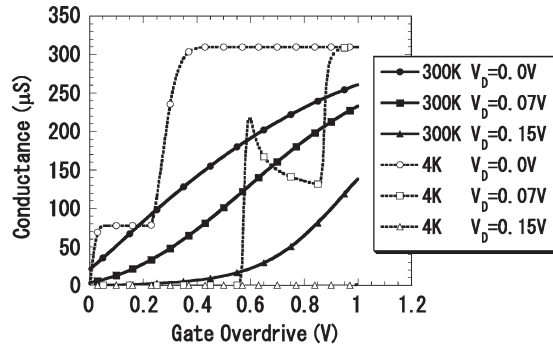
Fig. 10.   Drain conductance variation as a function of the gate overdrive. The low-temperature curve with characteristic step structures changes to the broad increasing curve at room temperature.
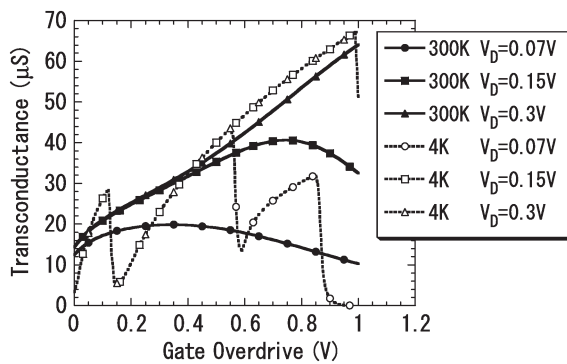

Fig. 11.   Transconductance variation as a function of the gate overdrive. Peak structure at low temperature caused by a variation in the quantum capacitance is rolled at room temperature.
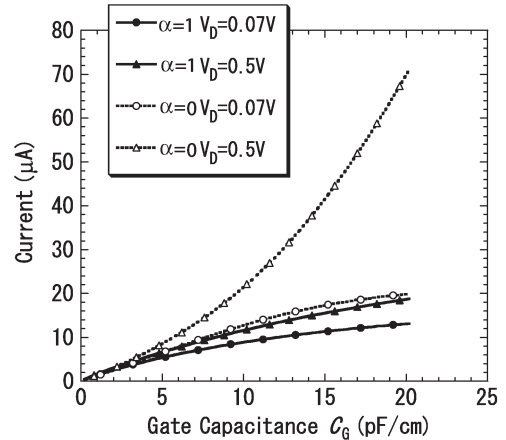

Fig. 12.   Drain current variation as a function of the gate capacitance. Significant overestimation of saturation current results at $V_D = 0.5$ V, when the quantum capacitance is neglected. Dominance of the quantum capacitance masks the effects of the gate capacitance increase.

rather than by the μ_S dependence, and the current tends to be independent of the gate capacitance.

IV. CONCLUSION

A compact model of I–V characteristics for a ballistic nanowire MOSFET is proposed, and the device property is discussed. A new type of electrostatic relationship that connects the gate overdrive to the carrier density at the current bottleneck point is derived, and it is combined with the current expression to yield a closed current formula of a ballistic nanowire MOSFET. The result is applied to a silicon nanowire MOSFET with the use of the nanowire subband parameter reported by a group at the University of Bologna. The electric current level amounts to a few tens of microamperes for the applied bias of several tenths of a volt. Roughly speaking, the magnitude weakly depends on the temperature, and the conductance is on the order of magnitude of the quantum conductance. In particular, the conductance in the linear region at low temperatures is dominated by the quantum conductance. The subthreshold characteristics are analyzed, and the gate-all-around structure is shown to yield the ideal S factor of around 60 meV/dec when the short channel effect is neglected. The planar gate structure includes the parasitic capacitance associated with the substrate, and the value of the S factor will degrade. The contribution of the quantum capacitance (due to the density of states) in the current formula is identified, and the critical importance of quantum capacitance in the nanowire MOSFET operation is pointed out. Neglect of the capacitance component causes a serious overestimation of the saturation current. The current magnitude of the nanowire MOSFET is primarily controlled by the quantum capacitance when the gate capacitance is large, and the increase of the gate capacitance in such a region is ineffective for the improvement of the device performance.

of the curve $V_D = 0.07$ V is also caused by the Fermi level arrangement associated with Fig. 3(b). The zero level of the curve $V_D = 0.15$ V is due to the current saturation. Fig. 11 shows a similar plot of the transconductance dI_D/dV_G as a function of the gate overdrive. The low-temperature curves show discontinuous decreases here and there, and this is also due to the fact that the Fermi level increases across subband bottoms. As we see in (5), $dI_D/dV_G \propto (1/q)d\mu/dV_G$, but (22) and (26) yield that $(1/q)d\mu/dV_G = C_G/(C_G + C_Q)$. Therefore, a sudden increase in C_Q due to the jump-over of Fermi level to a new subband causes a sudden decrease of the transconductance. The discontinuity is due to the similar effect as in Fig. 10, but the present case is not directly related to the quantum conductance. The transconductance at room temperature is basically understood as the smoothing out of the low-temperature bumps. Fig. 12 shows the current variation as a function of the gate capacitance C_G. The gate and quantum capacitances are connected in series. In a region where C_G is sufficiently large, therefore, the current is principally controlled by the quantum capacitance, and the gate capacitance affects only a little, as suggested by the curves for $\alpha = 1$. If we neglect the quantum capacitance (as in the curves with $\alpha = 0$), the current in the saturation region shows a significant increase in response to the C_G increase. However, the current in the linear region shows only a weak dependence on C_G, as suggested by the curve $\alpha = 0$ and $V_D = 0.07$ V. As we can infer from (33), I_D in the small V_D region is controlled by G_0 and V_D

REFERENCES

[1] K. Natori, Y. Kimura, and T. Shimizu, "Characteristics of a carbon nanotube field-effect transistor analyzed as a ballistic nanowire field-effect transistor," *J. Appl. Phys.*, vol. 97, no. 3, pp. 034 306-1–034 306-7, Jan. 2005.

[2] D. Jimènez, J. J. Sáenz, B. Iniquez, J. Suné, L. F. Marsal, and J. Parllarès, "Unified compact model for the ballistic quantum wire and quantum well metal–oxide–semiconductor field-effect-transistor," *J. Appl. Phys.*, vol. 94, no. 2, pp. 1061–1068, Jul. 2003.

[3] J. Wang, A. Rahman, A. Ghosh, G. Klimeck, and M. Lundstrom, "On the validity of the parabolic effective-mass approximation for the *I–V* calculation of silicon nanowire transistors," *IEEE Trans. Electron Devices*, vol. 52, no. 7, pp. 1589–1595, Jul. 2005.

[4] M. Lundstrom and J. Guo, *Nanoscale Transistors: Device Physics, Modeling and Simulation*. New York: Springer-Verlag, 2006, ch. 5.

[5] B. C. Paul, R. Tu, S. Fujita, M. Okajima, T. H. Lee, and Y. Nishi, "An analytical compact circuit model for nanowire FET," *IEEE Trans. Electron Devices*, vol. 54, no. 7, pp. 1637–1644, Jul. 2007.

[6] A. Godoy, F. Ruiz, C. Sampedro, F. Gámiz, and U. Ravaioli, "Calculation of the phonon-limited mobility in silicon gate-all-around MOSFETs," *Solid State Electron.*, vol. 51, no. 9, pp. 1211–1215, Sep. 2007.

[7] S. Jin, M. V. Fischetti, and T-W. Tang, "Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity," *J. Appl. Phys.*, vol. 102, no. 8, p. 083 715-1, Oct. 2007.

[8] E. Gnani, A. Gnudi, S. Reggiani, M. Rudan, and G. Baccarani, "Band structure effects on the current–voltage characteristics of SNW-FETs," in *IEDM Tech. Dig.*, 2007, pp. 737–740.

[9] A. Pecchia, L. Salamandra, L. Latessa, B. Aradi, T. Frauenheim, and A. D. Carlo, "Atomistic modeling of gate-all-around Si-nanowire field-effect transistors," *IEEE Trans. Electron Devices*, vol. 54, no. 12, pp. 3159–3167, Dec. 2007.

[10] H. Sakaki, "Scattering suppression and high-mobility effect of size-quantized electrons in ultrafine semiconductor wire structures," *Jpn. J. Appl. Phys.*, vol. 19, no. 12, pp. L735–L738, Dec. 1980.

[11] T. Ohno, K. Shiraishi, and T. Ogawa, "Intrinsic origin of visible light emission from silicon quantum wires: Electronic structure and geometrically restricted exciton," *Phys. Rev. Lett.*, vol. 69, no. 16, pp. 2400–2403, Oct. 1992.

[12] R. Kotlyar, B. Obradovic, P. Matagne, M. Stettler, and M. D. Giles, "Assessment of room-temperature phonon-limited mobility in gated silicon nanowires," *Appl. Phys. Lett.*, vol. 84, no. 25, pp. 5270–5272, Jun. 2004.

[13] Y. Zheng, C. Rivas, R. Lake, K. Alam, T. B. Boykin, and G. Klimeck, "Electronic properties of silicon nanowires," *IEEE Trans. Electron Devices*, vol. 52, no. 6, pp. 1097–1103, Jun. 2005.

[14] E. Gnani, S. Reggiani, A. Gnudi, P. Parruccini, R. Colle, M. Rudan, and G. Baccarani, "Band-structure effects in ultrascaled silicon nanowires," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2243–2254, Sep. 2007.

[15] K. Natori, "Ballistic metal–oxide–semiconductor field effect transistor," *J. Appl. Phys.*, vol. 76, no. 8, pp. 4879–4890, Oct. 1994.

[16] K. Natori, "Ballistic/quasi-ballistic transport in nanoscale transistor," *Appl. Surf. Sci.*, vol. 254, no. 19, pp. 6194–6198, Jul. 2008. DOI: 10.1016/j.apsusc.2008.02.150.

[17] K. Natori, "High-field transport in semiconductors by transmission formalism," *Phys. Stat. Sol. (C)*, vol. 5, no. 1, pp. 111–114, Jan. 2008.

[18] J. M. Luttinger and W. Kohn, "Motion of electrons and holes in perturbed periodic fields," *Phys. Rev.*, vol. 97, no. 4, pp. 869–883, Feb. 1955.

[19] R. Landauer, "Conductance from transmission: Common sense points," *Phys. Scr.*, vol. T42, pp. 110–114, 1992.

[20] K. Natori, "The capacitance of microstructures," *J. Appl. Phys.*, vol. 78, no. 7, pp. 4543–4551, Oct. 1995.

Kenji Natori (M'84) received the B.S., M.S., and Ph.D. degrees in physics from the University of Tokyo, Tokyo, Japan, in 1969, 1971, and 1974, respectively.

He was with Toshiba Research and Development Center, Kawasaki, Japan, in 1975. In Toshiba, he was engaged in the research on device physics of downscaled MOSFETs and the design of high-density dynamic memories. He supervised the design team of Toshiba 1-Mb CMOS DRAM, which was the first commercial high-density CMOS DRAM, as a Manager. While he was a Senior Manager with the Toshiba ULSI Research Center, his responsibilities included the basic research and development of the device and process for the advanced ULSI. In July 1992, he moved to the University of Tsukuba. Since 1994, he has been a Professor with the Institute of Applied Physics, University of Tsukuba, Tsukuba, Japan. His research interests are in the device physics of downscaled devices and their scaling limits, including ballistic carrier transport, the capacitance of microstructures, and the scaling limit of digital circuits. He is currently interested in nanoscale device transport.

Dr. Natori is a member of the IEEE Electron Devices Society, the Japan Society of Applied Physics, the Physical Society of Japan, and the IEICE, Japan. He is a member of the Organizing Committee of the International Conference on Solid State Devices and Materials.