第 29 卷　第 3 期
2008 年 3 月

半　导　体　学　报
CHINESE JOURNAL OF SEMICONDUCTORS

Vol. 29　No. 3
Mar. , 2008

The Theory of Field-Effect Transistors: XI. The Bipolar Electrochemical Currents (1-2-MOS-Gates on Thin-Thick Pure-Impure Base)*

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

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

Abstract: The field-effect transistor is inherently bipolar, having simultaneously electron and hole surface and volume channels and currents. The channels and currents are controlled by one or more externally applied transverse electric fields. It has been known as the unipolar field-effect transistor for 55-years since Shockley's 1952 invention, because the electron-current theory inevitably neglected the hole current from over-specified internal and boundary conditions, such as the electrical neutrality and the constant hole-electrochemical-potential, resulting in erroneous solutions of the internal and terminal electrical characteristics from the electron channel current alone, which are in gross error when the neglected hole current becomes comparable to the electron current, both in subthreshold and strong inversion. This report presents the general theory, that includes both electron and hole channels and currents. The rectangular (x, y, z) parallelepiped transistors, uniform in the width direction (z-axis), with one or two MOS gates on thin and thick, and pure and impure base, are used to illustrate the two-dimensional effects and the correct internal and boundary conditions for the electric and the electron and hole electrochemical potentials. Complete analytical equations of the DC current-voltage characteristics of four common MOS transistor structures are derived without over-specification: the 1-gate on semi-infinite-thick impure-base (the traditional bulk transistor), the 1-gate on thin impure-silicon layer over oxide-insulated silicon bulk (SOI), the 1-gate on thin impure-silicon layer deposited on insulating glass (SOI TFT), and the 2-gates on thin pure-base (FinFETs).

Key words: bipolar field-effect transistor theory; MOS field-effect transistor; simultaneous electron and hole surface and volume channels and currents; surface potential; two-section short-channel theory; double-gate impure-base theory

PACC: 7340Q　　**EEACC**: 2560S; 2560B

CLC number: TN386. 1　　**Document code**: A　　**Article ID**: 0253-4177(2008)03-0397-13

1　Introduction

The principle of conductivity modulation by a transverse electric field to provide a 3-terminal solid-state device to amplify electrical signal was conceived by Lilienfeld in his three United States patents filed and issued during 1926～1933[1]. The practical field-effect transistor (FET) was invented and mathematically analyzed by Shockley in 1952[2]. Shockley's transistor consisted of a p-type semiconductor slice (known as base) with two p-type contacts to the two ends of the slice, denoted as the drain and source of holes for the hole current flowing through p-type slice or p-type <u>volume</u> conduction channel. The transistor also contains two n-type regions (known as gates) built into the bulk of the base-slice through the two opposite surfaces of the p-type base-slice. The two n-Gate/p-Base n/p-junctions are reverse biased to reduce the holes in the p-type base-slice, hence the electrical conductance of the volume channel of holes. This Junction-Gate FET (JG-FET or JGFET) was further analyzed by Prim and Shockley in 1953[3] using Shockley's two-section model in which the length of the transistor is divided into two sections, the source emitter section where Shockley's gradual channel approximation was used and the drain collector section where the carrier depletion approximation was used. The boundary separating the two sections is the point where the channel thickness is pinched off by the thickened and carrier-depleted space-charge region of each of the two reverse-biased n-Gate/p-Body junctions. The transistor was reduced to practice by Dacey and Ross in 1953[4]. Shockley's 1952 unipolar theory

---

* This investigation and Jie Binbin have been supported by the CTSAH Associates (CTSA) founded by the late Linda Su-Nan Chang Sah, in memory on her 70th year.

† Corresponding author. Email: bb_jie@msn. com and tom_sah@msn. com

Received 22 February 2008, revised manuscript received 27 February 2008

and two-section model have been used in the last 55 years to analyze the electrical characteristics of four FETs with insulated gates for volume applications: (i) The most widely used FETs with one metal-conductor/oxide-insulator/semiconductor-silicon (MOS) gate[5,6] on a semi-infinite thick and impure-base of semiconducting-silicon (Bulk MOSFET)[5] to develop compact transistor models[6] for use with circuit simulators such as SPICE[5,6]; (ii) The FETs with one MOS gate on polycrystalline silicon thin film deposited on a insulator such as glass (SOI TFT), recently made popular, as the pixel driver of Liquid Crystal Display (LCD) panels of personal computers and televisions[7,8]; (iii) The FETs with one MOS gate on thin crystalline silicon layer which is insulated from the crystalline silicon substrate by a thick layer of SiO_2 obtained from a high concentration of ion implanted oxygen, subsequently thermally activated to form a buried SiO_2 layer (SOI SIMOX MOSFET), for higher switching speed from reduced overlap and bulk-junction parasitic capacitances that load down the interconnect lines[9,10]; and (iv) The FETs with two MOS gates on fin-shaped thin silicon pure-base (FinFET)[11] as the technology advances towards nanometer dimensions[12].

While reviewing the literature on the FinFET in order to referee a manuscript for the IEEE Transactions on Electron Devices in February, 2007, we noticed the distinct presence of simultaneous electron and hole currents in the experimental DC current voltage characteristics of the latest silicon nanometer FinFET, reported by IMEC with four groups of Assignees from four semiconductor integrated circuit manufacturers (Intel, Phillips, Samsung, TI)[13], which we reported as two Late News at the 2007-May-23 Workshop on Compact Modeling[14] and as an invited paper, 5 months later with additional discoveries about the history, at the 2007-October-14 IEEE Electron Device Society Mini-Colloquium at Bucharest, Romania[15]. It was at once obvious to us, in March 2007[14], that the simultaneous presence of electron and hole channels, or the nMOST (nFET) and pMOST (pFET), could increase the circuit function density by a factor of two to six, aside from speed gain, since one physical BiFET can be connected internally to give the 2-Transistor nFET-pFET CMOS inverter and the 6-Transistor SRAM-like bistable memory. To help advance rapidly the circuit design technology without years of delay from lacking a compact transistor model based on the correct transistor theory, we embarked on an exposition to provide the complete bipolar theory of the bipolar FET to replace Shockley's 55-year-old unipolar theory of the bipolar 'field-effect' transistor[2]. Our results have been reported in five journal articles[16~20] which did include the simultaneous presence of the electron and hole currents. However, our numerical solutions only reached the zeroth-order or initial-guess using the unipolar solution (actually hole-current-less bipolar solution that includes the mobile or configurable spatial distribution of the hole carrier concentration) in each of the two emitter sections, such as the source electron-emitter section, without further numerical iteration to take into account of the hole current; similarly, the drain hole-emitter section, without further taking into account of the electron current. These unipolar approximations to give the bipolar currents-voltage characteristics of a bipolar MOSFET are actually quite accurate when the electron or the hole current dominates, which we recognized[14], not unlike the Matthiessen's Rule, but gross error is expected in bias voltage ranges where the electron and hole currents are comparable or equal, such as the intersection of subthreshold current-voltage characteristics of the nFET and pFET, around half of the threshold voltage of each transistor, and the high drain/source voltage range beyond the drain electron current saturation voltage of the electron channel where the electron channel current is saturated while the hole channel current is continuously increasing parabolically in the saturation range with increasing drain to source voltage, reaching the electron channel current value at $V_{DS} = 2V_{GS} > 0$ for the initially electron channel nFET-UniFET, which changes over to BiFET with constant nFET current and parabolically rising pFET current, both in their current saturation range. (See the I_D-V_{DS} characteristics in the inset of Fig. 1 and also on the issue cover of our first BiFET article in October 2007[16] and in Fig. 4(a) computed by the electrochemical potential-gradient electron and hole channels and currents theory[17].)

In the following sections we present the bases that lead to the general theory, including both electron and hole channels, but also electron and hole currents, because including just the surface inversion channels, as some may have implicitly done, and not the currents, is inadequate when both the electron and hole currents are of same order of magnitudes. The rectangular (x, y, z) parallelepiped transistors, uniform in the width direction (z-axis), is used to give the generic solution. Such a rectilinear coordinate system is designed to facilitate the inclusion of the two 2-D terms, from the longitudinal gradient of the longitudinal electric field in the gate oxide, $\int[\partial E_y(-x_O \leqslant x \leqslant 0,\ y)/\partial y]\partial x$, and in the base-channel,

$\int[\partial E_y(0 \leqslant x \leqslant x_B, y)/\partial y]\partial x$, by additional iterations. The central issue on the correctness of the theory concerns the internal and boundary conditions imposed on the three independent variables, the electric potential and the electron and hole electrochemical potentials, for which we shall give the rigorous theoretical derivations in the following sections for the four transistor geometries of volume applications: one or two MOS gates on thin and thick, pure and impure base, and all uniform in width (z-axis).

2 Electrochemical Potential Theory

The DC electrical characteristics of the transistor are governed by the six Shockley Equations[21]. They reduce to three equations for the case of zero or negligible electron-hole generation-recombination-trapping (GRT) which would otherwise provide the interaction link between the electron and hole concentrations. These are the Poisson Equation that couples the electron and hole charge densities or concentrations via the divergence of the electric field, and the two current continuity equations one each for electrons and holes or the Kirchoff's Current Law of conservation of the rate of change of charges (electrons and holes) with time and the currents carried by the moving charges (electrons or holes). The current continuity equations are derivable from the Boltzmann transport Equation, one each for the electrons and holes. The two current components are the drift current due to the electric field force on the charged electrons and holes, and the diffusion currents due to the concentration gradients of the electrons and holes. These are derived here for the analytical characterization of the BiFET at the DC Steady-State. The algebra steps are given by the following equations numbered ($\#\alpha$) where $\# = 1$ to 8 and $\alpha =$ A to D or H and the final eight equations used in this report for the GRT-free case are numbered by ($\#$). The derivation has not been presented in books and articles in sufficient details to provide the justifications and approximations necessary for the analytical solution of the BiFET, so they are described in this report. The rigorous results included the two-dimensional terms (with uniform properties in the z-direction) which have not been included in previous theories but which are increasingly important, to the dominant proportion, as the transistor dimension, such as the channel Length in the FET, L, decreases towards nanometers, the semiconductor dopant impurity concentrations decreases towards pure or zero, and the local Debye screening length, L_D, of the fixed and mobile charges by the mobile charges, increases towards the pure-silicon value of about $L_D \sim 25\mu$m at room temperatures, since the non-geometrical 2-Dimensional effects such as the longitudinal electric field gradient, i. e., the divergence of the longitudinal electric field, is proportional to or scaled by $(L_D/L)^2$. To simplify the notations and results, we use the exponential representation of the three-dimensional electron and hole volume concentrations, $N(x,y,z)$ and $P(x,y,z)$, by their thermal-voltage-(k_BT/q)-normalized electrochemical potentials $U_N(x,y,z)$ and $U_P(x,y,z)$ and the electrostatic potential or DC electric potential represented by the intrinsic or pure-semiconductor Fermi potential, $U(x,y,z) = U_I(x,y,z)$. In this normalized notation, we have $N = n_i\exp(U - U_N)$ and $P = n_i \times \exp(U_P - U)$ where n_i is the electron and hole concentration in a pure or intrinsic sample, k_B is the Boltzmann constant, and T is the lattice or semiconductor temperature. We also use the exponential representation for the net-positive-charge of the ionized acceptor over donor impurity concentrations, defined by $P_{IM} = P_F - N_F = n_i\exp(+U_F) - n_i\exp(-U_F)$ in which F is a $\underline{f}$ar away or $\underline{r}$emote location, $\underline{R}$, where the potential may be taken as the universal or global reference from the Coulombic or the $1/r$ electrostatic point-charge potential, $U(x_R, y_R, z_R) = 0$. The use of $R \to \infty$, is selfconsistent also, since $\text{limit}(r \to \infty)(1/r) = 0$ for the potential, $\text{limit}(r \to \infty)(\partial/\partial r)(1/r) \to 0$ for the electric field which is the gradient of the potential, and $\text{limit}(r \to \infty)(\partial^2/\partial r^2)(1/r) \to 0$ for the divergence of the electric field, which is proportional to the space charge, giving the remote charge neutrality condition. Thus, at this remote reference point, R, thermal equilibrium[21] is reached with no electric field and no electrical net charge or with electrical neutrality. Then, the mass action law or detailed balance holds, giving $P_F N_F = P_R N_R = P_\infty N_\infty = n_i^2$ which can be solved simultaneously with $P_{IM} = n_i\exp(+U_F) - n_i\exp(-U_F)$ to give the Fermi potential or energy U_F and the equilibrium hole and electron concentrations as a function of the net acceptor impurity ion concentration, P_{IM}, which are

$$P_\infty/n_i = \exp(+U_F) = (1/2)(P_M/n_i)\{1 + [1 + (2n_i/P_{IM})^2]^{1/2}\} \to P_{IM}/n_i \qquad \text{when } P_{IM}/n_i > 20$$

and

$$N_\infty/n_i = \exp(-U_F) = (1/2)(P_{IM}/n_i)\{-1 + [1 + (2n_i/P_{IM})^2]^{1/2}\} \to n_i/P_{IM} \qquad \text{when } P_{IM}/n_i > 20.$$

In the decades of traditional solid-state and transistor electronic device engineering practices, the subscript B has been used instead of R and F, to write N_B, U_B, P_B, for the basewell-channel of MOSFET, the $\underline{b}$ase of the BJT, and the $\underline{b}$ulk or body of a semiconductor slice or slab, but B is also frequently used as the contact to the outside world and as the reference in the FET and BJT (Bipolar Junction Transistor) terminology, which creates a notation conflict, but worse, causes the theorists to use erroneous boundary conditions, such as the notion of floating base that contradicts the rigorously defined DC Steady-State condition, leading to incorrect computed current-voltage characteristics, that confuses the issue with the transient responses of the transistor during switching and turn-on and turn-off transients which must be treated as time-dependent phenomena, rather than piecewise DC Steady-State conditions. And such neglect and arbitrary set of an internal variable among the three potentials (electric and two electrochemical) becomes detrimental in DC characteristics analyses of the pure-base dual gate MOSFET characteristics[5,11,14~20]. For this presentation of the DC analysis of the BiFET characteristics, we assume near thermal equilibrium[21] with no hot carrier effects so the electron and hole mobilities and diffusivities closely follows the equilibrium Einstein Relationship $D_n/\mu_n = k_B T/q = D_p/\mu_p$. This can be easily extended to take into account of non-thermal-equilibrium distributions of electrons and holes in high electric fields by means of hot electron and hole temperatures, T_E and T_H using the Maxwellian energy distributions or the exponential representations extended to high concentrations of hot carriers, in addition to the near-thermal-equilibrium lattice or heat-transfer temperature, $T_L \equiv T$, in which, the hot carrier temperatures are calculated from the power transfer balance or rate of energy exchange between the carriers and the lattice vibrations or phonons via the energy-conservation (with acoustic phonons) and power-dissipation (with intervalley-interband acoustic and optical phonons) scattering transitions first treated by Shockley in 1951[21,22]. A simple kinetic description of the hot carriers using the ballistic or the Lucky Electron (Hole) Model given by Shockley[21,23] can also be used for low densities of or a few lucky hot carriers to account for generation of the secondary electrons and holes by energetic primary electrons and holes accelerated to the phonon energies by the high electric field. Both high-field reduction of the electron and hole mobilities and electron-hole pair generation and recombination are excluded in this report.

A rigorous derivation of the three equations, (1), (2) and (3) listed below, to be used by the BiFET theory, are given in the following paragraphs. The Cartesian coordinate unit vectors are defined by $\boldsymbol{r} = \boldsymbol{i}_X x + \boldsymbol{i}_Y y + \boldsymbol{i}_Z z$ and the volume element is given by $\mathrm{d}\mho = \mathrm{d}x\mathrm{d}y\mathrm{d}z$. In order to simplify the notation and also to bring out the device and materials physics in scaling, we normalize the x-axis to the Debye Length of pure silicon, $L_D = (\varepsilon_S k_B T/2q^2 n_i)^{1/2}$, $X = x/L_D$, and the y-axis to the channel length L, $Y = y/L$.

Shockley Equations (Three of Six at DC for No-Generation-Recombination-Trapping[21].)

Poisson Equation

$$\nabla \cdot (\varepsilon_S \boldsymbol{E}) = \rho(x, y, z) = q(P - N - P_{IM}) \qquad \text{Poisson Equation 3-Dimensional Impure (1A)}$$

$$= \varepsilon_S[+(\partial E_X/\partial x) + (\partial E_Y/\partial y)] = q(P - N - P_R + N_R) \qquad \text{2-Dimensional Impure (1B)}$$

$$= \varepsilon_S(k_B T/q)[-(\partial^2 U/\partial x^2) - (\partial^2 U/\partial y^2)] = q(P - N - P_F + N_F) \qquad \text{2-Dimensional Impure (1C)}$$

$$= qn_i[\exp(U_P - U) - \exp(U - U_N) - \exp(+U_F) + \exp(-U_F)] \qquad \text{2-Dimensional Impure (1D)}$$

$$[-(\partial^2 U/\partial x^2) - (\partial^2 U/\partial y^2)] = (q^2/\varepsilon_S k_B T)(P - N - P_F + N_F) \qquad \text{2-Dimensional Impure (1E)}$$

$$= (1/2L_D^2)[\exp(U_P - U) - \exp(U - U_N) - \exp(+U_F) + \exp(-U_F)] \qquad \text{2-Dimensional Impure (1F)}$$

$$[-(\partial^2 U/\partial X^2) - (L_D/L)^2(\partial^2 U/\partial Y^2)] = (P - N - P_F + N_F)/2n_i \qquad \text{2-Dimensional Impure (1G)}$$

$$= [\exp(U_P - U) - \exp(U - U_N) - \exp(+U_F) + \exp(-U_F)]/2 \qquad \text{2-Dimensional Impure (1H)}$$

$$[-(\partial^2 U/\partial X^2) - (L_D/L)^2(\partial^2 U/\partial y^2)] = [\exp(U_P - U) - \exp(U - U_N)]/2 \qquad \text{2-Dimensional Pure (1)}$$

Electron Current Continuity Equation

$$\nabla \cdot \boldsymbol{j}_N + q(g_N - r_N) = \partial n/\partial t \qquad \text{Time-Dependent Electron 3-D Continuity Equation (2A)}$$

$$\nabla \cdot \boldsymbol{J}_N + q(G_N - R_N) = 0 \qquad \text{DC Steady State Electron 3-D Continuity Equation (2B)}$$

$$\nabla \cdot \boldsymbol{J}_N = 0 \qquad \text{GRT-free DC Steady State Electron 3-D Continuity Equation (2C)}$$

$$\oiiint \nabla \cdot \boldsymbol{J}_N \,\mathrm{d}\mho = \oiint \boldsymbol{J}_N \cdot \mathrm{d}\boldsymbol{S} = 0 \qquad \text{Divergence Theorem, Kirchoff Law 3-D Continuity Equation (2D)}$$

$$\oiiint \nabla \cdot \boldsymbol{J}_N \,\mathrm{d}\mho = \oiint \boldsymbol{J}_N \cdot \mathrm{d}\boldsymbol{S} = \oint J_{NX}\partial yZ + \oint J_{NY}\partial xZ = 0 \qquad \text{Electron 2-D Continuity Equation (2E)}$$

$$\therefore \int J_{NX}\partial yZ + \int J_{NY}\partial xZ = \text{constant} = -I_{DN} \qquad \text{Kirchoff Law 2-D Electron Current Continuity (2)}$$

Hole Current Continuity Equation

$$\nabla \cdot \boldsymbol{j}_P + q(g_P - r_P) = \partial p/\partial t \qquad \text{Time-Dependent Hole 3-D Continuity Equation (3A)}$$

$$\nabla \cdot \boldsymbol{J}_P + q(G_P - R_P) = 0 \qquad \text{DC Steady State Hole 3-D Continuity Equation (3B)}$$

$$\nabla\cdot\boldsymbol{J}_{\mathrm{P}}=0 \qquad \text{GRT-free DC Steady State Hole 3-D Continuity Equation (3C)}$$

$$\oiiint\nabla\cdot\boldsymbol{J}_{\mathrm{P}}\,\mathrm{d}\mho=\oiint\boldsymbol{J}_{\mathrm{P}}\cdot\mathrm{d}\boldsymbol{S}=0 \qquad \text{Divergence Theorem, Kirchoff Law 3-D Continuity Equation (3D)}$$

$$\oiiint\nabla\cdot\boldsymbol{J}_{\mathrm{P}}\,\mathrm{d}\mho=\oiint\boldsymbol{J}_{\mathrm{P}}\cdot\mathrm{d}\boldsymbol{S}=\oint J_{\mathrm{P}X}\partial yZ+\oint J_{\mathrm{P}Y}\partial xZ=0 \qquad \text{Hole 2-D Continuity Equation (3E)}$$

$$\therefore\int J_{\mathrm{P}X}\partial yZ+\int J_{\mathrm{P}Y}\partial xZ=\text{constant}=-I_{\mathrm{DP}} \qquad \text{Kirchoff Law 2-D Hole Current Continuity (3)}$$

Analytical Solution of the Poisson Equation

Voltage Equation, X-Equation. (1966-Sah-Pao Model[5,24])

Integrating the Poisson Equation (1F) and (1H) by quadrature along the X-axis from X_1 to X_2 and U_1 to U_2 with the assumption of spatially constant impurity concentration, $P_{\mathrm{IM}}(x,y)=P_{\mathrm{IM}}$, we get the general solution as follows[5]. Let $X_2=X$ be a variable, then

$$(\partial U/\partial X)_2^2-(\partial U/\partial X)_1^2=F^2(U_2,U_{\mathrm{P}},U_{\mathrm{N}},U_1,P_{\mathrm{IM}},E_Y) \qquad \text{2-Dimensional Impure (4A)}$$

$$\begin{aligned}(\partial U/\partial X)^2 &= F^2(U,U_{\mathrm{P}},U_{\mathrm{N}},U_1,P_{\mathrm{IM}},E_Y)+(\partial U/\partial X)_1^2\\ &= F^2(U,U_{\mathrm{P}},U_{\mathrm{N}},U_1,P_{\mathrm{IM}},E_Y,E_{X1}) && \text{2-Dimensional Impure (4B)}\\ &= +\exp(U-U_{\mathrm{N}})-\exp(U_1-U_{\mathrm{N}})+\exp(U_{\mathrm{P}}-U)-\exp(U_{\mathrm{P}}-U_1)\\ &\quad -(P_{\mathrm{M}}/n_{\mathrm{i}})\times(U-U_1)+2(L_{\mathrm{D}}/L)^2\int(\partial^2U/\partial Y^2)\partial_X\mathrm{U}+(\partial U/\partial X)_1^2 && \text{2-Dimensional Impure (4C)}\\ &= +\exp(U-U_{\mathrm{N}})-\exp(U_1-U_{\mathrm{N}})+\exp(U_{\mathrm{P}}-U)-\exp(U_{\mathrm{P}}-U_1) && \text{1-Dimensional Pure (4)}\end{aligned}$$

Integrating the Poisson Equation (1F) and (1H) twice along X axis, from X_1 to X_2, and let $X_2=X$ as a variable, we get the general Voltage Equation or X-equation as follows[5] where

$$\iint\rho\partial x\partial x=\int\varepsilon E_X\partial x-\int\varepsilon E_{X1}\partial x+\varepsilon\iint(\partial E_Y/\partial y)\partial x\partial x=-\varepsilon(V-V_1)-(x-x_1)\varepsilon E_{X1}+\varepsilon\iint(\partial E_Y/\partial y)\partial x\partial x$$

$$\begin{aligned}&U(X,Y)-U_1(X_1,Y)=\mathrm{sign}(U-U_1)\times(C_{\varepsilon}/C_{X-X1})\times(\partial U/\partial X_1)_1\\ &-(x_{\varepsilon}/\mathrm{L})^2\iint(\partial^2U/\partial Y^2)\partial X\partial X-(1/2)(x_{\varepsilon}/L_{\mathrm{D}\varepsilon})^2\iint(\rho/qn_{\mathrm{i}})\partial X\partial X \qquad \text{2-Dimensional terms (5)}\end{aligned}$$

Analytical Solution of the Current Equations

Current Equation, Y-Equation. (1966-Sah-Pao Model[5,25])

Electron Current

$$\begin{aligned}\boldsymbol{J}_{\mathrm{N}} &= +q\mu_{\mathrm{n}}N\boldsymbol{E}+qD_{\mathrm{n}}\nabla N=-qD_{\mathrm{n}}N\nabla U_{\mathrm{N}} && \text{3-Dimensional (6A)}\\ &= J_{\mathrm{N}X}\boldsymbol{i}_X+J_{\mathrm{N}Y}\boldsymbol{i}_{\mathrm{Y}}=-qD_{\mathrm{n}}N[(\partial U_{\mathrm{N}}/\partial x)\boldsymbol{i}_X+(\partial U_{\mathrm{N}}/\partial y)\boldsymbol{i}_Y] && \text{2-Dimensional (6B)}\end{aligned}$$

$$\mathrm{J}_{\mathrm{N}X}\boldsymbol{i}_X=-qD_{\mathrm{n}}N(\partial U_{\mathrm{N}}/\partial x)\boldsymbol{i}_X=0 \quad \text{Gradual-Channel} \quad \therefore U_{\mathrm{N}}(x,y)=U_{\mathrm{N}}(y) \qquad (6\mathrm{C})$$

$$J_{\mathrm{N}Y}\boldsymbol{i}_Y=-qD_{\mathrm{n}}N(\partial U_{\mathrm{N}}/\partial y)\boldsymbol{i}_{\mathrm{Y}}\neq 0 \qquad (6\mathrm{D})$$

$$\oiint\boldsymbol{J}_{\mathrm{N}}\cdot\mathrm{d}S=\oint J_{\mathrm{N}X}\partial yZ+\oint J_{\mathrm{N}Y}\partial xZ=\oint J_{\mathrm{N}Y}\partial xZ=0 \quad \therefore\int J_{\mathrm{N}Y}\partial xZ=\text{constant}=-I_{\mathrm{DN}} \qquad (6)$$

Hole Current

$$\begin{aligned}\boldsymbol{J}_{\mathrm{P}} &= +q\mu_{\mathrm{p}}P\boldsymbol{E}+qD_{\mathrm{p}}\nabla N=-qD_{\mathrm{p}}P\nabla U_{\mathrm{P}} && \text{3-Dimensional (7A)}\\ &= J_{\mathrm{P}X}\boldsymbol{i}_X+J_{\mathrm{P}Y}\boldsymbol{i}_Y=-qD_{\mathrm{p}}N[(\partial U_{\mathrm{N}}/\partial x)\boldsymbol{i}_X+(\partial U_{\mathrm{N}}/\partial y)\boldsymbol{i}_Y] && \text{2-Dimensional (7B)}\end{aligned}$$

$$J_{\mathrm{P}X}\boldsymbol{i}_X=-qD_{\mathrm{p}}P(\partial U_{\mathrm{P}}/\partial x)\boldsymbol{i}_X=0 \quad \text{Gradual-Channel} \quad \therefore U_{\mathrm{P}}(x,y)=U_{\mathrm{P}}(y) \qquad (7\mathrm{C})$$

$$J_{\mathrm{P}Y}\boldsymbol{i}_Y=-qD_{\mathrm{p}}P(\partial U_{\mathrm{P}}/\partial y)\boldsymbol{i}_Y\neq 0 \qquad (7\mathrm{D})$$

$$\oiint\boldsymbol{J}_P\cdot\mathrm{d}\boldsymbol{S}=\oint J_{\mathrm{P}X}\partial yZ+\oint J_{\mathrm{P}Y}\partial xZ=\oint J_{\mathrm{P}Y}\partial xZ=0 \quad \therefore\int J_{\mathrm{P}Y}\partial xZ=\text{constant}=-I_{\mathrm{DP}} \qquad (7)$$

Total Terminal Current

(Flowing into the Drain and Source Nodes from Outside)

$$I_{\mathrm{D}}=I_{\mathrm{DN}}+I_{\mathrm{DP}}=-I_{\mathrm{S}} \qquad (\text{integrate } x=0 \text{ to } x_{\mathrm{B}};\ y \text{ can be any value between 0 to } L.)\ (8)$$

Since the total electron and hole currents are each spatially constant, independent of y, the integrated current densities over the cross-sectional area is equal to the current flowing out of the drain and source terminals as shown in (6) and (7). Thus, integrating the electron and hole current densities given by (6) and (7), and using the normalized $X=x/L_{\mathrm{D}}$ and $Y=y/L$, we get the following analytical solutions for the electron and hole currents flowing into the drain from the outside. The solution of (10) and (12) are obtained by integration $\int\partial Y$ from $Y=Y_3\geqslant 0$ to $Y\leqslant 1$.

$$I_{\mathrm{DN}} = -\iint J_{\mathrm{N}Y}\partial x\partial z = +qD_{\mathrm{n}}(\partial U_{\mathrm{N}}/\partial y)\int N\partial xW \tag{9A}$$

$$= +qD_{\mathrm{n}}n_{\mathrm{i}}(W/L)L_{\mathrm{D}}(\partial U_{\mathrm{N}}/\partial Y)\exp(-U_{\mathrm{N}})\int \exp(+U)\partial_X U/F(U,U_{\mathrm{P}},U_{\mathrm{N}},U_1,P_{\mathrm{IM}},E_Y,E_{X1}) \tag{9}$$

$$I_{\mathrm{DN}} = -\iiint J_{\mathrm{N}Y}\partial x\partial y\partial z/L = +qD_{\mathrm{n}}\int(\partial U_{\mathrm{N}}/\partial y)\partial y\int N\partial x(W/L) \tag{10A}$$

$$= +qD_{\mathrm{n}}n_{\mathrm{i}}(W/L)[L_{\mathrm{D}}/(Y-Y_3)]\int\partial_Y U_{\mathrm{N}}\exp(-U_{\mathrm{N}})\int \exp(+U)\partial_X U/F(U,U_{\mathrm{P}},U_{\mathrm{N}},U_1,P_{\mathrm{IM}},E_Y,E_{X1}) \tag{10}$$

$$I_{\mathrm{DP}} = -\iint J_{\mathrm{P}Y}\partial x\partial z = +qD_{\mathrm{p}}(\partial U_{\mathrm{P}}/\partial y)\int P\partial xW \tag{11A}$$

$$= +qD_{\mathrm{p}}n_{\mathrm{i}}(W/L)L_{\mathrm{D}}(\partial U_{\mathrm{P}}/\partial Y)\exp(+U_{\mathrm{P}})\int \exp(-U)\partial_X U/F(U,U_{\mathrm{P}},U_{\mathrm{N}},U_1,P_{\mathrm{IM}},E_Y,E_{X1}) \tag{11}$$

$$I_{\mathrm{DP}} = -\iiint J_{\mathrm{P}Y}\partial x\partial y\partial z/L = +qD_{\mathrm{p}}\int(\partial U_{\mathrm{P}}/\partial y)\partial y\int P\partial x(W/L) \tag{12A}$$

$$= +qD_{\mathrm{p}}n_{\mathrm{i}}(W/L)[L_{\mathrm{D}}/(Y-Y_3)]\int\partial_Y U_{\mathrm{P}}\exp(+U_{\mathrm{P}})\int \exp(-U)\partial_X U/F(U,U_{\mathrm{P}},U_{\mathrm{N}},U_1,P_{\mathrm{IM}},E_Y,E_{X1}) \tag{12}$$

Flatband Boundary of the Two Sections[26]

If the terminal voltages or electric potentials satisfy (i) $0 \leqslant U_{\mathrm{GS}} \leqslant U_{\mathrm{DS}}$ for electron-channel and electron-current initially dominant in the low drain voltages range $0 \leqslant U_{\mathrm{DS}} \leqslant U_{\mathrm{GS}}$, to be denoted as nBiFET, or (ii) $0 \geqslant U_{\mathrm{GS}} \geqslant U_{\mathrm{DS}}$ for hole-channel and hole-current initially dominant in the low drain voltage range $0 \geqslant U_{\mathrm{DS}} \geqslant U_{\mathrm{GS}}$, to be denoted as pBiFET, which is when both the electron and hole currents are in their ‘saturation’ ranges, either in subthreshold or in strong inversion, then the channel can be divided into two sections, a model first used by Shockley in 1952[2] for the volume-channel JGFET, obviously followed that of his 1949's bipolar junction transistor's collector/base/emitter section separation idea. But the 2-section volume-channel idea for JGFET, is not so difficult to apply to the surface-channel MOSFET, as the boundary point or plane where the oxide electric field reverses, which has been used as figures in textbook to teach undergraduates about MOS transistors by Sah in 1993[26] to illustrate this important location. There is also a general analytical equation, just discovered on 20080202 and reported here, that determines the flatband location or *x-z* plane, $y = y_0$, which can be obtained using the Y-equations or Current-equations given by (9A) to (12) above and noting that at this flatband plane, $N(x,y) = N(y_0) = n_{\mathrm{i}}\exp[U(y_0) - U_{\mathrm{N}}(y_0)] \equiv n_{\mathrm{i}}\exp(U_0 - U_{\mathrm{N0}})$ and $P(x,y) = P(y_0) = n_{\mathrm{i}}\exp[U_{\mathrm{P}}(y_0) - U(y_0)] \equiv n_{\mathrm{i}}\exp(U_{\mathrm{P0}} - U_0)$, i.e. both are spatially constant in x and y. Then letting $Y = Y_0$ and $Y_3 = 0$ for the electron currents in (9) and (10), and $Y = 1$ and $Y_3 = Y_0$ for the hole currents in (11) and (12), and eliminating the electron current between (9) and (10), and eliminating hole current between (11) and (12), we get the following two pairs of analytical equations to calculate the location of this flatband plane which separates the source electron-emitter section, with length Y_0, from the drain electron-collector section, with length $(1 - Y_0)$, and similarly, the source hole-collector section with length Y_0 from the drain hole-emitter section with length $(1 - Y_0)$. Of course, the four equations give the identical value of Y_0, when the recursive iterations reach convergence.

$$(y_0/x_{\mathrm{B}}) = (\partial U_{\mathrm{N0}}/\partial y)\cdot\exp(U_0 - U_{\mathrm{N0}})/\int(\partial U_{\mathrm{N}}/\partial y)\partial y\int\exp(U - U_{\mathrm{N}})\partial x\ (x = 0 \text{ to } x_{\mathrm{B}};\, y = 0 \text{ to } y_0) \tag{13}$$

$$(L - y_0)/x_{\mathrm{B}} = (\partial U_{\mathrm{N0}}/\partial y)\cdot\exp(U_0 - U_{\mathrm{N0}})/\int(\partial U_{\mathrm{N}}/\partial y)\partial y\int\exp(U - U_{\mathrm{N}})\partial x\ (x = 0 \text{ to } x_{\mathrm{B}};\, y = y_0 \text{ to } L) \tag{14}$$

$$(L - y_0)/x_{\mathrm{B}} = (\partial U_{\mathrm{P0}}/\partial y)\cdot\exp(U_{\mathrm{P0}} - U_0)/\int(\partial U_{\mathrm{P}}/\partial y)\partial y\int\exp(U_{\mathrm{P}} - U)\partial x\ (x = 0 \text{ to } x_{\mathrm{B}};\, y = y_0 \text{ to } L) \tag{15}$$

$$(y_0/x_{\mathrm{B}}) = (\partial U_{\mathrm{P0}}/\partial y)\cdot\exp(U_{\mathrm{P0}} - U_0)/\int(\partial U_{\mathrm{P}}/\partial y)\partial y\int\exp(U_{\mathrm{P}} - U)\partial x\ (x = 0 \text{ to } x_{\mathrm{B}};\, y = 0 \text{ to } y_0) \tag{16}$$

Simultaneous Equations Needed for Solving Specific Transistor Structures

There are 3 potential variables, $U(x,y)$, $U_{\mathrm{N}}(y)$ and $U_{\mathrm{P}}(y)$ which are used to obtain the terminal currents as a function of terminal voltages or electric potentials, $U(x,y)$ applied to the terminals ($x = X_\alpha$, $y = Y_\alpha$) with $a = \mathrm{S, G1, G2, D}$ which indicate the coordinate location of the source, gate 1, gate 2 and drain terminals. For any transistor structure, the problem is completely defined by specifying the electric potentials applied to these terminals. The solution or the current-voltage characteristics are obtained by knowing the internal and boundary conditions of the electric potentials and their space derivatives or the electric fields, based on the Coulomb electron physics. The electrochemical potentials of electrons and holes at the terminals and inside the device cannot be specified since they are determined by the electric potentials applied to the terminals which determine spatial distribution of the electron and hole charge concentrations, hence, their electrochemical potentials. Such fundamental errors have been made in all previous analytical and numerical solutions, traceable to Shockley's constant quasi-Fermi potential approximation for majority carriers in minority carrier devices, such as the p/n junction, although some of the MOS transistor numerical solutions turned out to be nearly correct which can be understood from the dominance of the particular charge transport phenomena which is picked out to give the analytical solution by neglecting all others via arbitrary setting of the electrochemical potentials of the others, for example, the unipolar nFET solution with hole and trapped electron electrochemical potentials set to zero or some constant, which would give zero electrical current from the holes and trapped electrons while still allowing them to change their spatial charge density distribution to meet the demand of the applied voltages or electric fields and the electrode configurations. But, when their currents become comparable to the electron current, obviously gross errors would occur in the calculated terminal currents. In the following sections, the correct internal and boundary conditions are derived and listed for the four volume-application-and-produced MOSFET structures.

3 Applications to MOS Transistors (Internal and Boundary Conditions)

For ease of presentation, we collect the previous equations that are used to completely define the solution of a specific transistor structure. They are listed and renumbered below, as the Voltage or X or electric potential $U(X,Y)$ equations and the Currents and Y or electrochemical potentials $U_N(X,Y)$ and $U_P(X,Y)$ equations. They are to be solved simultaneously, analytically as much as possible, and numerically by recursive iterations, for a given transistor structure to give the drain terminal electron and hole currents, I_{DN} and I_{DP}, as a function of the terminal electric potentials U_{G1S}, U_{G2S} and U_{DS} of gate 1, gate 2 and drain relative to source. In order to give analytical solutions, we use the thin transistor model, which Shockley called the 'gradual channel' approximation or just 'gradual approximation' but it can also be called the long channel approximation, although none can be justified for all transistor geometries and impurity distributions. Nevertheless, it gives us the start and it has provided simple, useful and easy to understand analytically results for more than 40 years since their first use by Sah during the mid-1960's[24,25]. So the best and new (just realized to write this report) physical-picture-based condition to give the iterativeable analytical solutions is to assume that both the electron and hole currents flow along the channel direction (y-axis), which were used in (6C) for the electron current namely, $J_{NX} = -qD_n N(\partial/\partial x)U_N(x,y) = 0$ or $(\partial/\partial x)U_N(x,y) = 0$, then, since $N(x,y) \neq 0$, so $U_N(x,y) = U_N(y)$, and similarly in (7C) for the hole current, $J_{PX} = -qD_p P(\partial/\partial x)U_P(x,y) = 0$ or $(\partial/\partial x)U_P(x,y) = 0$, then, since $P(x,y) \neq 0$, so $U_P(x,y) = U_P(y)$.

The general voltage and X equations from (4B) and (4C) and (5) are

$$\begin{aligned}(\partial U/\partial X)^2 &= F^2(U, U_P, U_N, U_1, P_{IM}, E_Y) + (\partial U/\partial X)_1^2 \\ &= F^2(U, U_P, U_N, U_1, P_{IM}, E_Y, E_{X1}) && \text{2-D Impure (4B)} \quad (17) \\ &= +\exp(U - U_N) - \exp(U_1 - U_N) + \exp(U_P - U) - \exp(U_P - U_1) \\ &\quad - (P_{IM}/n_i)\times(U - U_1) + 2(L_D/L)^2\int(\partial^2 U/\partial Y^2)\partial_X U + (\partial U/\partial X)_1^2 && \text{2-D Impure (4C)} \quad (18)\end{aligned}$$

$$\begin{aligned}U(X,Y) - U_1(X_1,Y) &= \mathrm{sign}(U - U_1)\times(C_\varepsilon/C_{X-X1})\times(\partial U/\partial X_1)_1 \\ &\quad - (x_\varepsilon/L)^2\iint(\partial^2 U/\partial Y^2)\partial X\partial X - (1/2)(x_\varepsilon/L_D)^2\iint(\rho/qn_i)\partial X\partial X && \text{2-D terms (5)} \quad (19)\end{aligned}$$

The general current and Y equations from (9) to (12) are

$$I_{DN} = +qD_n n_i(W/L)L_D(\partial U_N/\partial Y)\exp(-U_N)\int\exp(+U)\partial_X U/F(U, U_P, U_N, U_1, P_{IM}, E_Y, E_{X1}) \ (9) \qquad (20)$$

$$= +qD_n n_i(W/L)[L_D/(Y-Y_3)]\int\partial_Y U_N\exp(-U_N)\int\exp(+U)\partial_X U/F(U, U_P, U_N, U_1, P_{IM}, E_Y, E_{X1}) \ (10) \quad (21)$$

$$I_{DP} = +qD_p n_i(W/L)L_D(\partial U_P/\partial Y)\exp(+U_P)\int\exp(-U)\partial_X U/F(U, U_P, U_N, U_1, P_{IM}, E_Y, E_{X1}) \ (11) \qquad (22)$$

$$= +qD_p n_i(W/L)[L_D/(Y-Y_3)]\int\partial_Y U_P\exp(+U_P)\int\exp(-U)\partial_X U/F(U, U_P, U_N, U_1, P_{IM}, E_Y, E_{X1}) \ (12) \quad (23)$$

We just gave the four equations in (13) to (16) to determine the boundary, $Y = Y_0$, separating the two sections when $0 \leqslant U_{GS} \leqslant U_{DS}$ for nBiFET or $0 \geqslant U_{GS} \geqslant U_{DS}$ for pBiFET. So, they are not repeated here.

3.1 1-Gate MOSFETs

Finite-Thickness Impure-Base（SOI SIMOX and SOI TFT）

The general 1-gate MOS transistor has a finite silicon base thickness of x_B from $x=0$ to $x=x_B$ with a dielectric constant of $\varepsilon_S=\kappa_{\text{Silicon}}\varepsilon_0$ where $\varepsilon_0=8.854\times10^{-14}$ F/m is the permittivity or electric constant of free space and $\kappa_{\text{Si}}=11.7$ is the relative dielectric constant of silicon. The 1-gate or front gate is composed of a gate conductor over a gate insulator or oxide of thickness x_O with dielectric constant of $\varepsilon_O=3.9\varepsilon_0$, which is located from $x=-x_O$ to $x=0$ and the oxide/silicon interface plane is located at $x=0$. Oxide and interface trapped charges can be readily included in the general voltage equation but is implicitly understood as part of the flat-band gate voltage which is absorbed into the applied gate voltage. The distance between the drain and source contact is L, which is the active channel length. The back surface, at plane $x=x_B$ at all y from $y=0$ to $y=L$, is taken as a free surface or a surface covered by an infinite thick oxide with no interface charge. Thus, $U(x=x_B,y)=U_0(y)$ and $E_X=(k_BT/q)[(\partial/\partial x)U(x=x_B,y)]=0$. One cannot and must not assign a condition for the second derivative, $\partial^2U(x,y)/\partial x^2$ at the back interface, $x=x_B$, or any (x,y) since it not only forces electrical neutrality at all y at the back interface which must be determined by the applied potential to the terminals which determines the spatial variation of the electron and hole concentrations, which is determined by the Poisson Equation, but such a assumption on $\partial^2U(x,y)/\partial x^2$ also over-specifies the problem which usually cannot find a physically realistic situation to give such specification. Furthermore, the insulating back boundary at $x=x_B$ is indeed not inconsistent with the general assumption that the electron and hole currents both flow in the y direction, i.e. at $x=x_B$, both $\partial U_N/\partial x=0$ and $\partial U_P/\partial x=0$. Finally, the most important, is the location of the source and drain contacts for the applied potentials to the drain and source. They are at $(x_B,y=L)$ for the drain and $(x_B,y=0)$ for the source. We do not specify how these are connected to the external lead, which is covered by the contact theory. Therefore the current-voltage solutions are those limited by the ‘intrinsic’ transistor with perfect drain and source contacts or with very long ‘intrinsic’ transistor which is current limiting to dominate the electrical characteristics while the source and drain contacts appear as nearly short circuits or zero resistances. In real transistors, the contacts must be taken into account in the analysis, which has never been taken into account in the transistor theory, only treated as an after-thought as a series resistance or a junction diode, one each to the source and the drain, which is not a bad approximation but which is increasingly poor as one approaches the pure-base with one or two gates. So, based on the above, with the source as the reference for the potentials applied to the terminals, $U_{GS}=qV_{GS}/k_BT$ and $U_{DS}=qV_{DS}/kT$, the X-equations for this 1-Gate thin-base or finite-thickness-base but impure, MOS transistor (with $\rho_{OX}=0$ and $\rho_{\text{INTERFACE}}=0$), valid for all y values between the source $(x=x_B,y=0)$ and the drain $(x=x_B,y=L)$, are

$$U_{GS}-U_S=\text{sign}(U_S-U_0)\times(C_D/C_O)\times(\partial U/\partial X)_S-(x_O/L)^2\iint(\partial^2U/\partial Y^2)\partial X\partial X \text{ from (19)} \tag{24}$$

$$\begin{aligned}(\partial U/\partial X)^2=F^2(U,U_P,U_N,U_0,P_{IM},E_Y)&=+\exp(U-U_N)-\exp(U_0-U_N)+\exp(U_P-U)-\exp(U_P-U_0)\\&\quad-(P_{IM}/n_i)\times(U-U_0)+2(L_D/L)^2\int(\partial^2U/\partial Y^2)\partial_XU \text{ from (17)}\end{aligned} \tag{25}$$

$$X_B=\int\text{sign}(U-U_0)\partial_XU/F \text{ (Integrated from } U=U_0 \text{ to } U=U_S\text{) (The Thickness Equation)} \tag{26}$$

From the current equations (20) to (23), with $U=U_{DS}$, $E_{X1}=0$, $Y_3=0$ and $Y=1$, we have

$$I_{DN}=+qD_nn_i(W/L)L_D(\partial U_N/\partial Y)\exp(-U_N)\int\exp(+U)\partial_XU/F(U,U_P,U_N,U_0,P_{IM},E_Y) \text{ from (20)} \tag{27}$$

$$=+qD_nn_i(W/L)L_D\int\partial_YU_N\exp(-U_N)\int\exp(+U)\partial_XU/F(U,U_P,U_N,U_0,P_{IM},E_Y) \text{ from (21)} \tag{28}$$

$$I_{DP}=+qD_pn_i(W/L)L_D(\partial U_P/\partial Y)\exp(+U_P)\int\exp(-U)\partial_XU/F(U,U_P,U_N,U_0,P_{IM},E_Y) \text{ from (22)} \tag{29}$$

$$=+qD_pn_i(W/L)L_D\int\partial_YU_P\exp(+U_P)\int\exp(-U)\partial_XU/F(U,U_P,U_N,U_0,P_{IM},E_Y) \text{ from (23)} \tag{30}$$

To give the internal variations, i.e. the position dependence of the four potentials, $U_N(y)$, $U_P(y)$, $U_S(y)$ and $U_0(y)$, and their derivatives, $(\partial/\partial Y)$, we use (20) to (23), after the convergent electron and hole currents, I_{DN} and I_{DP}, are obtained. Let $Y_3=0$, then the Y variations are given by the following where the ∂Y integration is carried from $Y=0$ at the source to a point Y in the channel between the source at $Y=0$ and drain at $Y=1$.

$$I_{DN}=+qD_nn_i(W/L)L_D(\partial U_N/\partial Y)\exp(-U_N)\int\exp(+U)\partial_XU/F(U,U_P,U_N,U_0,P_{IM},E_Y) \text{ (20)} \tag{31}$$

$$= + qD_n n_i (W/L)(L_D/Y)\int \partial_Y U_N \exp(-U_N)\int \exp(+U)\partial_X U / F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (21) \tag{32}$$

$$I_{DP} = + qD_p n_i (W/L) L_D (\partial U_P/\partial Y)\exp(+U_P)\int \exp(-U)\partial_X U / F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (22) \tag{33}$$

$$= + qD_p n_i (W/L)(L_D/Y)\int \partial_Y U_P \exp(+U_P)\int \exp(-U)\partial_X U / F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (23) \tag{34}$$

The problem is now solved. There are just enough equations to compute the y variations of the four potentials, $U_N(y)$, $U_P(y)$, $U_S(y)$, $U_0(y)$, and to obtain the electron and hole currents from the double integration given by (28) and (30) or (32) and (34) from $y = 0$ to $y = L$ using the boundary condition for the electric potential $U(x = x_B, y = 0) = U_{SB} = 0$ and $U(x = x_B, y = L) = U_{DB} = U_{DS}$ = voltage applied to the drain contact relative to the source contact. There is no other internal and boundary condition to be specified. For example, the y-integration of (28) to give the electron current is carried out by integrating the electron electrochemical potential $U_N(y)$ from $y = 0$ to $y = L$. This integration variable can be transformed analytically or numerically to $U_0(y)$ since the terminal values of $U_0(y = 0)$ and $U_0(y = L)$ are known, or alternatively, the terminal values of $U_N(y = 0)$ and $U_N(y = L)$ can be computed from the voltage equation (24) and the thickness equation (26), which, with (25), (27) = (28) and (29) = (30), give $U_0(y) = U_0[U_N(y)]$, $U_S(y) = U_S[U_N(y)]$ and $U_P(y) = U_P[U_N(y)]$. The important point to note is that the hole current equations (29) = (30) provides the missing boundary condition that was substituted by the over-specified electrical neutrality condition[5,11,27] and the constant U_P also in y, such as setting $U_P = 0$[11,27], or $(U_P + U_N)/2$ = constant = 0 at all locations, resulting in incorrect solutions, not only when the neglected hole current is comparable to the electron current but also when the neglected hole current is small but still significantly alters the spatial distribution of holes to modify the spatial distribution of the electrons and hence the electron currents, in both subthreshold and strong inversion.

The strategy for numerical solution could be one of recursive iterations by starting with the unipolar solution, namely, for the electron channel or $V_{GS} \geqslant 0$ and $V_{DS} \geqslant 0$, by solving the electron equations given by (24) to (28) using (31) and (32) with some assumed values for U_P and $\partial U_P/\partial Y$, such as zero as some have done, or $U_P = U_F$ as many have done, or just plainly drop P all together[11], all completely missed the presence of the hole currents. Then, the unipolar solution for $U_N(y)$, $U_S(y)$ and $U_0(y)$ and the single integration in X or $\partial_X U$ and double integrations in ∂X and ∂Y or $\partial_X U$ and $\partial_Y U_N$ to get the electron current can be carried out analytically or numerically, as the zeroth-order solution. Similarly, the unipolar zeroth order solution for $U_P(y)$ can be obtained, which can then be used to get the first order solution of $U_N(y)$, $U_S(y)$ and $U_0(y)$. This recursive iteration can be continued until reaching ‘external’ convergence, as defined by a less than 1% or 0.1% change of the electron current, I_{DN}, and the hole current, I_{DP}, in the next iteration.

These results for the impure base can be readily applied to pure base, by letting $P_{IM} = 0$.

Infinite-Thickness Impure-Base (Bulk MOSFET)

The preceding results are readily applied to the bulk MOSFET, i.e. infinite or very thick base of a 12-inch silicon wafer, which has been the basic building block (BBB) (First used in Ref. [21].) device of all integrated circuits. Our previous analysis showed that the characteristic length of the FET is the Debye screening length of the semiconductor, which for the pure silicon base is about 25μm at room temperature, and which decreases as the impurity concentration increases, by the ratio of $(2n_i/P_{IM})^{1/2}$. So, for the traditional bulk MOSFET with ion-implanted impure base layer of $P_{IM} = 10^{17\sim18}\text{cm}^{-3}$ and a room temperature intrinsic carrier concentration of $n_i = 10^{10}\text{cm}^{-3}$, the impure Debye length drops to $25\mu\text{m} \times (2 \times 10^{10}/2 \times 10^{18})^{1/2} = 2.5\text{nm}$. Thus, the current nanometer technology of 10 to 50nm dimensions would still qualify as long and thick transistors. For this case, $E_X(x = x_B \rightarrow \infty, y) = 0$ still applies, and $U_0(x = x_B \rightarrow \infty, y) = U_0(y)$ can be assumed to take some boundary values determined by the contacts, such as $U_0(y = 0) = U_{SB}$ and $U_0(y = L) = U_{DB}$ which replaces the meaningless or no longer viable thickness equation (26) so there is still enough conditions for the number of unknowns. In the past, it has also been assumed that there is no hole current flowing out through the terminals, so $\partial U_P(y)/\partial y = 0$ or $U_P(y) = U_P$ = constant such as U_F. Alternative boundary conditions could be $U_0(x = x_B \rightarrow \infty, y = 0) = 0$ and $U_S(x = 0, y = L) = U_{DB}$. Another pair was $U_0(x = x_B \rightarrow \infty, y) = 0$ and $\partial^2 U(x = x_B \rightarrow \infty, y)/\partial X^2 = 0$, the so-called remote charge neutrality condition to remove the imaginary electric field near flatband[5]. Thus, the current-potential equations from finite-base-thickness (24) to (30) reduces to the following for the infinite-base-thickness Bulk transistor, with $U_0 = 0$ and the thickness equation removed. Or, more generally, for a body or source bias of $U_0(y) = U_{SB}$ = constant.

$$U_{GS} - U_S = \text{sign}(U_S - U_0)\times(C_D/C_O)\times(\partial U/\partial X)_S - (x_O/L)^2\int_{\text{oxide}}(\partial^2 U/\partial Y^2)\partial X\partial X \text{ from (24)} \tag{35}$$

$$\begin{aligned}(\partial U/\partial X)^2 &= F^2(U, U_P, U_N, U_0, P_{IM}, E_Y) = +\exp(U - U_N) - \exp(U_0 - U_N) + \exp(U_P - U) - \exp(U_P - U_0)\\ &\quad - (P_{IM}/n_i)\times(U - U_0) + 2(L_D/L)^2\int(\partial^2 U/\partial Y^2)\partial_X U \text{ from (25)}\end{aligned} \tag{36}$$

From the current equations, and set $U = U_{DS}$, $E_{X1} = 0$, $Y_3 = 0$ and $Y = 1$, we have the boundary equations or terminal current-voltage equations given by

$$I_{DN} = +qD_n n_i(W/L)L_D(\partial U_N/\partial Y)\exp(-U_N)\int\exp(+U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y) \text{ from (27)} \tag{37}$$

$$= +qD_n n_i(W/L)L_D\int\partial_Y U_N\exp(-U_N)\int\exp(+U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y) \text{ from (28)} \tag{38}$$

$$I_{DP} = +qD_p n_i(W/L)L_D(\partial U_P/\partial Y)\exp(+U_P)\int\exp(-U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y) \text{ from (29)} \tag{39}$$

$$= +qD_p n_i(W/L)L_D\int\partial_Y U_P\exp(+U_P)\int\exp(-U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y) \text{ from (30)} \tag{40}$$

To give the internal variations, i.e. the position dependence of the four potentials, $U_N(y)$, $U_P(y)$, $U_S(y)$ and $U_0(y)$, and their derivatives, $(\partial/\partial Y)$, we use (20) to (23), after the convergence of the electron and hole currents, I_{DN} and I_{DP}, are obtained. With $E_{X1} = 0$ and let $Y_3 = 0$, then the Y variations are given by the following where the ∂Y integration is carried from $Y = 0$ at the source to a point Y in the channel between the source at $Y = 0$ and drain at $Y = 1$.

$$I_{DN} = +qD_n n_i(W/L)L_D(\partial U_N/\partial Y)\exp(-U_N)\int\exp(+U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y) \text{ (20)} \tag{41}$$

$$= +qD_n n_i(W/L)(L_D/Y)\int\partial_Y U_N\exp(-U_N)\int\exp(+U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y) \text{ (21)} \tag{42}$$

$$I_{DP} = +qD_p n_i(W/L)L_D(\partial U_P/\partial Y)\exp(+U_P)\int\exp(-U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y) \text{ (22)} \tag{43}$$

$$= +qD_p n_i(W/L)(L_D/Y)\int\partial_Y U_P\exp(+U_P)\int\exp(-U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y) \text{ (23)} \tag{44}$$

3.2 2-Gate MOSFETs

Finite-Thickness Impure-Base

The results just obtained for thick and thin 1-Gate can be easily extended to the 2-Gate transistor. We apply the general results, listed in (17) to (23), with the two gates labeled by 1 in the front at $x = 0$ and 2 in the back at $x = x_B$, thus $U_2 = U_{S2}$, and we apply the same kind of boundary condition to the second gate as we have applied to the first gate. Thus, by inspection of (17) to (23), we get the following results for the 2-Gate Impure-Base asymmetrical transistor.

$$(\partial U/\partial X)^2 = F^2(U, U_P, U_N, U_2, P_{IM}, E_Y) + (\partial U/\partial X)_2^2 = F^2(U, U_P, U_N, U_2, P_{IM}, E_Y, E_{X2}) \text{ from (17)} \tag{45}$$

$$\begin{aligned}&= +\exp(U - U_N) - \exp(U_{S2} - U_N) + \exp(U_P - U) - \exp(U_P - U_{S2})\\ &\quad - (P_{IM}/n_i)\times(U - U_{S2}) + 2(L_D/L)^2\int(\partial^2 U/\partial Y^2)\partial_X U + (\partial U/\partial X)_2^2 \text{ from (18)}\end{aligned} \tag{46}$$

$$\begin{aligned}U_{G1}(X, Y) - U_{S1}(Y) &= \text{sign}(U_{G1} - U_{S1})\times(C_D/C_O)\times(\partial U/\partial X)_1\\ &\quad - (x_{O1}/L)^2\int_{\text{oxide1}}(\partial^2 U/\partial Y^2)\partial X\partial X - (1/2)(x_{O1}/L_D)^2\iint(\rho_{\text{oxide1}}/qn_i)\partial X\partial X \quad \text{from (19)}\end{aligned} \tag{47}$$

$$\begin{aligned}U_{G2}(X, Y) - U_{S2}(Y) &= \text{sign}(U_{G2} - U_{S2})\times(C_D/C_O)\times(\partial U/\partial X)_2\\ &\quad - (x_{O2}/L)^2\int_{\text{oxide2}}(\partial^2 U/\partial Y^2)\partial X\partial X - (1/2)(x_{O2}/L_D)^2\iint(\rho_{\text{oxide2}}/qn_i)\partial X\partial X \quad \text{from (19)}\end{aligned} \tag{48}$$

$$X_B = \int\text{sign}(U - U_{S2})\partial_X U/F \text{ (Integrated from } U = U_{S1} \text{ to } U = U_{S2}\text{) (The Thickness Equation)} \tag{49}$$

The general current and Y equations from (20) to (23) or (9) to (12) are

$$I_{DN} = +qD_n n_i(W/L)L_D(\partial U_N/\partial Y)\exp(-U_N)\int\exp(+U)\partial_X U/F(U, U_P, U_N, U_{S2}, P_{IM}, E_Y, E_{X2}) \text{ (9)} \tag{50}$$

$$= +qD_n n_i(W/L)L_D\int\partial_Y U_N\exp(-U_N)\int\exp(+U)\partial_X U/F(U, U_P, U_N, U_{S2}, P_{IM}, E_Y, E_{X2}) \text{ (10)} \tag{51}$$

$$I_{DP} = +qD_p n_i(W/L)L_D(\partial U_P/\partial Y)\exp(+U_P)\int\exp(-U)\partial_X U/F(U, U_P, U_N, U_{S2}, P_{IM}, E_Y, E_{X2}) \text{ (11)} \tag{52}$$

$$= +qD_p n_i(W/L)L_D\int\partial_Y U_P\exp(+U_P)\int\exp(-U)\partial_X U/F(U, U_P, U_N, U_{S2}, P_{IM}, E_Y, E_{X2}) \text{ (12)} \tag{53}$$

The y-variation equations to give $U_N(y)$, $U_P(y)$, $U_{S1}(y)$ and $U_{S2}(y)$, and their derivatives, $\partial/\partial y$, are just the double integration of (50) and (52) multiplied by ∂Y and evaluated at Y, invoking the y-independence of I_{DN} and I_{DP}.

$$I_{DN} = + qD_n n_i (W/L)(L_D/Y)\int \partial_Y U_N \exp(-U_N)\int \exp(+U)\partial_X U/F(U, U_P, U_N, U_{S2}, P_{IM}, E_Y, E_{X2})\ (10) \tag{54}$$

$$I_{DP} = + qD_p n_i (W/L) L_D/(1-Y)\int \partial_Y U_P \exp(+U_P)\int \exp(-U)\partial_X U/F(U, U_P, U_N, U_{S2}, P_{IM}, E_Y, E_{X2})\ (12) \tag{55}$$

The flatband boundary separating the two sections is given by (13) to (16) with (45) for $\partial U/\partial X$.

Symmetrical-Gate Thin Pure-Base (FinFET)

This is a special case with simpler solutions than the general asymmetrical case just described. It is the FinFET geometry that has been analyzed by many[11] due to its technological importance or promise[12]. Due to the complete symmetry between the two gates, including applied gate voltages which contain the offset from workfunction differences of the contacts, by virtue of connecting the two gates, the solution is symmetrical with respect to the bisection plane, $x = x_B/2$ where $U(x = x_B/2, y) = U_0(y)$ is a maximum (when $U_{GS} > 0$) or minimum (when $U_{GS} < 0$) and at this plane, we have $\partial U(x = x_B/2, y)/\partial x = 0$. The symmetry of two gates also gives $U_{S1}(y) = U_{S2}(y) = U_S(y)$, while $\partial U_{S1}/\partial X = -\partial U_{S2}/\partial X$. The solutions are obtained by treating just half of the thickness of the channel, $x = 0$ to $x = x_B/2$, giving the following equations from the general results just obtained for the asymmetrical gates. The dominant 2-D term is the longitudinal field gradient given by (56).

$$2(L_D/L)^2\int(\partial^2 U/\partial Y^2)\partial_X U \quad \text{(Dominant 2-D term)} \tag{56}$$

$$\begin{aligned}(\partial U/\partial X)^2 &= F^2(U, U_P, U_N, U_0, P_{IM}, E_Y) = +\exp(U-U_N) - \exp(U_0-U_N) + \exp(U_P-U) - \exp(U_P-U_0)\\ &\quad - (P_{IM}/n_i)\times(U-U_0) + 2(L_D/L)^2\int(\partial^2 U/\partial Y^2)\partial_X U \text{ from (46)}\end{aligned} \tag{57}$$

$$U_{GS} - U_S = \text{sign}(U_S - U_0)\times(C_D/C_O)\times(\partial U/\partial X)_S \quad \text{(1-D term)} \tag{58}$$

$$- (x_O/L)^2\int_{\text{oxide}}(\partial^2 U/\partial Y^2)\partial X\partial X - (1/2)(x_O/L_D)^2\iint(\rho_{\text{oxide}}/qn_i)\partial X\partial X \text{ (2-D terms) from (47)} \tag{59}$$

$$X_B = 2\int \text{sign}(U - U_0)\partial_X U/F \text{ (Integrated from } U = U_0 \text{ to } U = U_S\text{) (The Thickness Equation)} \tag{60}$$

The general current and Y equations are

$$I_{DN} = + qD_n n_i (W/L) L_D (\partial U_N/\partial Y)\exp(-U_N)\int \exp(+U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (9) \tag{61}$$

$$= + qD_n n_i (W/L) L_D \int \partial_Y U_N \exp(-U_N)\int \exp(+U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (10) \tag{62}$$

$$I_{DP} = + qD_p n_i (W/L) L_D (\partial U_P/\partial Y)\exp(+U_P)\int \exp(-U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (11) \tag{63}$$

$$= + qD_p n_i (W/L) L_D \int \partial_Y U_P \exp(+U_P)\int \exp(-U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (12) \tag{64}$$

The y-variation equations to give $U_N(y)$, $U_P(y)$, $U_S(y)$ and $U_0(y)$, and their derivatives, $\partial/\partial y$, are just (61) and (63) multiplied by ∂Y and evaluated at Y, invoking the y-independence of I_{DN} and I_{DP}.

$$I_{DN} = + qD_n n_i (W/L) L_D/(Y)\int \partial_Y U_N \exp(-U_N)\int \exp(+U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (10) \tag{65}$$

$$I_{DP} = + qD_p n_i (W/L) L_D/(1-Y)\int \partial_Y U_P \exp(-U_P)\int \exp(-U)\partial_X U/F(U, U_P, U_N, U_0, P_{IM}, E_Y)\ (12) \tag{66}$$

The location of the flatband plane, $y = y_0$, is given by (13) to (16) which are

$$(y_0/x_B) = (\partial U_{N0}/\partial y)\cdot\exp(U_0 - U_{N0}) \Big/ \int(\partial U_N/\partial y)\partial y\int\exp(U - U_N)\partial x \ (x = 0 \text{ to } x_B; y = 0 \text{ to } y_0) \tag{67}$$

$$(L - y_0)/x_B = (\partial U_{N0}/\partial y)\cdot\exp(U_0 - U_{N0}) \Big/ \int(\partial U_N/\partial y)\partial y\int\exp(U - U_N)\partial x \ (x = 0 \text{ to } x_B; y = y_0 \text{ to } L) \tag{68}$$

$$(L - y_0)/x_B = (\partial U_{P0}/\partial y)\cdot\exp(U_{P0} - U_0) \Big/ \int(\partial U_P/\partial y)\partial y\int\exp(U_P - U)\partial x \ (x = 0 \text{ to } x_B; y = y_0 \text{ to } L) \tag{69}$$

$$(y_0/x_B) = (\partial U_{P0}/\partial y)\cdot\exp(U_{P0} - U_0) \Big/ \int(\partial U_P/\partial y)\partial y\int\exp(U_P - U)\partial x \ (x = 0 \text{ to } x_B; y = 0 \text{ to } y_0) \tag{70}$$

4 Summary

The general bipolar DC current-voltage solutions of the bipolar field effect transistor are derived rigorously based on the electrostatic theory from the Coulomb Law, which avoids over-specifying the internal and boundary conditions, such as the erroneous traditional and recent assumptions made on the spatial variations of the electrochemical potentials and electrical neutrality. The internal and boundary conditions of only the electric potential and its gradient or the electric field can be specified. The divergence of the electric field is not and cannot be specified since it is determined by the spatial distribution of the mobile and fixed charges. The general analytical solutions are obtained for MOS transistors with 1-gate and asymmetrical and symmetrical 2-gates, pure and impure, and thick and thin base. The general analytical solution is employed to obtain the current-voltage equations by imposing the boundary conditions of only the electric potential for four popular MOS transistor structures, including the Bulk MOSFET, the SOI on bulk Silicon, the SOI TFT on glass slides, and the FinFET.

Acknowledgment We thank Professor Xing Zhou (Nanyang Technological University, Singapore) who suggested this problem and has continued to suggest and consult about this work. We would also like to thank Professors 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 (University 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 and ICCDCS.

References

[1] Chih-Tang Sah, "Evolution of the MOS Transistor-from Conception to VLSI," IEEE Proceedings, 76(10), 1280－1326, October 1988.

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

[3] 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.

[4] George C. Dacey and Ian M. Ross, "Unipolar 'Field-Effect' Transistor," Proceedings of the IRE, 41(8), 970－979, August, 1953.

[5] Chih-Tang Sah and Bin B. Jie, "A History of MOS Transistor Compact Modeling," Keynote. Technical Proceedings Workshop on Compact Modeling (WCM), Abstract on pp. 1－2 and full text on pp. 349－390. Editors: Xing Zhou, Matthew Laudon and Bart Romanowicz. NSTI Nanotech 2005. The NSTI Nanotechnology Conference and Trade Show, May 8－12, 2005. Nano Science and Technology Institute, Cambridge, MA 02139, USA. # PCP05040394. On-line: http://www. nsti. org/publ/Nanotech2005WCM/1429. pdf

[6] J. Watts, C. McAndrew, C. Enz , C. Galup-Montoro, G. Gildenblat, C. Hu, R. van Langevelde, M. Miura-Mattausch, R. Rios and C.-T. Sah, "Advanced Compact Models for MOSFETs," Technical Proceedings Workshop on Compact Modeling (WCM), 3－12. Editors: Xing Zhou, Matthew Laudon and Bart Romanowicz. NSTI Nanotech 2005. The NSTI Nanotechnology Conference and Trade Show, May 8－12, 2005. Nano Science and Technology Institute, Cambridge, MA 02139, USA. # PCP05040394. On-line: See [5] above for address.

[7] The thin film transistor of semiconductor on insulator (TFT SOI) was the original form of Lilienfeld's 1930 transistor. Recent literature has described the large-dimension single-MOS-gate on evaporated-thin-silicon-film transistors (Silicon on Insulator Thin-Film Transistor, SOI TFT) to drive the liquid crystal light modulator cells of the large-area Liquid Crystal Display (LCD) panel for computers and televisions with 1080 × 1920 × 3color = 6-Million or more LCD cells and transistors and in other display applications. See [8] and recent references to be cited.

[8] Tiao-Yuan Huang, I-Wei Wu, Alan G. Lewis, Anne Chang, and Richard H. Bruce, "A Simpler 100-V Polysilicon TFT with Improved Turn-On Characteristics," IEEE Electron Device Letters, 11(6), 244－246, June 1990.

[9] Ghavam G. Shahidi, Carl A. Anderson, Barbara A. Chappell, Terry I. Chappell, James H. Comfort, Bijan Davari, Robert H. Dennard, Robert L. Franch, Patricia A. McFarland, James S. Neely, Tak H. Ning, Michael R. Polcari and James D. Warnock, "A Room Temperature 0. 1μm CMOS on SOI," IEEE Transaction on Electron Devices, 41(12), 2405－2412, December 1994.

[10] Terukazu Ohno, Mitsutoshi Takahashi, Yuichi Kado and Toshiaki Tsuchiya, "Suppression of Parasitic Bipolar Action in Ultra-Thin-Film-Depleted CMOS/SIMOX Devices by Ar-Ion Implantation into Source/Drain Regions," IEEE Transaction on Electron Devices, 45(5), 1071－1076, May 1998.

[11] 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.

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

[13] T. Hoffmann, G. Doornbos, I. Ferain, N. Collaert, P. Zimmerman, M. Goodwin, R. Rooyackers, A. Kottantharayil, Y. Yim, A. Dixit, K. De Meyer, M. Jurczak, and S. Biesemans, "GIDL (Gate-Induced Drain Leakage) and Parasitic Schottky Barrier Leakage Elimination in Aggressively Scaled HfO2/TiN FinFET Devices." IEDM Technical Digest, 743－746, Dec. 2005.

[14] Chih-Tang Sah and Bin B. Jie, "Double-Gate Thin-Base MOS Transistor: The Correct Theory," and Bin B. Jie and Chih-Tang Sah, "Double-Gate Thin-Base MOS Transistor: Characteristics for the Short Channel," Late-News Papers Presented on May 23, 2007 at the Workshop on Compact Modeling (WCM20070523) of the NSTI Nanotechnology Conference and Trade Show, May 22－24, 2007, hosted by the Nano Science and Technology Institute, Cambridge, MA 02139, USA. http://www. nsti. org/Nanotech2007/WCM2007/.

[15] Bin B. Jie and Chih-Tang Sah, "Bipolar Theory of MOS Field-Effect Transistors and Experiments," Invited Paper, IEEE Electron Device Society, Mini-Colloquium NADE-Nano-electronic Devices Present and Perspectives, Sinaia, Bucharest, Romania, October 14, 2007. Twenty-eight PowerPoint slides were displayed during the internet-presentation by Sah which were reproduced in the Conference Proceedings. Editor: Marcel D. Profirescu, P. O. Box

57-112, Bucharest 74500, Romania. profires@edil.pub.ro

[16] Chih-Tang Sah and Bin B. Jie, "Bipolar Theory of MOS Field-Effect Transistors and Experiments," Chinese Journal of Semiconductors, 28(10), 1534－1540, October 2007.

[17] 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.

[18] Chih-Tang Sah and Bin B. Jie, "The Bipolar Field-Effect Transistor: II. Drift-Diffusion Current Theory (Two-MOS-Gates on Pure-Base)," Chinese Journal of Semiconductors, 28(12), 1849－1859, December 2007.

[19] 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 (same as Chinese Journal of Semiconductors), 29(1), 1－11, January 2008.

[20] Binbin Jie and Chihtang Sah, "The Bipolar Field-Effect Transistor: IV. Short Channel Drift-Diffusion Current Theory (Two-Gates on Pure-Base)" Journal of Semiconductors, 29(2), 193－200, February 2008.

[21] 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 on 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. For the Basic Building Block (BBB) devices and circuits, see Section 670－674 on pages 596 to 643, Circuit Applications of MOSFET. BBB was first used in Section 663 starting on page 583 on the Charge Transfer Gate and later in Section 672 starting on page 614 on the MOS Inverter Circuits.

[22] W. Shockley, "Hot Electrons in Germanium and Ohm's Law," Bell System Technical Journal, 30, 990－1034, October 1951. Hot Electrons and the Ohmic Law based on the Maxwellian distribution of carrier concentration in energy or the exponential representation was first used in this 1951 article to analyze the high electric field effects on electron mobility.

[23] William Shockley, "Problems Related to p-n Junctions in Silicon," Solid-State Electronics, 2(1), 35－67, January 1, 1961. The Shockely Lucky Electron Model or Ballistic Model was first presented in this 1961 article to analyze the energetic electrons (and holes) accelerated by high electric fields to high kinetic energies characteristic of the energy of the phonons that are emitted during the electron-phonon scattering transition.

[24] Chih-Tang Sah and Henry C. Pao, "The Effects of Fixed Bulk Charge on the Characteristics of Metal-Oxide-Semiconductor Transistor," IEEE Trans Elec Dev, 13(4), 410－415, April 1966. [24] and [25] have been referred to by compact modelers as 1966-Sah-Pao.

[25] Henry C. Pao and Chih-Tang Sah, "The Effects of Diffusion Current on the Characteristics of MOS Transistors," Solid-State Electronics, 9(10), 927－938, October 1, 1966. (1966-Pao-Sah)

[26] Chih-Tang Sah, Fundamentals of Solid-State Electronics-Study Guide, 423pp, World Scientific Publishing Company, Singapore, New Jersey, London. 1993. The Shockley 1952 two-section-model applied to the MOSFET using the flatband boundary to separate the two sections, is shown in "Appendix B, Examples of Sub-Half-Micron ULSI Transistors," Fig. B2.5(b) on p. 411 for nMOST and Fig. B2.6(c) on p. 412 for pMOST.

[27] Xing Zhou, Zhaomin Zhu, Subhash C. Rustagi, Guan Huel See, Guojun Zhu, Shihuan Lin, Chengquing Wei, and Guan Hui Lim, "Rigorous Surface-Potential Solution for Undoped Symmetric Double-Gate MOSFETs Considering Both Electrons and Holes at Quasi Nonequilibrium," IEEE Transaction on Electron Devices, 55(2), 616－623, February 2008.

场引晶体管理论:XI.双极电化电流(薄及厚、纯及不纯基体,单及双 MOS 栅极)*,**

薩支唐[1,2,3,†]　揭斌斌[3,†]

(1 中国科学院外籍院士,北京 100864)
(2 佛罗里达大学,佛罗里达州,Gainesville FL 32605,美国)
(3 北京大学,北京 100871)

摘要: 场引晶体管本质双极,包括电子和空穴表面和体积沟道和电流,一或多个外加横向控制电场.自 1952 年 Shockley 发明,55 年来它被认为单极场引晶体管,因电子电流理论用多余内部和边界条件,不可避免忽略空穴电流.多余条件,诸如电中性和常空穴电化电势,导致仅用电子电流算内部和终端电学特性的错误解.当忽略的空穴电流与电子电流可比,可在亚阈值区和强反型区,错误解有巨大误差.本文描述普适理论,含有电子和空穴沟道和电流.用 z 轴宽度方向均匀的直角平行六面体(x,y,z)晶体管,薄或厚、纯或杂基体,一或二块 MOS 栅极,描述两维效应及电势、电子空穴电化电势的正确内部和边界条件.没用多余条件,导出四种常用 MOS 晶体管,直流电流电压特性完备解析方程:半无限厚不纯基上一块栅极(传统的 Bulk MOSFET),与体硅以氧化物绝缘的不纯硅薄层上一块栅极(SOI),在沉积到绝缘玻璃的不纯硅薄层上一块栅极(SOI TFT),和薄纯基上两块栅极(FinFETs).

关键词: 双极场引晶体管理论;MOS 场引晶体管;并存电子和空穴表面和体积沟道和电流;表面势;两区短沟道理论;双栅不纯基理论

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)03-0397-13

---

* 该研究及揭斌斌由 CTSAH Associates (CTSA) 资助.CTSA 由萨故夫人张淑南创建.纪念她七十周年.

** 萨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

2008-02-22 收到,2008-02-27 定稿