The Bipolar Theory of the Field-Effect Transistor: X. The Fundamental Physics and Theory (All Device Structures) *

Sah Chih-Tang[1,2,3,†] and Jie Binbin[3,†]

(1 *University of Florida*, *Gainesville*, *Florida* 32605, *USA*)
(2 *Chinese Academy of Sciences*, *Foreign Member*, *Beijing* 100864, *China*)
(3 *Peking University*, *Beijing* 100871, *China*)

Abstract: This paper describes the foundation underlying the device physics and theory of the semiconductor field effect transistor which is applicable to any devices with two carrier species in an electric field. The importance of the boundary conditions on the device current-voltage characteristics is discussed. An illustration is given of the transfer DCIV characteristics computed for two boundary conditions, one on electrical potential, giving much higher drift-limited parabolic current through the intrinsic transistor, and the other on the electrochemical potentials, giving much lower injection-over-the-barrier diffusion-limited current with ideal 60mV per decade exponential subthreshold roll-off, simulating electron and hole contacts. The two-MOS-gates on thin pure-body silicon field-effect transistor is used as examples

Key words: bipolar field-effect transistor theory; MOS field-effect transistor; electric potential; electrochemical potential; boundary conditions.
PACC: 7340Q **EEACC**: 2560S; 2560B
CLC number: TN386.1 **Document code**: A **Article ID**: 0253-4177(2008)04-0613-07

1 Introduction

Electrical conduction in semiconductor devices contains two currents from two electrical charge carrier species, the electrons and the holes, which are quasi-particles in the one-particle representation of the energy band model of the many electrons and ions in a crystalline solid. Most devices are electrical and operate in externally applied electric fields. We will exclude the magnetic devices in this report. The theory of the electrical characteristics or terminal currents versus terminal voltages of semiconductor devices has been obtained by simultaneously solving six coupled differential equations known as the Shockley Equations[1,2], usually given in the Cartesian coordinate, $\boldsymbol{r} = (x, y, z)$. The first of the six Shockley Equations is the Poisson Equation of the electromagnetic theory[3], given in (1) below. It relates the divergence of the electric field, $\boldsymbol{E}(r, t)$ [which is derivable by the gradient of a scalar potential (Magnetic field is not included.), $\boldsymbol{E}(r, t) = -\nabla v(\boldsymbol{r}, t)$, known as the electric potential $v(r, t)$] to the net space-charge density from the mobile electrons and holes, $qn(\boldsymbol{r}, t)$ and $qp(\boldsymbol{r}, t)$, and the fixed and variable trapped charges, $qP_{IM}(\boldsymbol{r}) - qn_T(\boldsymbol{r}, t)$, in the volume and on the surface of the physical structure of the device or in the interface layer between the device and its surrounding. The current continuity equations, given by (2), (3), and (4) below, relate the divergence of the densities of the electron, hole and trapped electron currents, $\boldsymbol{j}_N(r, t)$, $\boldsymbol{j}_P(r, t)$ and $\boldsymbol{j}_T(r, t)$ flowing in the physical structure and through the surface and interface of the device. Included are the net generation, or generation less recombination, of electrons $= g_N(\boldsymbol{r}, t) - r_N(\boldsymbol{r}, t)$, holes $= g_P(\boldsymbol{r}, t) - r_P(\boldsymbol{r}, t)$, and net electron trapping $= r_N(\boldsymbol{r}, t) - g_N(\boldsymbol{r}, t) - r_P(\boldsymbol{r}, t) + g_P(\boldsymbol{r}, t)$[1,2,5] which is frequently omitted.

$$+\nabla \cdot (\varepsilon_S \boldsymbol{E}) = \rho(\boldsymbol{r}, t) = q(p - n - P_{IM} + n_T)$$
Poisson Equation (1)

$$+\nabla \cdot \boldsymbol{j}_N + q(g_N - r_N) = q\partial n/\partial t$$
Electron Current Continuity Equation (2)

$$-\nabla \cdot \boldsymbol{j}_P + q(g_P - r_P) = q\partial p/\partial t$$
Hole Current Continuity Equation (3)

$$+\nabla \cdot \boldsymbol{j}_T + q(g_P - r_P) - q(g_N - r_N) = q\partial n_T/\partial t$$
Electron Current Continuity Equation (4)

The remaining two are the electron and hole current equations, which show explicitly the origins of the electrical currents; the current due to drift of the charged particles from the electrical force arisen from

* This investigation and Jie Binbin have been supported by the CTSAH Associates (CTSA) founded by the late Linda Su-Nan Chang Sah, in memory of her 70th year.
† Corresponding author. Email: bb_jie@msn.com and tom_sah@msn.com
Received 24 March 2008, revised manuscript received 26 March 2008

the Coulomb Law and the current due to diffusion of the charged particles in the concentration gradient due to the statistics from random scattering of the many particles:

$$\boldsymbol{j}_N = + q\mu_n n\boldsymbol{E} + qD_n \nabla n$$

Electron Drift and Diffusion Currents (5)

$$= - q\mu_n n \nabla v_N$$

Electron Electrochemical Current (5A)

$$\boldsymbol{j}_P = + q\mu_p p\boldsymbol{E} - qD_p \nabla p$$

Hole Drift and Diffusion Currents (6)

$$= - q\mu_p p \nabla v_P$$

Hole Electrochemical Current (6A)

The drift and diffusion currents, (5) and (6), are represented in an alternative form shown in (5A) and (6A), which are the gradients of the electron and hole electrochemical potentials, $v_N(\boldsymbol{r}, t)$ and $v_P(\boldsymbol{r}, t)$, or the electrochemical field for electrons and holes, $\boldsymbol{E}_N = - \nabla v_N$ and $\boldsymbol{E}_P = - \nabla v_P$, and which are just the exponential representations or the Maxwellian-Boltzmann representations of the electron and hole concentrations, given by

$$n(\boldsymbol{r}, t) = n_i \exp\{q[v(\boldsymbol{r}, t) - v_N(\boldsymbol{r}, t)]/k_B T\} \quad (7)$$

$$p(\boldsymbol{r}, t) = n_i \exp\{q[v_P(\boldsymbol{r}, t) - v(\boldsymbol{r}, t)]/k_B T\} \quad (8)$$

These two electrochemical potentials, one for electrons and one for holes, were coined by Shockley in his 1949 classic, the 64-page invention paper of the bipolar junction transistors[4] as the quasi-Fermi potentials, which he doubly emphasized as the non-equilibrium quasi Fermi levels, since they approach the Fermi level (potential) defined at electrical equilibrium or zero currents[1]. The coefficients and constants here have the traditional meaning, q = electron charge, k_B = Boltzmann constant, (μ_n, D_n, μ_p, D_p) = electron and hole mobilities and diffusivities, which are related by the Einstein relationship at thermal equilibrium, $D/\mu = k_B T/q$, T = electron and hole temperatures or lattice temperature at thermal equilibrium[1]. At high electric fields, the near-equilibrium Boltzmann representation (or energy distribution) of the electrons and holes is extended to the Maxwellian representation in which the T would be the hot electron temperature T_E and hot hole temperature T_H respectively in the two exponential Maxwellian distributions[1].

There are actually a seventh electrical current equation and an eighth neutral particle flux equation[5]. The seventh is the continuity equation of the electrical current due to the mobile trapped charges or the mobile charged traps, by drift and diffusion, in the same form as the two continuity equations of the electrical currents due to the two mobile charges, electrons and holes. The eighth continuity equation of the neutral particle flux is due to the mobile neutral traps (due to impurity atoms and lattice defects, such as vacancies) via diffusion from their concentration gradients. Often we assume the trapped charges or the charged traps are immobile, i. e. with zero mobility and diffusivity so $\boldsymbol{j}_T(\boldsymbol{r}, t) = 0$ in (4)[5]. But, in a most important branch of diode-transistor device physics and engineering and circuit applications, the charged and neutral traps, which consist of foreign impurities and lattice defects, can move, although very slowly in years. But their movements change the composition of the device which alters the electrical characteristics. These two continuity equations, the seventh and eighth, quantify the lifetimes (operating life and shelf life) of the device and the integrated circuits composed of many devices.

2 Boundary Conditions Determine the Currents

For a given device, such as the two-terminal diodes (p/n, metal/semiconductor or m/s, and metal/oxide/semiconductor or m/o/s or MOS), and the three- or four-terminal transistors, the electrical characteristics (such as currents flowing into the device terminals as a function of the voltages applied to the terminals) are completely determined by the values of three variables at the boundaries of the device. In the traditional representation of the model, these three variables are the electron and hole concentrations, $n(\boldsymbol{r}, t)$ and $p(\boldsymbol{r}, t)$ and the electric field $\boldsymbol{E}(\boldsymbol{r}, t) = - \nabla v(\boldsymbol{r}, t)$ or the electric potential $v(\boldsymbol{r}, t)$. This is the preference of the chemists, because measurements of concentrations are routine to them. Electrical engineers are more familiar with potentials, in concept and in measurement capabilities. This electrical engineers' model, given by (5A) and (6A), using the two electrochemical potentials to represent the electron and hole concentrations, had been used extensively by Shockley when he introduced the quasi-Fermi potential concept in his 1949 invention of the p/n junction transistor [4], and by the subsequent transistor engineers to design diode-transistor devices for integrated circuits and by students and teachers to learn transistor physics. These three electrical engineers' variables are the three potentials: the electric potential $v(\boldsymbol{r}, t)$ and the electrochemical potentials for electrons and holes, $v_N(\boldsymbol{r}, t)$ and $v_P(\boldsymbol{r}, t)$.

We shall now give some examples on how the boundary conditions or the boundary values of these three potentials can affect the D.C. steady-state current-voltage (DCIV) characteristics of the 2-MOS-Gate on thin pure-Base silicon field-effect transistor, known as the FinFET[2,12].

To ease referencing, we list below the FinFET DCIV equations which were derived from the three equations of the six Shockley equations for the generation-recombination-trapping-free D. C. steady-state condition. Their derivations were given in the preceding article of this series, CJS11, which appeared in the last month issue[2]. The two-dimensional (2-D) terms from the longitudinal gradient of the longitudinal electric field (longitudinal electrical-field gradient or **LEG**) are excluded so this illustration is for long transistor with the channel physical length L much larger than the Debye length L_D since the LEG terms are scaled by $(L_D/L)^2$ where $L_D = [\varepsilon_{Si} k_B T/q(n+p)]^{1/2}$, first recognized and extensively illustrated by Shockley in 1949 as the crucial device parameter[4], is about $L_D \sim 25\mu m$ in pure silicon where $n = p = n_i = 1.33\times10^{10}cm^{-3}$ at $T = 300K$. So, for pure-silicon devices, the device is always small and multi-dimensional unless its dimensions are much larger than the pure Debye length of $\sim 25\mu m$. But for impure silicon, say $p = P_{IMPURITY} = 10^{18}cm^{-3}$ so $n = n_i^2/p = (1.33)^2\times10^2$, the Debye length is reduced to $\sim 25\mu m \times (10^{10}/10^{18})^{1/2} \sim 2.5nm$, making the 2D effects in state-of-the-art or next generation 25nm transistors not important, if built on heavily-impurity-doped silicon base-body. The effects on the DCIV characteristics from the neglected longitudinal electric field gradient, LEG, will be described in a later report in this series. The DCIV equations for the FinFET listed below are copied from [2], retaining the equation numbers of [2] but with the 2-D and impurity terms omitted. (For these omitted terms, please compare the following with the original Equations (57) to (70) in [2].) The coordinates are normalized: $X = x/L_D$ from $x = 0$ or $X = 0$ to the thickness of the base-body silicon film, $x = x_B$ or $X = X_B$; $Y = y/L$ from $Y = 0$ to Y_0 to 1, where Y_0 is the internal flatband boundary which is the electrical boundary that divides the physical length of the base-body silicon film of length L or $Y = y/L = L/L = 1$ into two electrical sections of electrical lengths Y_0 and $1 - Y_0$.

$$(\partial U/\partial X)^2 = F^2(U, U_P, U_N, U_0)$$
$$= +\exp(U - U_N) - \exp(U_0 - U_N) + \exp(U_P - U) - \exp(U_P - U_0) \quad (57)$$

$$U_{GS} - U_S = \mathrm{sign}(U_S - U_0)\times(C_D/C_O)\times(\partial U/\partial X)_S \quad \text{(Valid for all } Y.\text{) (The Voltage Equation.)} \quad (58)$$

Integration of U is from $U = U_0$ to $U = U_S$ ($x = x_B/2$ to x_B) and Y, from $Y = 0$ to $Y = 1$. (59)

$$X_B = 2\int \mathrm{sign}(U - U_0)\partial_X U/F \quad \text{(Valid for all } Y.\text{)} \quad \text{(The Thickness Equation)} \quad (60)$$

$$I_N = +qD_n n_i(W/L)L_D(\partial U_N/\partial Y)\exp(-U_N)\int[\exp(+U)\partial_X U/F(U, U_P, U_N, U_0)]\times 2 \quad (61)$$

$$= +qD_n n_i(W/L)L_D\int\partial_Y U_N\exp(-U_N)\int[\exp(+U)\partial_X U/F(U, U_P, U_N, U_0)]\times 2 \quad (62)$$

$$I_P = +qD_p n_i(W/L)L_D(\partial U_P/\partial Y)\exp(+U_P)\int[\exp(-U)\partial_X U/F(U, U_P, U_N, U_0)]\times 2 \quad (63)$$

$$= +qD_p n_i(W/L)L_D\int\partial_Y U_P\exp(+U_P)\int[\exp(-U)\partial_X U/F(U, U_P, U_N, U_0)]\times 2 \quad (64)$$

$$I_N = +qD_n n_i(W/L)[L_D/(Y)]\int\partial_Y U_N\exp(-U_N)\int[\exp(+U)\partial_X U/F(U, U_P, U_N, U_0)]\times 2 \quad (65)$$

$$I_P = +qD_p n_i(W/L)[L_D/(1-Y)]\int\partial_Y U_P\exp(+U_P)\int[\exp(+U)\partial_X U/F(U, U_P, U_N, U_0)]\times 2 \quad (66)$$

$$(y_0/x_B) = (\partial U_{N0}/\partial y)\cdot\exp(U_0 - U_{N0})/\int(\partial U_N/\partial y)\partial y\int[\exp(U - U_N)\partial x]\times 2$$
$$(x = 0\to x_B/2; y = 0\to y_0) \quad (67)$$

$$(L - y_0)/x_B = (\partial U_{N0}/\partial y)\cdot\exp(U_0 - U_{N0})/\int(\partial U_N/\partial y)\partial y\int[\exp(U - U_N)\partial x]\times 2$$
$$(x = 0\to x_B/2; y = y_0\to L) \quad (68)$$

$$(L - y_0)/x_B = (\partial U_{P0}/\partial y)\cdot\exp(U_{P0} - U_0)/\int(\partial U_P/\partial y)\partial y\int[\exp(U_P - U)\partial x]\times 2$$
$$(x = 0\to x_B/2; y = y_0\to L) \quad (69)$$

$$(y_0/x_B) = (\partial U_{P0}/\partial y)\cdot\exp(U_{P0} - U_0)/\int(\partial U_P/\partial y)\partial y\int[\exp(U_P - U)\partial x]\times 2$$
$$(x = 0\to x_B/2; y = 0\to y_0) \quad (70)$$

There are three potential variables, $U_N(Y)$, $U_P(Y)$ and $U(X, Y)$ with their two internal values at the two SiO_2/Si interfaces, $x = 0$ and $x = x_B$, and at the mid-base minimum or maximum, respectively given by $U(X = 0, Y) = U(X = X_B, Y) = U_S(Y)$ and $U(X = X_B/2, Y) = U_0(Y)$. To obtain the electron

and hole channel currents from (61) to (64), we need to relate the three of the four potentials to the fourth, the Y integration potential. For example, to compute the electron surface and volume channel current, I_N, using the differential equation (61) with a single integration in X or $\partial_X U$ for the electron charge, Q_N, or using the integral equation (62), which is the solution of the differential equation (61) with two integrations, in X or $\partial_X U(X, Y)$ then in Y or $\partial_Y U_N(Y)$, we need to relate $U_S(Y)$, $U_0(Y)$ and $U_P(Y)$ to $U_N(Y)$. We have two internal equations, the Voltage Equation given by (58) and the Thickness Equation given by (60). We need a third equation, to relate $U_P(Y)$ to $U_N(Y)$. It has been arbitrarily set without a device physics basis in all previous work. For our newly discovered bipolar nature of all FETs, this third equation or relation to relate U_P to U_N in order to compute the electron channel current in the Bipolar FET, is obtained from the hole current equations, either (63) or (64), and the implicit coupling of the electrons with the holes is through the space-charge density of the Poisson Equation. So, this is a recursive iteration exercise by successive numerical approximation within each recursive iteration cycle and by repeating the cycle (so named recursive, if not in conflict with another usage) until a specified convergence is reached, for example, the electron and hole currents, I_N and I_P, no longer change more than a user specified percentage from the previous recursive iteration cycle to the present recursive iteration cycle. The obvious starting point is the unipolar solution, that is, for I_N let $P = 0$, and for I_P let $N = 0$, because there is enough equations, i.e., the Voltage and Thickness Equations, (58) and (60), for the three variables, the electric and electron-electrochemical potentials, $U_S(Y)$, $U_0(Y)$, and $U_N(Y)$, at any Y in the internal channel and at its two boundaries at $Y = 0$ and $Y = 1$. For convenience and on account of device physics, we have called it in our CJS05 report[11] the 0^{th} cycle or the initial guesses or guess of the recursive iteration algorithms. However, these initial solutions based on $N = 0$ or $P = 0$, correctly called the Unipolar solutions, are not strictly the correct initial solutions of the problem formulated in the three potentials, but rather they are those for the Drift and Diffusion formulation. The senior author also recognized the two obvious algorithms or routes to carry out the recursive iteration solution and we coined them the cross-linked double-string algorithm and the zig-zag single-string algorithm[11]. The final and converged numerical DCIV characteristics from the two algorithms should be identical if the stop criterion is set to a high degree of accuracy.

In addition, the internal boundary of the two-section model, first defined by the first author, in his 1991 – 1993 textbook to teach junior students in the device core course, $0 \leqslant Y_0 \leqslant 1$, can be computed from the four equations, (67) to (70), all should give the same numerical result when the user-defined convergence is reached and if the convergence criterion is set at a sufficiently high accuracy. This internal boundary, defined by Y_0, is necessary to compute the DCIV characteristics in the current saturate range when the field-effect transistor is electrically divided into two sections with electrical lengths of y_0 and $L - y_0$, where y_0 or $Y_0 = y_0/L$ is the flatband boundary inside the transistor, which are called by us[11] as the emitter section and collector section, just like those in the Bipolar Junction Transistor coined by Shockley in 1949[4]. This electrical division is also important to compute the presently omitted 2-D LEG effects on the DCIV characteristics, which was not recognized by device modeling engineers, although the two section model was a prominent feature introduced and exhaustively described, based on device physics, by Shockley in his 1952 invention article on the Junction-Gate Field-Effect Transistor (JGFET) which has two identical n/p junction gates and a p-type volume channel[6] and which was analyzed with high-power exotic analytical mathematics by Prim and Shockley in 1953 as an introductory exercise by the young Prim to learn the analytical mathematical solution of transistors[7] in the pre-desktop computer era, without the canned numerical computer programs as device simulators of today, which has replaced the physics-based analytical solutions, removing the incentive of the education to recent students, teachers and practicing engineers on understanding fundamental diode-transistor device physics.

The DCIV curves are highly dependent on the Y boundary conditions set by the user for the four variables $U_S(Y)$, $U_0(Y)$, $U_N(Y)$ and $U_P(Y)$ at the two ends of the transistor's thin base, $Y = 0$ at the source and $Y = 1$ at the drain. Two cases are described below to illustrate.

The least restrictive case is the boundary condition of $U_0(Y = 0) = U_{SB} = 0$ and $U_0(Y = 1) = U_{DB} \geqslant 0$ with $U_{GB} \geqslant 0$ for the electron channel current, I_N, and the corresponding bias configuration or boundary condition of $U_0(Y = 0) = U_{SD} \leqslant 0$ and $U_0(Y = 1) = U_{DB} = 0$, with $U_{GD} \leqslant 0$ for the hole channel current, I_P. In this case, source and drain electric potential boundary contacts are the two end points located at the mid-plane of the thin silicon base-body, the source point at $(x = x_B/2, y = 0)$ and the drain point at $(x = x_B/2, y = L)$ of the "intrinsic" transistor. "Intrinsic"

means the geometry definition of the ideal transistor structure, excluding the fringe fields at the corners. They are exactly the boundary points assumed in solving the y-independent Poisson equation so there is no physical reality conflict. Since $U_N(Y)$ and $U_P(Y)$ are not specified in the base-body volume and at the boundaries, $Y = 0$ and $Y = 1$, the electron and hole currents, I_N and I_P, and the magnitude and spatial variations of their concentrations, P and N, are not restricted and are set only by the voltage applied to the two boundary points, labeled Source and Drain. Since the boundary conditions are imposed on the electric potential, $U(x, y)$, that controls the drift current, and not the electrochemical potentials of electrons and holes, that controls the diffusion currents, the channel currents are mobile carrier concentration limited drift currents, and the concentration gradient limited diffusion currents are negligible. So, the electron and hole currents would be higher than those from another set of boundary conditions imposed on $U_N(Y)$ and $U_P(Y)$, that would bring out the diffusion limitation, such as those computed by us for the bipolar theory of the MOS field effect transistor using the electrochemical potential approach[8,9] with the boundary conditions of $U_N(Y=0) = U_{SB} = 0$ and $U_N(Y=1) = U_{DS} \geqslant 0$, and gate voltage of $U_{GB} \geqslant 0$ for the electron channel current I_N; and $U_P(Y=0) = U_{SB} \leqslant 0$ and $U_P(Y=1) = U_{DB} = 0$, and gate voltage of $U_{GB} \leqslant 0$ for the hole channel current I_P. These boundary conditions on $U_N(Y)$ and $U_P(Y)$ at $Y = 0$ and $Y = 1$, correspond to electron and hole contacts, which would then limit the electron and hole currents and also the spatial variations of the electron and hole concentrations by the current-voltage characteristics of the contacts, whose physical structures, such as the p^+/i, n^+/i and m/i diodes, are excluded from the intrinsic transistor. Even more restricted conditions on $U_N(Y)$ and $U_P(Y)$ have been imposed by others[10], not just at the boundaries, but also in the interior or the base-body, which are out of reach and physically unrealizable.

Numerical illustration to compare the transfer DCIV, I_N and I_P versus V_{GS}, at $V_{DS} = +0.5$V is given in Fig. 1, for the two different boundary conditions: (BC-1) BC－U_0 of $U_0(Y=0) = U_{SB} = 0$ and $U_0(Y=1) = U_{DB} \geqslant 0$, $U_N(Y)$ and $U_P(Y)$ are not restricted; and (BC-2) consisting of BC-U_N of $U_N(Y=0) = 0$ and $U_N(Y=1) = U_{DB} \geqslant 0$ with $U_{GB} \geqslant 0$, and BC-U_P of $U_P(Y=0) = U_{SB} \leqslant 0$ and $U_P(Y=1) = U_{DB} = 0$ with U_{GD} and $U_{GB} \leqslant 0$. The (BC-2) curves show the contact controlled DCIV characteristics with lower surface channel currents in both the large drift-dominated "inversion" current range and the small diffusion-

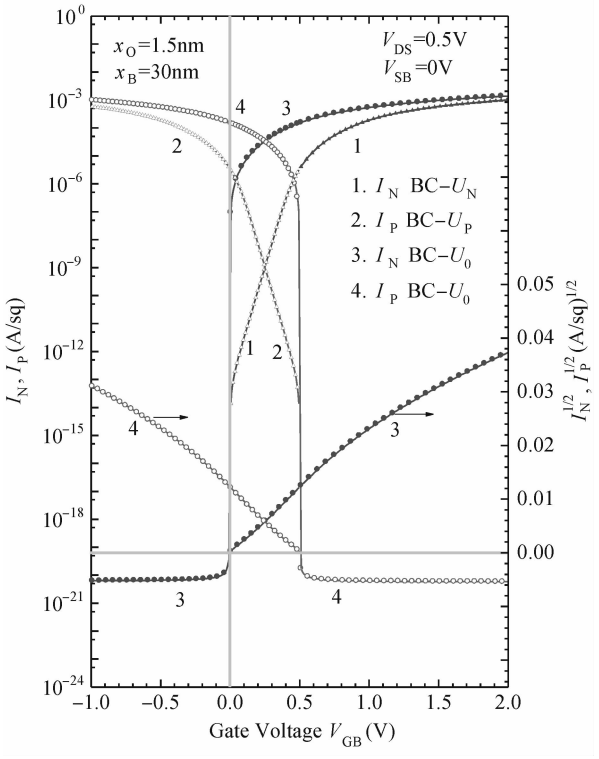

Fig. 1　Transfer DCIV characteristics for the two different boundary conditions shown in the figure.

dominated subthreshold range of the classical 60mV per decade decrease, which is limited by the electron contact n^+/i for I_N and hole contact p^+/i for I_P when the U_N and U_P values at the boundaries are specified by the user. The (BC-1) curves show much higher currents since the electrons and holes are not limited by boundary or contact conditions, resulting in the dominance of the space-charge-limited drift current that is power-law (parabolic) dependence on the gate voltage. The very abrupt cutoff, is an artifact from graphing on the semi-log scale of an ohmic or linear, or a power-law (parabolic) single-carrier mobile space-charge-limited I-V characteristics. The DCIV characteristics covering extended range of parameters are reported in the following article[11].

3　Summary

The fundamental basis of the bipolar solutions of the bipolar MOS field effect transistor is reviewed to help show the importance of boundary conditions of the electric and electrochemical potentials on the DCIV characteristics of the transistor. The electron and hole surface channel currents are not limited by diffusion from concentration gradient, only drift from electric field when the boundary conditions are applied only to the electric potential. When the boundary values are imposed on the electrochemical potentials, the currents are limited by the assumed electron

or hole contact types to the source and drain, such as the diffusion-limited over-the-barrier injection-current from the n^+/i and p^+/i contacts, simulated by imposing the boundary values on $U_N(Y)$ and $U_P(Y)$ at the source $Y=0$ and drain $Y=1$ contacts, giving rise to the 60mV per decade current roll-off from the classical 1949-Shockley p/n junction diode $\exp(U_{PN})$.

Acknowledgment We thank Professors Xing Zhou (Nanyang Technological University, Singapore) Chenming Hu (University of California, Berkeley), Gennady Gildenblat (Arizona State University) and Mitiko Miura-Mattausch (Hiroshima University) and Dr. Colin McAndrew (Freescale Semiconductor Corporation), and Drs. Jin Cai, Tak H. Ning, Lewis M. Terman, and Hwa-Nien Yu (all of IBM Thomas J. Watson Research Center) for encouragements, comments and suggestions. We further thank Professors Marcel D. Profirescu (Univerisity Politechnica of Bucharest, Romania), Adelmo Ortiz-Conde and Francisco J. Garcia Sanchez (Universidad Simon Bolivar, Venezuela), Juin J. Liou (University of Central Florida, USA) for inviting us to present our results at their IEEE-EDS-sponsored mini-colloquium and international conferences, NADE, mini @ Cancun and ICCDCS.

References

[1] Chih-Tang Sah, Fundamentals of Solid-State Electronics, 1001pp, World Scientific Publishing Company, Singapore, New Jersey, London. 1991. Section 202 on page 153 – 159 on equilibrium. Section 350 on p. 268 listing the six Shockley Equations. Sections 313 on pages 241 – 249 for phonon-scattering limited mobilities and Section 314 on pages 251 – 252 and Section 684 on pages 667 – 670 on hot-carrier limited mobility effects on the DC current-voltage characteristics of MOSFETs. Section 384 on pages 302 and 303 on generation of secondary electron-hole pairs by lucky hot primary electrons and holes at high energies or electric fields and its effects on p/n junction characteristics in Section 536 on pages 441 to 450.

[2] Chih-Tang Sah and Binbin Jie, "The Theory of Field-Effect Transistors: XI. The Bipolar Electrochemical Currents (1-2-MOS-Gates on Thin-Thick Pure-Impure Base," Journal of Semiconductors, 29(3), 397 – 409, March 2008. (CJS11)

[3] See the following classics and modern textbooks. Julius Adams Stratton, Electromagnetic Theory, McGraw-Hill Book Company, New York, 1941; Gaylord P. Harnwell, Principles of Electricity and Electromagnetism, McGraw-Hill Book Company, New York, 1938, Second Edition, 1941; Wolfgang K. H. Panofsky and Melba Phillips, Classical Electricity and Magnetism, Addison-Wesley Publishing Company, Cambridge, Massachusetts, 1955; Hermann A. Haus and James R. Melcher, Electromagnetic Fields and Energy, Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[4] William Shockley, "The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors," Bell System Technical Journal 28, 435 – 498, July 1949. See Eq. (2.4) on page 439 for the definitions of the non-equilibrium quasi Fermi potentials using the exponential representation of the electron and hole concentrations. See Equation (2.12) on page 441 for the definition of the Debye Length, in CGS units, which many if not all subsequent device theorists have missed and used the wrong definition from omission of the factor of 2 inside the square root.

[5] Chih-Tang Sah, Fundamentals of Solid-State Electronics-Study Guide, 423 pages, World Scientific Publishing Company, Singapore, New Jersey, London, 1993. The generalized Shockley Equations including mobile traps are given in Section 430 on pages 160 – 161. The two general analytical solutions of the Poisson Equation via integration by quadrature and via integrating twice are given on pages 140 to 141 of Section 412, which were for the equilibrium MOS capacitor but are also applicable to the general non-equilibrium case if the electrochemical potentials of electrons and holes are independent of the coordinate of integration, such as the x-axis and z-axis in the two-dimensional device model commonly analyzed by others and in this series of articles. The flatband location, $0 \leqslant Y_0 \leqslant 1$, to define the boundary jointing the two electrical sections of the MOS transistor in the current saturation range of gate-source and drain-source DC bias voltage ranges, was first illustrated graphically as a A1G (At One Glance) teaching aid by the author in Fig. B. 2.5(b) on page 411 for the bulk nMOST and Fig. B. 2.6(c) on page 412 for the bulk pMOST.

[6] William Shockley, "A Unipolar 'Field Effect' Transistor," Proceedings of the IRE, 40(11), 1365 – 1376, November 1952.

[7] R. C. Prim and W. Shockley, "Joining Solutions at the Pinch-Off Point in 'Field-Effect' Transistors," Transactions of the IRE, Professional Group on Electron Devices, PGED-4, 1 – 14, December 1953, Private communication from Shockley to Sah in one of one-on-one weekly tutorial sessions during 1956 – 1958.

[8] Chih-Tang Sah and Bin B. Jie, "The Bipolar Field-Effect Transistor: I. Electrochemical Current Theory (Two-MOS-Gates on Pure-Base)," Chinese Journal of Semiconductors 28(11), 1661 – 1673, November 2007. (CJS01)

[9] Binbin Jie and Chih-Tang Sah, "The Bipolar Field-Effect Transistor: III. Short Channel Electrochemical Current Theory (Two-MOS-Gates on Pure-Base)." Journal of Semiconductors, 29(1), 1 – 11, January 2008. (CJS03)

[10] A. Ortiz-Conde, F. J. Garcia-Sanchez, J. Muci, S. Malobabic, J. Liou, "A Review of Core Compact Models for Undoped Double-Gate SOI MOSFETs," IEEE Trans Elec Dev, 54(1), 131 – 139, January 2007.

[11] Binbin Jie and Chih-Tang Sah, "The Bipolar Field-Effect Transistor: V. Bipolar Electrochemical Current Theory (Two-MOS-Gates on Pure-Base)." Journal of Semiconductors, 29(4), 620 – 627, April 2008. (Following paper in this issue.) (CJS05)

[12] Chenming Hu, "From CMOS to Nanotechnology," Keynote, 17th Annual IEEE/SEMI Advanced Semiconductor Manufacturing Conference, May 22 – 24 2006, Boston, Massachusetts.

场引晶体管双极理论：X. 基本物理和理论(所有器件结构)*,**

薩支唐[1,2,3,†]　揭斌斌[3,†]

(1 佛罗里达大学，佛罗里达州，Gainesville　FL32605，美国)
(2 中国科学院外籍院士，北京　100864)
(3 北京大学，北京　100871)

摘要： 本文描述半导体场引晶体管器件物理和理论所用的根本原则，它适用电场中有两种载流子的器件. 讨论边界条件对器件电流电压特性的重要性. 作为例子，计算两种边界条件下的转移直流电压特性：电势边界给出很高、流进内禀晶体管、飘移限制抛物型电流，电化学势边界仿真电子和空穴接触，给出很低、越过势垒注入、扩散限制电流，具有理想、每量级 60mV、指数型亚阈值区倾斜. 双 MOS 栅 薄纯基硅场引晶体管为典型结构.

关键词： 双极场引晶体管理论；MOS 场引晶体管；电势；电化学势；边界条件
PACC：7340Q　　**EEACC**：2560S；2560B
中图分类号： TN386.1　　**文献标识码：** A　　**文章编号：** 0253-4177(2008)04-0613-07

* 该研究及揭斌斌由 CTSAH Associates (CTSA) 资助. CTSA 由薩故夫人张淑南创建. 记念她七十周年.
** 薩支唐写成此摘要基于揭斌斌的现代语初稿. 感谢潘胜和北京大学原物理系教师赵立群和潘桂明的修改建议.
† 通信作者. Email：bb_jie@msn.com and tom_sah@msn.com
2008-03-24 收到，2008-03-26 定稿