EQUIVALENT-CIRCUIT MODELING OF THE
LARGE-SIGNAL TRANSIENT RESPONSE
OF FOUR-TERMINAL MOS FIELD EFFECT TRANSISTORS

By

JOSE IGNACIO ARREOLA

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1978
to
IGNACIO
and
CELIA,
my parents
ACKNOWLEDGMENTS

I am deeply indebted to Prof. Fredrik A. Lindholm for his contribution to this work and for his continued guidance, support and encouragement. I also wish to thank Dr. D. R. MacQuigg for his help in doing experimental measurements and for many interesting discussions. I would like to express my appreciation to Prof. A. D. Sutherland for allowing me to study the results of his two-dimensional calculations which broadened my understanding of the MOSFET.

The financial support of Consejo Nacional de Ciencia y Tecnología (México) throughout this work is gratefully acknowledged. I must also thank Mrs. Vita Zamorano for her careful typing of the manuscript. Finally, I owe a special debt of gratitude to my wife, Patricia, for her forbearance and encouragement.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACKNOWLEDGMENTS</strong></td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
</tr>
<tr>
<td><strong>I</strong></td>
</tr>
<tr>
<td><strong>II</strong></td>
</tr>
<tr>
<td>2.1</td>
</tr>
<tr>
<td>2.2</td>
</tr>
<tr>
<td>2.3</td>
</tr>
<tr>
<td>2.4</td>
</tr>
<tr>
<td><strong>III</strong></td>
</tr>
<tr>
<td>3.1</td>
</tr>
<tr>
<td>3.1.1</td>
</tr>
<tr>
<td>3.2</td>
</tr>
<tr>
<td>3.3</td>
</tr>
<tr>
<td>3.3.1</td>
</tr>
<tr>
<td>3.3.2</td>
</tr>
<tr>
<td>3.4</td>
</tr>
<tr>
<td>3.5</td>
</tr>
<tr>
<td>3.6</td>
</tr>
<tr>
<td><strong>IV</strong></td>
</tr>
<tr>
<td>4.1</td>
</tr>
<tr>
<td>4.2</td>
</tr>
<tr>
<td>4.2.1</td>
</tr>
<tr>
<td>4.2.2</td>
</tr>
<tr>
<td>4.2.3</td>
</tr>
</tbody>
</table>
### CHAPTER IV (continued)

4.3 Drain Current and Charge Components in a Model Merging Weak, Moderate and Strong Inversion .......................... 51

4.3.1 Drain Current ........................................... 51
4.3.2 Charge Components .............................. 57
4.3.3 Limits for the Strong, Weak, and Moderately Inverted Portions of the Channel .................. 59

4.4 Results and Evaluation of the Model .......... 62
4.5 Conclusions ............................................. 74

### V FUNCTIONAL DEPENDENCIES FOR THE ELEMENTS IN THE FOUR-TERMINAL EQUIVALENT-CIRCUIT .......... 76

5.1 Introduction ........................................... 76
5.2 Source-Drain Current Source ................. 77
5.3 Capacitances ............................................ 78

5.3.1 Expression for the Capacitances .......... 79
5.3.2 Physical Interpretation of the Results for the Capacitances ................. 84
5.3.3 An Engineering Approximation for the Functional Dependencies of the Intrinsic Substrate Capacitances $C_{SB}$ and $C_{DB}$ ................. 92
5.3.4 Engineering Importance of the Intrinsic Substrate Capacitances $C_{SB}$ and $C_{DB}$ .......... 97

5.4 Transcapacitors ........................................ 98

5.4.1 Expressions for the Transcapacitors ........ 98
5.4.2 Engineering Importance of the Transcapacitance Elements .......... 100
5.4.3 Transcapacitances in a Three-Terminal Equivalent-Circuit .......... 106

5.5 Conclusions ............................................. 109

### VI SCOPE AND FUTURE WORK ......................... 111

### APPENDIX

A PROPERTIES OF QUASI-FERMI POTENTIALS .......... 115

B APPROXIMATED EXPRESSION FOR THE DIFFUSION/DRIFT RATIO IN THE MOSFET .......... 119

C COMPUTER SUBPROGRAM TO CALCULATE THE VALUE OF THE ELEMENTS IN THE EQUIVALENT-CIRCUIT .......... 123
<table>
<thead>
<tr>
<th>LIST OF REFERENCES</th>
<th>128</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOGRAPHICAL SKETCH</td>
<td>132</td>
</tr>
</tbody>
</table>
Abstract of Dissertation Presented to the Graduate Council of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

EQUIVALENT-CIRCUIT MODELING OF THE LARGE-SIGNAL TRANSIENT RESPONSE OF FOUR-TERMINAL MOS FIELD EFFECT TRANSISTORS

By
José Ignacio Arreola

March 1978

Chairman: Fredrik A. Lindholm
Major Department: Electrical Engineering

An approach is proposed that yields equivalent-circuit models for the large-signal transient response for all electronic devices described by charge-control. The approach is applied to derive an improved equivalent-circuit model for the four-terminal MOSFET. It is suggested that the model proposed gives a better description of the physics internal to the device than was previously available.

A static characterization of current and charges in the MOSFET is also proposed that unifies the descriptions of the weak, moderate and strong inversion modes of operation. Predictions of this characterization agree better with experimental results than previous work of similar complexity. The static characterization of current and charges is used to derive functional dependencies for the equivalent-circuit components in terms of applied voltages and physical make-up of the MOSFET.
CHAPTER I
INTRODUCTION

Computer simulations of MOSFET digital circuits can disagree severely with measured performance. A particular case of such a disagreement, which results in suboptimal circuit design, is the poor simulation of transient currents flowing in a substrate terminal of MOS field effect transistors [1]. The sources of such disagreements are either in the computer programs in use or in the inadequacies of existing large-signal equivalent-circuit models for the four-terminal MOSFET.

The purpose of this dissertation is to derive an improved equivalent-circuit model for the four-terminal MOSFET. Improvements are made in the following aspects of the equivalent-circuit model:

(a) the representation of capacitive effects in a four-terminal device;
(b) the characterization of the dc steady-state currents and charges;
(c) the inclusion, in principle, of two- and three-dimensional effects present, for example, in short-channel MOSFETs.
As will be seen, all of these improvements are inter-related and result from basing the derivation of the equivalent-circuit model on the internal physics that determines the operation of the MOSFET.

We begin in Chapter III by proposing an approach that yields equivalent-circuit models for the large-signal transient response of all electronic devices described by charge control [2-4]. The relation of this approach to the indefinite admittance matrix of circuit theory offers advantages in the modeling of devices having more than three terminals.

Chapter III starts by discussing the problems arising from the four-terminal nature of the MOSFET. Such problems were apparently not previously recognized. For the intrinsic part of the device (see Fig. 3.1), we apply the systematic approach developed in Chapter II. This approach, whose power is emphasized because of the four terminals of the MOSFET, yields a general description of the device that offers improvements (a) and (c) listed earlier.

To define fully the equivalent-circuit model of Chapter III, one needs a suitable description of the dc steady-state behavior. Extensive work has been done in the past to characterize operation in the dc steady-state; however, none of this work is completely suitable for the purposes of equivalent-circuit modeling. In Chapter IV, a new model for the dc steady-state behavior is derived that unifies the description of the full range of operation of the device -
from weak to strong inversion and from cut-off to saturation. The model avoids discontinuities in the characteristics present in all previous characterizations of similar complexity, and shows good agreement with experimental results. The new model also improves the characterization of the charges in the device.

In Chapter V we derive, using the results of Chapter IV, the functional dependence of each circuit element in the equivalent-circuit developed in Chapter III. In Chapter V we also assess the engineering importance of the improvements introduced in the equivalent-circuit model for the MOSFET and propose possible simplifications of the model.

Chapter VI treats possibilities for future research.
CHAPTER II

A NONLINEAR INDEFINITE ADMITTANCE MATRIX
FOR MODELING ELECTRONIC DEVICES

2.1 Introduction

This chapter describes a new approach for developing equivalent-circuit models of electronic devices. The models developed by this approach represent the large-signal (hence nonlinear) response to transient excitation. The approach applies to all devices whose operation is described by the principles of charge control [2-4], including, therefore, field effect transistors of various kinds, bipolar transistors, and certain electron tubes.

The models yielded by the approach are compact, composed of few circuit elements. As a result of their compactness, the models are meant to be useful in the computer-aided analysis of electronic circuits. This intended use contrasts with that intended for equivalent-circuit models [5] containing many circuit elements, which pertain chiefly to detailed studies of the physics underlying electronic-device behavior.

The approach to be described applies independently of the number of device terminals. Indeed, the greater that number, the more the power of the approach is disclosed.
The approach applies also independently of multidimensional spatial dependence that may be present in the boundary-value problem describing the device. This generality is needed, for example, in modeling the MOS field effect transistor (MOSFET), because the substrate terminal constitutes a fourth terminal through which sizable transient currents flow in some circuit applications, and because short-channel devices give rise to multidimensional effects.

Models of four-terminal devices [6,7] and models that include multidimensional effects [8] have been proposed earlier. But this previous work has not focused on laying down systematic procedures for developing models, which is the aim of this chapter.

Systematic procedures exist for modeling the linear response of multiterminal circuits subjected to small-signal excitation. These procedures are linked to the indefinite admittance matrix (IAM), which we first shall review and then exploit to model the nonlinear response of multiterminal electronic devices to large-signal excitation.

2.2 Indefinite Admittance Matrix

Consider a lumped electrical network which has n terminals. Let an additional external node be the common reference. From the standpoint of its behavior at the terminals, the network, if linear, may be described by a set of equations as follows:
\[ I = yV \] \hspace{0.5cm} (2.1)

The required linearity is assured for any network operating under small-signal conditions. The matrix elements of \( y \) are
\[ y_{jk} = \left. \frac{I}{V} \right|_{V_i = 0, \ i \neq k} \] \hspace{0.5cm} (2.2)

where \( I \) and \( V \) correspond to the current and voltage at the terminals.

The matrix \( y \) defined in (2.1) and (2.2) is called the indefinite admittance matrix \([9,10]\), and its elements satisfy the following property imposed by Kirchhoff's laws
\[ \sum_j y_{jk} = \sum_k y_{jk} = 0 \] \hspace{0.5cm} (2.3)

that is, the elements in any row or any column sum to zero.

As will be seen, our development of large-signal models for electronic devices will make use of two special cases of the IAM. In the first case, the matrix \( y \) is symmetric and has one of the following forms:
\[ y = a \quad y = b \frac{d}{dt} \quad y = c/dt \] \hspace{0.5cm} (2.4)

Here \( a, b, \) and \( c \) are real symmetric matrices, and each matrix element corresponds to a single lumped resistor or capacitor or inductor connected between each pair of the \( n \) terminals. In the second case, the matrix \( y \) is nonsymmetric, but is the sum of two indefinite admittance matrices: a symmetric
matrix, like (2.4), and a residual nonsymmetric matrix, each element of which corresponds to a controlled current source placed between each pair of terminals. In this second case, then, the circuit representation of the IAM results from connecting the network corresponding to the symmetric matrix in parallel with that corresponding to the nonsymmetric matrix. In general, summing of indefinite admittance matrices corresponds to connecting their circuit representations in parallel.

2.3 Extension for Nonlinear Electronic Devices

Consider an electronic device having n terminals. The modeling begins by specifying the physical mechanisms relevant to the operation of the device. For many devices, only three such mechanisms, at most, are relevant: the transport of charged carriers between terminals; the net recombination of charged carriers within the device; the accumulation of these carriers within the device. Thus the current $i_J$ flowing at any terminal $J$ is the sum of three components: a transport current $(i_J)_T$, a recombination current $(i_J)_R$, and a charging current $(i_J)_C$. That is

$$i_J = (i_J)_T + (i_J)_R + (i_J)_C \quad (2.5)$$

We now characterize these components.

The transport mechanism consists of the injection of a charged carrier in one terminal, followed by its transport
across the device until it reaches any of the other terminals, where it recombines at a surface with a carrier of opposite charge. The recombination mechanism differs from the transport mechanism only in that the carriers recombine within the bulk of the device instead of at the terminals. Therefore, both mechanisms can be characterized by the same form

\[ (i_J)_T,R = \sum_{K \neq J} (i_{JK})_T,R. \]  

(2.6)

Here \( i_{JK} \) represents the current due to the charged carriers injected from terminal \( J \), which recombine, at a surface or in the bulk, with opposite-charged carriers injected from terminal \( K \). From this characterization, it follows that \( (i_{JK})_T,R \) satisfies the following properties:

\[ i_{JK} = -i_{KJ} \quad i_{JJ} = 0. \]  

(2.7)

These properties allow transport and net recombination to be represented by controlled current sources connected between pairs of terminals. The value of the current source between terminals \( J \) and \( K \) is \( i_{JK} \).

The last mechanism to be considered is the accumulation of mobile carriers within the device, which requires the charging current

\[ (i_J)_C = \frac{dq_J}{dt}. \]  

(2.8)

As Fig. 2.1 illustrates, \( dq_J \) is the part of the total charge accumulated within the device in time \( dt \) that is supplied
Fig. 2.1 The charging current $(i_J)$ at terminal $J$ produces the accumulated charge $dq_J$. 

\[ i_J = (i_J)^{TR}_R + (i_J)^C \]
from terminal J. The charge accumulation expressed in (2.8) is a mechanism basic to any electronic device that operates by charge control [2-4].

Now, using (2.6) and (2.8), we may rewrite (2.5) as

$$i_J = \sum_{K \neq J} (i_{JK})_{T,R} + \frac{dq}{dt}.$$  

(2.9)

Although (2.9) is valid, it does not correspond to a convenient network. To get a convenient network representation, we apply one additional constraint which costs small loss in generality in that it holds for all charge-control devices [2-4]. We apply the constraint that the overall device under study is charge neutral. Or, more exactly and less demanding, we assume the device accumulates no net overall charge as time passes. This constraint of overall charge neutrality requires a communication of the flux lines among the terminals to occur that maintains charge neutrality by coulomb forces and by drift and diffusion currents. The requisite overall neutrality may result either from neutrality occurring at each macroscopic point, as in a transistor base, or from a balancing of charges that are separated, as on the gate and in the channel of a MOSFET.

As a result of overall neutrality, the current at any terminal J becomes the sum of the currents flowing out of all of the other terminals

$$i_J = - \sum_{K \neq J} i_K.$$  

(2.10)
This global counterpart of the Kirchhoff current-node law implies for the charging currents of (2.8) that

\[
(i_J)_C = - \sum_{K \neq J} (i_K)_C
\]

(2.11)

which means that a charging current entering one terminal flows, in its entirety, out of all of the other terminals. Hence, as is true also for the transport and recombination mechanisms, charge accumulation can be represented by a controlled current source connected between each pair of terminals.

For a model to be useful in circuit analysis, the elements of the model must all be specified as functions of the terminal currents and voltages. To do this, we now make use of the principles of charge control [2-4] and of the closely allied quasi-static approximation [6,7,11]. For the transport and recombination mechanisms, charge control gives directly

\[
(i_{JK})_{T,R} = \frac{q_{JK}}{t_{JK}}.
\]

(2.12)

Here \(q_{JK}\) is the charge of the carriers that contribute to the current flowing between terminals \(J\) and \(K\). The recombination time \(t_{JK}\) is the time constant associated with that current: a transit time if the mechanism being described is transport, a lifetime if it is recombination. Then, to produce the desired functional dependence, a quasi-static approximation [6,7,11] is used that specifies each
(i_{JK})_{T,R} as a function of the instantaneous voltages at the device terminals.

This characterization of \((i_{JK})_{T,R}\) combined with the properties expressed in (2.7), can be manipulated to describe transport and recombination by an IAM, like \(a\) in (2.4). Because \(i_{JK} = -i_{KJ}\), the matrix is symmetric.

There are two network representations of transport and recombination described by this matrix. As noted before, just below (2.7), one of these consists of controlled current sources connected between pairs of terminals. Another network representation consists entirely of non-linear resistors, \(R_{JK} = (V_J - V_K)/i_{JK}\).

Similar procedures are applied to model charge accumulation. To the charging current defined in (2.8) a quasi-static approximation is applied [6,7,11], specifying the functional dependence of \(q_J\) on the terminal voltages and enabling thereby the employment of the chain rule of differentiation. The resulting characterization of \((i_J)_C\) describes charge accumulation by a matrix that has the form of \(b\) in (2.4), a matrix whose elements are

\[
b_{JK} = \frac{\partial q_J}{\partial V_K} \bigg|_{dV_I=0}, \quad I \neq K
\]  

(2.13)

Matrix \(b\) also satisfies the key properties of the indefinite admittance matrix that are given in (2.3). For a general \(n\)-terminal electronic device, this matrix describing charge accumulation is nonsymmetric, and is therefore the sum of a
symmetric and a residual nonsymmetric part. The symmetric part corresponds to an all-capacitor network; the network representation of the residual nonsymmetric matrix consists of controlled current sources.

2.4 Conclusions

From the properties of the IAM it follows that the three-branch circuit of Fig. 2.2 serves as a building block for model generation. Connecting a circuit of this form between each terminal pair yields the general network representation for an n-terminal electronic device. For any particular device of interest, certain of the circuit elements may vanish. In a MOSFET, for instance, no transport or recombination currents flow to the gate, and the corresponding circuit elements will be absent.

Any equivalent-circuit model generated by this approach can be regarded in two ways: either as a product of the building block of Fig. 2.2 or as a circuit described by a matrix which obeys the key properties of the IAM. Description by the IAM treats all terminals equally in that none is singled out as the reference node; the advantages of this will show up plainly in the modeling of a four-terminal device, such as the MOSFET.

From Fig. 2.2 notice that the mobile charge accumulation within a general n-terminal electronic device is not represented by the flow of displacement currents in an all-capacitor model. The residual nonsymmetric matrix, and
Fig. 2.2 General equivalent-circuit between each pair of terminals of an n-terminal electronic device.
the corresponding transcapacitance current source of Fig. 2.2, provides the needed correction. This correction has practical engineering consequences in certain MOSFET circuits although a discussion of that is postponed for a later chapter.

To use the approach set forth here in modeling any particular device requires that the static dependence on the terminal voltages be specified for the currents and charges defined in (2.12) and (2.13). This requires that a physical model for the device be chosen to describe the dc steady state. For the MOSFET this has been done, and the corresponding equivalent-circuit model is derived in the following chapters.
CHAPTER III
EQUIVALENT-CIRCUIT MODEL FOR
THE FOUR-TERMINAL MOSFET

The main contribution of this chapter is the derivation of an equivalent-circuit model for the four-terminal MOSFET by use of the method described in Chapter II. The resulting model is intended to represent with good accuracy the large-signal transient currents flowing through each of the four terminals of the device, including the substrate terminal.

3.1 Examples of Engineering Needs for a Model for the Large-Signal Transient Response

In many digital integrated-circuit applications of the MOSFET, the substrate terminal of each device is connected to a power supply. This connection serves at least two purposes: it provides a means to control the threshold voltage of the device, and it enables a good lay-out of the circuit [12,13]. In a large-scale integrated circuit, the large transient current flowing through a power supply can result in poor voltage regulation and poor circuit performance unless both the circuit and the power supply are properly designed. An optimum design of a circuit will provide the maximum density of components on the chip consistent with
the requirement that the voltage regulation of each power supply remains acceptable.

To design circuits using computer aids therefore requires that one has available a set of equivalent-circuit models for the MOSFET that adequately represent the transient currents flowing through the terminals in response to large-signal excitation of the devices. According to engineers involved in such designs, such models are not now available [1]. This absence of accurate models forces the engineer to suboptimal designs, by which we mean less densely packed circuits than those that could be designed if accurate enough device models were available.

3.1.1 Reasons for the Poor Modeling of the Transient Substrate Current by Existing MOSFET Models

The substrate current during transients arises from capacitive effects in two regions of the device (Fig. 3.1): the depletion region around the source and drain islands (extrinsic substrate capacitances); and the depletion region underneath the inversion channel (intrinsic substrate capacitance). In general, however, the substrate current is modeled as arising only from the p-n junction (extrinsic) capacitances around the source and drain islands. These capacitances have the form [13]:

\[
C_j = \frac{C_{jo}}{\left(1 - \frac{V}{\phi_B}\right)^n}
\]  
(3.1)
Fig. 3.1 An n-channel enhancement MOSFET divided into intrinsic and extrinsic parts.
where \( V \) is the applied junction voltage, \( \phi_B \) is the built-in potential and \( n \) is an exponential factor. The maximum value of these capacitances, given by \( C_{jo'} \), is estimated for typical doping concentrations to be in the order of \( 10^{-8} \) F/cm\(^2\) [13]. As we shall see in Chapter V this is also the order of magnitude of the intrinsic substrate capacitances. Because the area of the channel and the area of the source and drain islands are in many circuits comparable, the inclusion of the intrinsic capacitive effects to model the substrate current is essential. Moreover, in new fabrication technologies such as silicon on sapphire [14] considerable reduction of the substrate extrinsic capacitances can be achieved. These reductions can also be achieved by employing special circuit techniques in the conventional technology [15]. In both these cases, the intrinsic effects are dominant and must be included in the modeling.

3.2 Problems Involved in Modeling of Four-Terminal Devices

The modeling of the intrinsic effects of the four-terminal MOSFET presents special problems not previously considered. To lead into these problems, consider first a two-terminal device. As is shown in Fig. 3.2(a), we apply a small voltage dV. The figure illustrates that there is only a single path of communication between the terminals. That is, there is only one way the flux lines can link be-
Fig. 3.2 Illustration of the paths of communication between terminals in a two- and four-terminal device.
tween the terminals and thus there is no uniqueness in the charge that flows at the terminals. The charge that flows at each terminal is dQ. A nonuniqueness does occur, however, in devices with more terminals. Consider now a four-terminal device. From Fig. 3.2(b) one sees that there are six paths of communication of the flux lines among the terminals in a general four-terminal device. Thus, suppose one applies a small voltage between any two terminals while appropriately terminating the other terminals so that charge can flow through them. Then one must account properly for the apportionment of the charges amongst the terminals. Of the total charge dQ that flows, what will be the charges dQ₁, dQ₂, dQ₃ and dQ₄ flowing at each of the four terminals?

There is a second related problem. One way of seeing this problem is to suppose that within the box of Fig. 3.1(b), for the time being, is an all-capacitor network. Then apply a small voltage between terminals 1 and 3, having shorted the other terminals to an arbitrary reference. In response, a certain amount of charge flows at terminal 4. Now inter-change the roles of terminals 3 and 4. That is, apply the small voltage at terminal 4 and measure the amount of charge flowing past terminal 3. The result of this experiment is that one finds exactly the same amount of charge as before. That is a property of a reciprocal network, of which an all-capacitor configuration is an example [16].

Now if one does the same experiment with a MOSFET one finds that this reciprocity does not apply, as we shall prove
in Section 5.4. The reason is that terminal 3 represents
the gate and terminal 4 represents the substrate; and the
gate and substrate are highly different physical struc-
tures. This asymmetry in physical structure introduces
a nonreciprocity in the network properties not present in
an all-capacitor network. To account for this asymmetry,
therefore, one should expect that the network representa-
tion for a MOSFET must contain elements describing mobile
charge accumulations in addition to capacitors.

To manage these problems one requires a systematic
approach. In Chapter II we have developed a methodology
that permits one to obtain a lumped network representation
of multiterminal electronic devices obeying the principles
of charge control whose large-signal transient behavior
depends on three physical mechanisms: mobile charge trans-
port, net recombination within the device and mobile charge
accumulation. The result is the equivalent-circuit of
Fig. 2.2, which applies between any pair of terminals and
is the basic building block from which an equivalent-circuit
can be constructed for the overall multiterminal device.
The currents representing transport and net recombination
flow in the current source $i_{JK}$. The charging current re-
presenting mobile charge accumulation flow through the
capacitor $C_{JK} = \frac{\partial q_j}{\partial v_J}$ and through the controlled current
source characterized by $d_{JK} = \frac{\partial q_j}{\partial v_K} - \frac{\partial q_k}{\partial v_J}$.

To apply this methodology to the MOSFET, one needs
only to describe the components of charge accumulation $d q_j$.
in each region and the transport and recombination flow in terms of the physics underlying the device behavior. We will now apply this methodology to the MOSFET.

3.3 Equivalent-Circuit for the Intrinsic MOSFET

For concreteness, consider the enhancement-mode n-channel MOSFET illustrated in Fig. 3.1. A central idea in the equivalent-circuit modeling is to resolve the electronic device under study into two parts [11]: an intrinsic part where the basic mechanisms responsible for the operation of the device occur, and an extrinsic part which depends on the details of the device structure. For the particular MOSFET under consideration this is done in Fig. 3.1.

The behavior of the intrinsic region in the MOSFET is described by charge control [2-4], and thus an equivalent circuit of its operation can be obtained by applying the methodology described in Chapter II.

3.3.1 Transport Current

At normal operating voltages and temperatures the leakage current in the insulated gate is negligible and the recombination/generation rate in the channel and in the substrate can be neglected. Charge transport occurs, therefore, only along the highly conductive inversion channel induced at the semiconductor surface. This transport mechanism is represented in the equivalent circuit as a controlled current source $i_{SD}$ connected between source and drain. Its explicit functional dependence in terms of the
physical make-up and the terminal voltages is obtained by using a quasi-static approximation to extrapolate the steady-state functional dependence of the drain current $I_D$. This will be considered in Chapter V.

3.3.2 Charging Currents

If we neglect recombination and generation, the current flowing in the substrate terminal $i_B$ is solely a charging current, that is, current that changes the number of holes and electrons stored in the intrinsic device. Thus, if during time $dt$ a change $dq_B$ occurs in the hole charge stored in the substrate, then

$$i_B = \frac{dq_B}{dt}.$$  \hspace{1cm} (3.2)

Similarly, neglecting any leakage current in the insulator, the current flowing in the gate $i_G$ is only a charging current. If this current changes the charge of the metal gate by $dq_G$ in time $dt$, then

$$i_G = \frac{dq_G}{dt}.$$  \hspace{1cm} (3.3)

The current flowing at the source terminal consists of two components. The first component changes the electron charge stored in the channel by an amount $dq_S$ in time $dt$. The second component arises from electrons that, flowing in from the source, pass through the channel and then out of the drain terminal. Thus,
\[
i_S = \frac{dq_s}{dt} + i_{SD}.
\]

(3.4)

Similarly, on the drain side,

\[
i_D = \frac{dq_D}{dt} - i_{SD}.
\]

(3.5)

The total change of charge in the inversion channel \(dq_N\) is then

\[
i_S + i_D = \frac{dq_S}{dt} + \frac{dq_D}{dt} = \frac{dq_N}{dt}
\]

(3.6)

If we apply a quasi-static approximation \([6,11]\) and then use the chain rule of differentiation, equations (3.2) through (3.5) can be expressed in the following matrix form:

\[
\begin{bmatrix}
i_S \\
i_D \\
i_G \\
i_B
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial q_S}{\partial v_S} & \frac{\partial q_S}{\partial v_D} & \frac{\partial q_S}{\partial v_G} & \frac{\partial q_S}{\partial v_B} \\
\frac{\partial q_D}{\partial v_S} & \frac{\partial q_D}{\partial v_D} & \frac{\partial q_D}{\partial v_G} & \frac{\partial q_D}{\partial v_B} \\
\frac{\partial q_G}{\partial v_S} & \frac{\partial q_G}{\partial v_D} & \frac{\partial q_G}{\partial v_G} & \frac{\partial q_G}{\partial v_B} \\
\frac{\partial q_B}{\partial v_S} & \frac{\partial q_B}{\partial v_D} & \frac{\partial q_B}{\partial v_G} & \frac{\partial q_B}{\partial v_B}
\end{bmatrix}
\begin{bmatrix}
v_S \\
v_D \\
v_G \\
v_B
\end{bmatrix}
\begin{bmatrix}
\dot{v}_S \\
\dot{v}_D \\
\dot{v}_G \\
\dot{v}_B
\end{bmatrix}
\begin{bmatrix}
i_{SD} \\
-i_{SD} \\
0 \\
0
\end{bmatrix}
\]

(3.7)

Here the dot notation designates time derivatives.

By applying the constraint that the overall intrinsic device is charge neutral, one can prove as is done in Chapter II \([17,18]\) that the first matrix in (3.7) satisfies the properties of the indefinite admittance matrix of network theory \([9]\). That is, the sum of all the elements in any row or column is equal to zero.
The matrix description of (3.7) together with the building block of Fig. 2.2 yields, therefore, the general large-signal equivalent-circuit for the intrinsic four-terminal MOSFET. This network representation is shown in Fig. 3.3 and its elements are defined in Table 1.

Elements in addition to capacitors that represent charging currents appear in the circuit of Fig. 3.3. These transcapacitors would be zero only if the matrix in (3.7) were symmetric. That is, if \( \frac{\partial q_J}{\partial v_K} = \frac{\partial q_K}{\partial v_J} \) for all \( J \) and \( K \). The physical structure of the MOSFET, however, is nonsymmetric and hence one should expect that the elements \( d_{JK} \) are in general nonzero. This is the case, indeed, as it will be shown in Chapter IV where we calculate the functional dependencies of these elements in terms of the applied voltages and the device make-up.

The transcapacitive elements in the network representation can be also seen as related to error terms yielded by an ideal all-capacitor model. In this sense, we will study and assess their importance in Chapter V.

In the circuit of Fig. 3.3, the capacitive effects between source and drain are represented by a capacitor \( C_{SD} \) and a controlled current source characterized by \( d_{SD} \). In the theory of operation of the MOSFET based on the gradual case [19], it has been shown [20] that there are no capacitive effects between source and drain. In this work, we will consider this to be the case and therefore we will assume \( \frac{\partial q_S}{\partial v_D} = \frac{\partial q_D}{\partial v_S} = 0 \), implying
Fig. 3.3 General equivalent-circuit of the intrinsic MOSFET.
Table 1  Definitions for the elements of the general equivalent-circuit for the MOSFET.

<table>
<thead>
<tr>
<th>CAPACITANCES</th>
<th>TRANSCAPACITANCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{SG} = -\frac{\partial q_G}{\partial v_S}$</td>
<td>$d_{SG} = \frac{\partial q_S}{\partial v_G} - \frac{\partial q_G}{\partial v_S}$</td>
</tr>
<tr>
<td>$c_{SB} = -\frac{\partial q_B}{\partial v_S}$</td>
<td>$d_{SB} = \frac{\partial q_S}{\partial v_B} - \frac{\partial q_B}{\partial v_S}$</td>
</tr>
<tr>
<td>$c_{SD} = -\frac{\partial q_D}{\partial v_S}$</td>
<td>$d_{SD} = \frac{\partial q_S}{\partial v_D} - \frac{\partial q_D}{\partial v_S}$</td>
</tr>
<tr>
<td>$c_{DG} = -\frac{\partial q_G}{\partial v_D}$</td>
<td>$d_{DG} = \frac{\partial q_D}{\partial v_G} - \frac{\partial q_G}{\partial v_D}$</td>
</tr>
<tr>
<td>$c_{DB} = -\frac{\partial q_B}{\partial v_D}$</td>
<td>$d_{DB} = \frac{\partial q_D}{\partial v_B} - \frac{\partial q_B}{\partial v_D}$</td>
</tr>
<tr>
<td>$c_{GB} = -\frac{\partial q_B}{\partial v_G}$</td>
<td>$d_{GB} = \frac{\partial q_G}{\partial v_B} - \frac{\partial q_B}{\partial v_G}$</td>
</tr>
</tbody>
</table>
\[ C_{SD} = 0 , \quad d_{SD} = 0 . \] (3.8)

In more detailed characterizations of the device - for example, the ones including channel length modulation [21] and two-dimensional effects in short channel devices [8] - the drain voltage directly influences the charging of the channel and capacitive effects between source and drain as modeled by Fig. 3.3 may need to be included.

3.4 Special Considerations

For the equivalent circuit model in Fig. 3.3 to be useful in circuit analysis we require that all the elements, current sources and capacitors, must be specified as functions of the terminal current and voltages. In doing this, as indicated in Chapter II, we will use the quasi-static approximation [6,11], which is based on the steady state operation of the MOSFET. A particular detailed model for steady-state operation is considered in Chapter IV and the functional dependencies for the elements of Table 1 will be derived in Chapter V from this model. However, before approaching these problems, we must give special consideration to two charge components that are not described in the conventional steady-state characterization of the device: the contributions from the source and the drain, \( dq_S \) and \( dq_D \), to the total charging of the channel. To gain physical insight as to how \( dq_S \) and \( dq_D \) contribute to the charging of the channel, consider the following.
If we apply a change in the gate voltage, a change of the charge in the channel dq_N will occur. The electrons necessary to supply this additional charge are injected into the channel by charging currents flowing in from the source and drain, that is

\[ i_S + i_D = \frac{dq_S}{dt} + \frac{dq_D}{dt} = \frac{dq_N}{dt} \] (3.9)

The contributions of dq_S and dq_D to dq_N are, in general, unequal and depend, as we shall see, on the operating conditions of the device.

Figure 3.4 shows a simplified energy band diagram at the surface of an N-channel MOSFET under various operating conditions determined by the magnitude of V_D. Consider first the case when V_D = 0 and AV_G is applied. Because the barrier height that the electrons have to overcome in both sides of the channel is equal (Fig. 3.1(b)), we expect that charging currents flowing into the source and drain ends will be equal,

\[ \left[ \frac{dq_S}{dt} \right]_{V_D=0} = \left[ \frac{dq_D}{dt} \right]_{V_D=0} \] (3.10)

Now apply a small V_D > 0 and change the voltage by AV_G. As in the previous case, electrons are injected from both sides of the channel. The electric field produced by the application of V_D, however, will present an additional barrier height for the electrons injected from the drain side (Fig. 3.4(c)). Thus we expect the charging current in the source to be larger
Fig. 3.4 Energy band diagram at the surface of a MOSFET under the effect of applied drain and gate voltage.
than the charging current in the drain. That is,

\[
\left(\frac{dq_S}{dt}\right)_{V_D>0} > \left(\frac{dq_D}{dt}\right)_{V_D>0}.
\]

(3.11)

For larger values of \(V_D\) the device will be eventually driven into saturation. The high electric field produced near the drain will impede charging of the channel from that end (Fig. 3.4(d)). Hence, the additional electrons required when \(\Delta V_G\) is applied will be supplied mainly from the source end. That is,

\[
\left(\frac{dq_D}{dt}\right)_{\text{SATURATION}} \approx 0.
\]

(3.12)

A similar argument can be employed to explain the contributions of \(dq_S\) and \(dq_D\) to the charging of the channel due to changes in the substrate voltage.

From the above discussion we can define an apportionment function \(\lambda\) such that the source and drain charging currents can be expressed as

\[
\left(\frac{dq_S}{dt}\right)_{V_S', V_D} = \lambda \left(\frac{dq_N}{dt}\right)_{V_S', V_D}.
\]

(3.13)

and

\[
\left(\frac{dq_D}{dt}\right)_{V_S', V_D} = (1 - \lambda) \left(\frac{dq_N}{dt}\right)_{V_S', V_D}.
\]

(3.14)

The apportioning function \(\lambda\) takes values from 1/2 to 1 between the conditions of \(V_{DS} = 0\) and saturation.
A convenient expression for $\lambda$ results from combining its definition in (3.13) with the indefinite admittance matrix that characterize the charging currents in (3.7). Using chain rule differentiation in (3.13),

$$\frac{\partial q}{\partial V_G} \dot{V}_G + \frac{\partial q}{\partial V_B} \dot{V}_B = \lambda \left( \frac{\partial q}{\partial V_G} \dot{V}_G + \frac{\partial q}{\partial V_B} \dot{V}_B \right)$$

This equation must remain valid for any value of $\dot{V}_G$ and $\dot{V}_B$.

Thus

$$\lambda = \frac{\partial q}{\partial V_G} \frac{\partial q}{\partial V_B}$$

By using the properties of the indefinite admittance matrix, the numerator and denominator of (3.16) can be rewritten as:

$$\frac{\partial q}{\partial V_G} + \frac{\partial q}{\partial V_B} = -\left( \frac{\partial q}{\partial V_G} + \frac{\partial q}{\partial V_D} \right)$$

$$= \frac{\partial q}{\partial V_S} + \frac{\partial q}{\partial V_B} + \left( \frac{\partial q}{\partial V_S} - \frac{\partial q}{\partial V_D} \right)$$

and

$$\frac{\partial q}{\partial V_G} + \frac{\partial q}{\partial V_B} = -\left( \frac{\partial q}{\partial V_G} + \frac{\partial q}{\partial V_D} \right)$$

$$= \frac{\partial q}{\partial V_S} + \frac{\partial q}{\partial V_B} + \frac{\partial q}{\partial V_D}$$

Substituting (3.17) and (3.18) into (3.16) and, using the definitions in Table 1, we obtain,
\[ \lambda = \frac{1}{1 + \frac{C_{DG}}{C_{DB}} + \frac{C_{SG}}{C_{SB}}} \quad (3.19) \]

Here, we have used the assumption that no direct capacitive effects exist between source and drain \((\partial q_S/\partial v_D = \partial q_D/\partial v_S = 0)\).

Equation (3.19) can now be used to obtain the functional dependencies of equivalent-circuit elements involving \(dq_S\) and \(dq_D\) directly from an extrapolation of the steady-state behavior of the device. From Table 1, these elements are,

\[ d_{SG} = \frac{\partial q_S}{\partial v_G} - \frac{\partial q_G}{\partial v_S} = \lambda \frac{\partial q_G}{\partial v_G} + C_{SG} \quad (3.20) \]

\[ d_{DG} = \frac{\partial q_D}{\partial v_G} - \frac{\partial q_G}{\partial v_D} = \lambda \frac{\partial q_G}{\partial v_G} + C_{DG} \quad (3.21) \]

\[ d_{SB} = -d_{SG} \quad (3.22) \]

\[ d_{DB} = -d_{SB} \quad (3.23) \]

\[ d_{GB} = d_{SG} + d_{DG} \quad (3.24) \]

Equations (3.22)-(3.24) have been simplified by direct application of the properties of the indefinite admittance matrix.

3.5 Modeling of the Extrinsic Components

The extrinsic components depend on details of the fabrication of a specific type of MOSFET. In many cases,
the extrinsic part can be modeled by inspection of the geometry of the device. Elements commonly found are: overlapping capacitances due to the overlap of the gate oxide over the source and drain islands; bonding capacitances resulting from metalization over areas where the oxide is relatively thick; P-N junction capacitances arising from source-substrate and drain-substrate diffusions; and resistance components due to finite resistivity at the source, drain and substrate. In general, these elements are distributed capacitances and resistors but can be transformed to lumped elements by applying a quasi-static approximation. Lindholm [11] gives the details of the general approach for modeling extrinsic effects in a four-terminal MOSFET. For particular devices, the details of the extrinsic modeling have been worked out in the literature [7,22].

3.6 Relation to Existing Models

A wide variety of equivalent-circuit models of different complexity and accuracy have been advanced for the MOSFET [7,11,20,22,23]. The general development of these models follows a partially heuristic and partially systematic approach that consists in interpreting in circuit form the different terms of the equations describing the device physics. The definitions of the elements in these circuit models depend on the particular approximations of the physical model involved.
In contrast, the equivalent-circuit of Fig. 3.3 and the definitions of its elements in Table 1, having been developed from a methodology based on fundamentals, are quite general. For example, the new network representation can take into account two and three-dimensional effects such as those in the short-channel MOSFET. To use the model one needs only compact analytical descriptions of these effects in physical models for the dc steady-state. Such descriptions, we anticipate, will appear in the future. Indeed, as new physical models for dc behavior appear, such as the one presented in the next chapter, the equivalent-circuit developed here is designed to make immediate use of them to yield new and better network representations of the large-signal transient response of the MOSFET.

Most of the past work in equivalent-circuit modeling of the intrinsic MOSFET neglects the effect of charging currents flowing into the substrate terminal. Among the models that consider these effects, the treatment of Cobbold [6] is the most detailed. His model, derived for small-signal applications, involves an equivalent-circuit in which the charging effects are represented by four capacitors (source-gate, source-substrate, drain-gate and drain-substrate) and a controlled current source (gate-substrate). As can be observed in the general equivalent-circuit between any two terminals shown in Fig. 2.2, the representation by a capacitor alone of charging currents
between two terminals requires certain specific conditions related to symmetry and apportionment of charge in the device to be satisfied. For example, if the terminals are the source and the gate a single-capacitor representation between these terminals would require \( \frac{\partial q_s}{\partial v_G} = \frac{\partial q_G}{\partial v_S} \). Because of the physical asymmetry of the MOSFET, these requirements are, in general, not satisfied. This problem was apparently not recognized by any of the previous workers in the field.
CHAPTER IV
STEADY-STATE MOSFET THEORY MERGING
WEAK, MODERATE AND STRONG INVERSION

4.1 Introduction

In Chapter III we have developed a circuit representation for the transient behavior of the intrinsic four-terminal enhancement-mode MOSFET. Each circuit element in this representation depends on the constants of physical make-up of the MOSFET and on the voltages at the terminals of the intrinsic device in a way that is determined by the static model chosen to represent the current and the inversion, substrate and gate charges. To complete the modeling, therefore, one must choose a static model for this current and these charges that is general enough to be suited to whatever circuit application is under consideration. None of the static models previously developed are suitable for this purpose, for reasons that will be soon discussed. Thus the purpose of this chapter is to develop a model that has the properties required.

One necessary property of the static model is that it represents the entire range of operation to be encountered in various circuit applications, including the cut-off, triode, and saturation, including operation in weak, moderate and strong inversion, and including four-terminal operation.
The model of Pao and Sah [24] comes nearest to this ideal. It covers in a continuous form the entire range of operation. However, its mathematical detail makes it inconvenient for computer-aided circuit analysis, and it does not include the substrate charge and the influence of the substrate terminal.

The Pao and Sah model has provided the basis for other modeling treatments. Swanson and Meindl [25], and Masuhara et al. [26] have presented simplified versions covering the entire range of operation. Their approach consists in a piecewise combination of models for the limits of weak and strong inversion. This approach introduces discontinuities in the slopes of the characteristics for moderate inversions, which are computationally undesirable. These models, furthermore, do not include charge components and the influence of the substrate terminal.

Following a different line of reasoning El-Mansey and Boothroyd [27] have derived an alternative to the Pao and Sah model. Their work includes charge components and four-terminal operation. However it also is mathematically more complicated than is desirable for computer-aided circuit design.

The goal of this chapter is to develop a model that includes:

(a) four-terminal operation;
(b) cut-off, triode and saturation regions;
(c) weak, moderate and strong inversion;
(d) current and total charges.
The model, furthermore, should avoid the discontinuities of a piecewise description while maintaining enough mathematical simplicity for computer-aided circuit analysis.

In Section 4.2 a review of the general fundamental of MOSFET operation are presented. A discussion, in Section 4.2.3, of the relation between the surface potential and the quasi-Fermi level for electrons sets the basis of our approach. In Section 4.3 expressions for the drain current and the total charge components are derived. To assess the validity of our approach, the predictions of our model for the drain current are compared against experimental data in Section 4.4. In the last section we include a discussion of the limitations of the model.

4.2 Fundamentals

4.2.1 Drain Current

In an n-channel MOSFET, illustrated by Fig. 3.1, the steady-state drain current density \( J_D(x,y) \) is essentially the electron current density in the inversion channel [24]:

\[
J_D(x,y) = J_n(x,y) = q \mu_n N \frac{dN}{dy} + q D_n \frac{dN}{dy} = -q \mu_n N \frac{dV}{dy}
\]

(4.1)

where \( V = V_N - V_P \) is defined as the difference between the quasi-Fermi potential for electrons \( V_N \) and the quasi-Fermi potential for holes \( V_P \). Because there is no significant hole current flowing in the device [28] \( V_P \) is nearly constant.
and coincides with the bulk Fermi potential, 
\( \phi_F = \frac{kT}{q} \ln \frac{N_{AA}}{n_i} \). The voltage \( V \) is referred to as the "channel voltage" [20], and at the boundaries of the channel, \( y=0 \) and \( y=L \), it has the values \( V(0) = V_S \) and \( V(L) = V_D \). These and other properties of \( V \) will be derived in Appendix A.

The total drain current is obtained by using the gradual channel approximation [19]:

\[
I_D = -Z \int_0^\infty J_D(x,y) \, dx = -Z \nabla_n \frac{Q_n}{\tau_y} \cdot \frac{dV}{dy} . \tag{4.2}
\]

Here \( Z \) is the channel width, \( \nabla_n \) is an effective mobility, and \( Q_n \) is the electron charge per unit area in the inversion channel defined by

\[
Q_n = -q \int_0^\infty N \, dx . \tag{4.3}
\]

The differential equation in (4.2) is solved by integrating along the channel

\[
I_D = - \left( \frac{Z \nabla_n C_o}{L} \right) \int_{V_S}^{V_D} \frac{Q_n}{C_o} \, dv , \tag{4.4}
\]

where \( L \) is the effective channel length and \( C_o \) is the oxide capacitance per unit area.

The effects due to mobility reduction and channel length modulation have been studied in detail by different authors [21,29,30]. They could be included in this work by appropriately modifying \( \nabla_n \) and \( L \).
4.2.2 Charge Components

For the purposes of equivalent-circuit modeling it is convenient to divide the charge distribution within the intrinsic device in three components: charge associated with the gate, charge in the bulk and charge in the inversion channel.

In the charge associated with the gate we include: the actual charge in the metallic gate ($Q_{Vox}$), the fixed charge in the oxide $Q_{ox}$, and the charge due to surface states at the oxide-semiconductor interface $Q_{ss}$. Inspection of the energy band diagram of Fig. 4.1 shows that this effective gate charge $Q_g$ per unit area can be expressed as

$$\frac{Q_g}{C_o} = V_G - \phi_{MS} - \psi_S + \frac{Q_{ox}}{C_o} + \frac{Q_{ss}}{C_o} \quad (4.5)$$

where $\phi_{MS} = \phi_m - \chi_S \cdot q(E_C - E_I) - \phi_F$ is by definition the metal-semiconductor work function, $\psi_S$ is the surface potential and $V_G$ is the applied gate voltage. In this work we will assume that the charge in the surface states $Q_{ss}$ is independent of voltage. It has been demonstrated, however, that when the device is operating under low voltage conditions [31] the voltage dependence of $Q_{ss}$ becomes important in determining the relation between surface potential and external applied voltages. A typical characterization of $Q_{ss}$ is given by [32]:

$$Q_{ss} = -q\eta_{ss} (\psi_S - V) \quad (4.6)$$
Fig. 4.1 Energy band diagram under nonequilibrium conditions. All voltages are referred to the substrate. Note that $-qV_n = E_{Fn}$, $-qV_p = E_{Fp}$, and $-qV_I = E_I$. 
where $N_{ss}^s$, representing the surface state density per unit area, is used as a parameter to obtain improved fit with experiment. Typical values for $N_{ss}^s$ are on the order of $1 \times 10^{10}$ cm$^{-2}$ eV$^{-1}$ [25,31,32]. The work presented here can be modified to include this effect.

The charge in the bulk consists mainly of ionized atoms and mobile majority carriers. In a p-substrate device the bulk charge $Q_b$ per unit area can be approximated by

$$Q_b = \int_0^\infty q(P-N_{AA})dx$$  \hspace{1cm} (4.7)

To solve this integral equation, one can change the variable of integration to the potential $V_I(x)$ by using the solution to Poisson's equation for the electric field. This procedure requires numerical integration of (4.7). In the present analysis we will obtain an analytic solution by assuming that, because of its "spike-like" distribution [33], the mobile electron charge in the channel has a negligible effect on the potential distribution. Although this approximation is only strictly valid under depletion or weak inversion conditions, it serves also as a good approximation under strong inversion conditions because the major contribution to $Q_b$ in strong inversion comes from the uncompensated and ionized impurities in the depletion layer [34].

Using the approach described above, we obtain
\[
\frac{Q_b}{C_O} = -K \left[ \psi_S - V_B + \frac{kT}{q} \left( e^{-\beta (\psi_S - V_B)} - 1 \right) \right]^{1/2}
\]

\[
= -K \left( \psi_S - V_B - \frac{kT}{q} \right)^{1/2}
\]

Here,

\[
K = \left( \frac{2q \varepsilon S N_{AA}}{C_O^2} \right)^{1/2}
\]

is a constant that depends on fabrication parameters. The exponential term in (4.8) results from integrating the contribution to the charge density of the mobile holes in the substrate \( P/N_{AA} = \exp[-\beta (\psi_S - V_B)] \). This yields \( P_S/N_{AA} = \exp[-\beta (\psi_S - V_B)] \) where \( P_S \) is the density of holes at the surface. For the regions of interest, depletion to strong inversion, this exponential term can be neglected.

The charge in the inversion channel, defined by (4.3), the charge in the gate, and the bulk charge are all related through a one-dimensional Gauss' law, which requires

\[
Q_n + Q_g + Q_b = 0
\]

The total charge components are obtained by integrating \( Q_g, Q_b \) and \( Q_n \) along the channel:

Total gate charge,

\[
Q_G = Z \int_0^L Q_g \, dy.
\]
Total substrate charge,

$$Q_B = Z \int_0^L Q_b \, dy \quad (4.12)$$

Total inversion charge,

$$Q_N = -(Q_G + Q_B) \quad (4.13)$$

Or, alternatively, we may change the variable of integration to the channel voltage $V$ by using (4.2),

$$Q_G = \frac{-ZL}{I_D^*} \int_{V_S}^{V_D} Q_g Q_n \, dV$$

$$Q_B = \frac{-ZL}{I_D^*} \int_{V_S}^{V_D} Q_b Q_n \, dV \quad (4.14)$$

$$Q_N = \frac{-ZL}{I_D^*} \int_{V_S}^{V_D} Q^2_n \, dV$$

where

$$I_D^* = \frac{I_D}{2 \mu_n C_o L} \quad (4.15)$$

is a normalized drain current. The dimensions of $I_D^*$ are: (volts)$^2$. Similar expressions have been obtained by Cobbold [7] by assuming drift only. In contrast, (4.14) includes the effect of drift and diffusion which, as we shall see, is necessary in obtaining a model for the complete operating range of the MOSFET.
4.2.3 Surface Potential

The complete characterization of the charge components per unit area $Q_g$ and $Q_b$ requires the functional relation between the surface potential $\psi_S$ and the applied voltages. This relation is established by applying Gauss Law, ignoring the $y$-component of electric field, which requires that the effective charge in the gate be the source of the $x$-directed electric field in the semiconductor. That is,

$$Q_g = \varepsilon_S E_x \bigg|_{x=0} = K_{C_0} F(\psi_S, V, V_B, \phi_F) . \quad (4.16)$$

The function $F(\psi_S, V, V_B, \phi_F)$ is the normalized electric field at the surface obtained from the solution to Poisson's equation. This solution has been worked out by several authors for the case when $V_B = 0$ [24]. If extended now to the case when a bias voltage $V_B$ is applied, we find that

$$F(\psi_S, V, V_B, \phi_F) = \left( \frac{kT}{q} \right)^{1/2} \left[ e^{-\beta (\psi_S - V_B)} + \frac{\psi_S - V_B}{kT/q} - 1 \right]$$
$$+ e^{\frac{\beta (\psi_S - V - 2\phi_F)}{kT/q}} - \psi_S - V_B - e^{-\beta (V - V_B + 2\phi_F)} \right]^{1/2} \quad (4.17)$$

For the usual substrate doping, $V - V_B + 2\phi_F$ is always much larger than $kT/q$. Furthermore, if we neglect the majority carrier concentration at the surface ($p_S < N_S$), which is a good approximation in both the depletion and inversion modes, one can show from (4.17) that (4.16) reduces to

$$V_G' - \psi_S = K \left[ \psi_S - V_B + \frac{kT}{q} \left( e^{\beta (\psi_S - V - 2\phi_F)} - 1 \right) \right]^{1/2} \quad (4.18)$$
where

\[ V'_G = V_G - \psi_{MS} + \frac{Q_{ox}}{C_o} + \frac{Q_{ss}}{C_o}. \]  

(4.19)

The solution of the integral equations defining the current (4.4) and the charge components (4.14), in which the variable of integration is the channel voltage \( V \), requires the functional relation between \( \psi_S \) and \( V \). This relationship, however, has not been found in closed form and hence the possibility of direct integration of (4.4) and (4.5) is excluded. A numerical integration can be performed [24] but, because of the large computer times involved, we will look for an approximation that will yield an analytic solution.

Let us consider some important characteristics of the functional relation between \( \psi_S \) and \( V \) that will set the basis for our approach. Figure 4.2 shows the solution for \( \psi_S \) obtained from (4.18) for a specific device having \( x_o = 2000 \) Å and \( N_{AA} = 1 \times 10^{15} \) cm\(^{-3}\). Figure 4.2 shows that, for values of \( V'_G \) for which \( \psi_S(0) \) is below \( 2\phi_F \), \( \psi_S \) is nearly independent of \( V \). For \( V'_G \) such that \( \psi_S(0) > 2\phi_F \), \( \psi_S \) increases almost linearly with \( V \) provided as it is shown below, that drift dominates in determining the channel current. For \( V \) greater than a certain critical voltage, however, diffusion begins to dominate and \( d\psi_S/dV \rightarrow 0 \). This characteristic behavior can be explained by studying the relative importance of the drift and diffusion components along the channel [24]:
Fig. 4.2 Surface potential $\psi_s$ as function of channel voltage $V$. 
\[
\frac{D_n}{\nu_n} \frac{dN}{dy} = \frac{1 - \frac{d\psi_S}{dV}}{\frac{d\psi_S}{dV}} \approx \frac{N_{\text{AA}} - P_S}{N_S} \approx \frac{N_{\text{AA}}}{N_S}
\]  

which is derived in Appendix B. For \(V'_G\) such that \(\psi_S(0) < 2\phi_F\), the channel is weakly inverted \((N_S << N_{\text{AA}})\) diffusion dominates, and (4.20) implies that \(d\psi_S/dV \to 0\). When \(V'_G\) is such that \(\psi_S(0) > 2\phi_F\) the channel near the source is strongly inverted \((N_S(0) >> N_{\text{AA}})\); then near the source, (4.20) implies that drift dominates and thus \(d\psi_S/dV \to 1\). As we move toward the drain, the electron concentration decreases, the channel becomes weakly inverted and there again diffusion dominates and \(d\psi_S/dV \to 0\). The channel voltage for which the channel becomes weakly inverted corresponds approximately in the strong inversion theory \([7]\) to the pinch-off voltage. At higher gate voltages, the channel remains strongly inverted in its entire length and drift is the main mechanism. In the strong inversion theory this corresponds to nonsaturated operation.

The behavior of \(\psi_S\) as described above has been used to establish two approximations often used in characterizing MOSFET behavior: the strong inversion and the weak inversion approximations. In the strong inversion approximation, which is applied when \(N_S(0) >> N_{\text{AA}}\), the surface potential is assumed to be related to the channel voltage by \(\psi_S = V + 2\phi_F\) \([7]\). Because then \(d\psi_S/dV \approx 1\), this assumption is equivalent to neglecting contributions due to diffusion mechanisms near the drain. In the weak inversion approximation, which is applied when \(N_S(0) < N_{\text{AA}}\) the surface potential is assumed to be in-
dependent of voltage $\psi_S = \psi_S(0)$ [31]. Then $d\psi_S/dV = 0$, and, therefore, drift mechanisms near the source are neglected. Although these two approximations produce satisfactory agreement with experiment in the strong and weak inversion limits, they fail for moderate inversion ($N_S^{-} = N_{AA}^{-}$) where neither of the criteria used in strong or weak inversion can be applied.

In the following section we will relax the strong and weak inversion approximations by using the basic properties of $d\psi_S/dV$. As shown in the previous discussion, these properties relate to the degree of inversion in the channel. As we shall see, the resulting model not only will merge the operation in the strong and weak inversion modes, but also will provide a first-order approximation for moderate inversion.

4.3 Drain Current and Charge Components in a Model Merging Weak, Moderate, and Strong Inversion

4.3.1 Drain Current

In Section 4.2 we found that the drain current could be expressed as

$$I^*_D = \frac{I_D}{\frac{Z_{||}^n C_O}{L}} = - \int_{V_S}^{V_D} Q'_n \, dV \quad (4.21)$$

Here, and in the rest of the chapter, the notation $Q'$ is used to designate a charge component divided by the oxide capacitance.
per unit area \( C_0 \). The dimensions of \( Q' \) are volts. By using the condition of charge neutrality \( Q_n = -(Q_g + Q_b) \), (4.21) can be rewritten as

\[
I_D^* = \int_{V_S}^{V_D} Q'_g \, dV + \int_{V_S}^{V_D} Q'_b \, dV \quad (4.22)
\]

In equation (4.22) a very convenient change of variables can be introduced by noting from (4.5) and (4.8) that

\[
\frac{dQ'_g}{dy} = \frac{d}{dy} (V'_G - \psi_S) = -\frac{d\psi_S}{dv} \frac{dv}{dy}
\]

\[
\frac{dQ'_b}{dy} = \frac{d}{dy} \left[ -K \left( \psi_S - V_B - \frac{kT}{q} \right)^{1/2} \right] = \frac{K^2}{2Q'_b} \frac{d\psi_S}{dv} \frac{dv}{dy}
\]

Thus,

\[
dv = \frac{-dQ'_g}{(d\psi_S/dv)} = \frac{2Q'_b/K^2}{(d\psi_S/dv)} \quad (4.23)
\]

Substituting (4.24) in the expression for the current, we obtain

\[
I_D^* = \int_{Q'_g(V_S)}^{Q'_g(V_D)} \frac{-Q_g}{(d\psi_S/dv)} \, dQ'_g + \int_{Q'_b(V_S)}^{Q'_b(V_D)} \frac{2Q'_b^2/K^2}{(d\psi_S/dv)} \, dQ'_b \quad (4.25)
\]

Figure 4.3 shows the elements constituting the integrands in (4.25) for a specified device operating in the pinch-off mode. This represents the most general case because the channel is strongly inverted at the source and becomes
Fig. 4.3 Components of charge per unit area and surface potential as functions of the channel voltage $V$. 
weakly inverted toward the drain. As discussed in Section 4.2.3, $d\psi_S/dV$ has almost constant values along the channel; in the strongly inverted portion $d\psi_S/dV = 1$ while in the weakly inverted portion $d\psi_S/dV = 0$. In the transition between strongly and weakly inverted regions, where the channel is moderately inverted, $d\psi_S/dV$ is not constant. However, because this represents a small portion of the characteristic ($\psi_S$ vs. $V$), we will assume in a first-order approximation that $\psi_S$ is there linearly related to $V$ with the value for the slope $d\psi_S/dV$ lying between 0 and 1.

Our approach will consist then in dividing the channel into three regions by defining appropriate limits $V_1$ and $V_2$ as shown in Figure 4.3. Below $V_1$ the channel will be assumed to be strongly inverted with $d\psi_S/dV = S_S$, a constant. Above $V_2$ we will consider the channel to be weakly inverted with $d\psi_S/dV = S_W$, a constant. In the transition region the channel will be assumed to be moderately inverted with $d\psi_S/dV = S_M$, also a constant. These approximations allow us to write the expression for the current as the sum of the contributions in each region. Furthermore, because $d\psi_S/dV$ is assumed constant in each case, it can be taken out of the integrals which can then be directly evaluated. If we define a function, $F_I$, related to $I_D$ by
\[ F_I(V_a', V_b') = \int_{Q_g'(V_a)}^{Q_g'(V_b)} - Q_g' \, dQ_g' + \int_{Q_b'(V_a)}^{Q_b'(V_b)} 2Q_b'^2/K^2 \, dQ_b' \]

\[ = \frac{-Q_g'^2}{2} \int_{V_a}^{V_b} + \frac{2}{3} \frac{Q_b'^3}{K^2} \int_{V_a}^{V_b} \]

\[ = \frac{Q_g'^2}{2} \int_{V_a}^{V_b} + \frac{2}{3} \frac{Q_b'^3}{K^2} \int_{V_a}^{V_b} \]

The three components of (4.27) result from carrying out the details of the integration indicated in (4.21). Here, if we let \( V_1 = V_D \) and \( S_S = 1 \), (4.27) reduces to the conventional expression (obtained by using the strong inversion approximation) for the drain current of a device operating in the triode mode.

In computing the drain current from (4.27), a numerical problem could occur in evaluating the term corresponding to the weakly inverted channel because \( S_W \) is very small. To avoid this problem an alternate form for this term can be obtained as follows. The channel charge \( Q_n \) was defined in (4.3) as

\[ Q_n = -q \int_0^\infty N(x, y) \, dx \]

Taking derivatives on both sides with respect to \( y \) yields
\[
\frac{dQ_n^o}{dV} = -q \int_0^\infty \frac{dN}{dy} (x,y) dx \tag{4.28}
\]

but, because \( N = n_i \exp[\beta(V_I-V_N)] = n_i \exp[\beta(V-I-V_p)] \), it follows from the gradual approximation \[19\] that

\[
\frac{dN}{dy} = -\frac{N}{kT/q} (1 - \frac{d\psi_s}{dV}) \frac{dV}{dy} \tag{4.29}
\]

Substituting (4.29) in (4.28), using the definition of \( Q_n^o \) and reordering the terms, we obtain

\[
dV = -\frac{kT/q}{Q_n^o} \frac{dQ_n^o'}{(1 - \frac{d\psi_s}{dV})} \tag{4.30}
\]

From (4.30) the contribution to the drain current from the weakly inverted channel can then be alternatively written as

\[
I_{DW}^* = -\int_{V_2}^{V_D} Q_n^o dV
= \frac{kT}{q} \int Q_n^o(V_D) \frac{dQ_n^o'}{(1 - \frac{d\psi_s}{dV})} \tag{4.31}
\]

But since we are assuming that \( d\psi_s/dV \) has a constant value \( S \) in this region, we finally obtain:

\[
I_{DW}^* = \frac{kT}{q} \frac{Q_n^o'}{(1 - S)} \left[ V_D \right]_{V_2}^{V_D} \tag{4.32}
\]
Here, if we let $S_W = 0$ and $V_2 = V_S$, (4.32) reduces to the conventional expression for the drain current of a device operating in weak inversion [31].

### 4.3.2 Charge Components

The procedure to calculate the total charge components is entirely analogous to the one presented for the drain current. Combining (4.10) and (4.14) and using the change of variable indicated in (4.24), we obtain for the total charge components

$$Q_G' = \frac{ZL}{I_D} \int_{V_S}^{V_D} (Q_g' + Q_g'O_b')dV$$

$$= \frac{ZL}{I_D} \left[ Q_g'(V_D) - \frac{Q_g'2}{(1-d\psi_S/dV)} + \frac{2Q_g'O_b'/K Q_b'}{(1-d\psi_S/dV)} \right]$$

and

$$Q_B' = \frac{ZL}{I_D} \int_{V_S}^{V_D} (Q_b' + Q_g'O_b')dV$$

$$= \frac{ZL}{I_D} \left[ Q_b'(V_D) + \frac{2Q_b'3/K Q_b}{(1-d\psi_S/dV)} + \frac{2Q_g'O_b'/K Q_b'}{(1-d\psi_S/dV)} \right]$$

Again, if the channel is divided in three regions and we assume that $d\psi_S/dV$ is constant in each region, the charges can be obtained by direct integration. Let us define functions $F_{QG}$ and $F_{QB}$ such that,
\[ F_{QG}(V_a, V_b) = \int_{Q_g(V_a)}^{Q_g(V_b)} - Q_g'^2 \, dQ_g' + \int_{Q_b'(V_a)}^{Q_b'(V_b)} 2Q_g'Q_b'^2/K^2 \, dQ_b' \]

\[ = \left[ \frac{-Q_g'^3}{3} + \frac{2}{3K^2} \left( Q_g'^2 + \frac{2Q_b'^5}{K^5} \right) \right]^{V_b}_{V_a} \]  

(4.35)

and

\[ F_{QB}(V_a, V_b) = \int_{Q_b'(V_a)}^{Q_b'(V_b)} 2Q_b'^3/K^2 \, dQ_b' + \int_{Q_b'(V_a)}^{Q_b'(V_b)} 2Q_g'Q_b'^2/K^2 \, dQ_b' \]

\[ = \left[ \frac{1}{2K^2} Q_b'^4 + \frac{2}{3K^2} \left( Q_g'^2 + \frac{2Q_b'^5}{K^5} \right) \right]^{V_b}_{V_a} \]  

(4.36)

where the second integral in (4.35) and (4.36) was evaluated using integration by parts with \( u = Q_g' \) and \( dv = Q_b'^2 \, dQ_b' \).

Then the total charges can be expressed as

\[ Q_G' = \frac{ZL}{t_D} \left[ \frac{F_{QG}(V_S, V_1)}{S_S} + \frac{F_{QG}(V_1, V_2)}{S_M} + \frac{F_{QG}(V_2, V_D)}{S_W} \right] \]  

(4.37)

\[ Q_B' = \frac{ZL}{t_D} \left[ \frac{F_{QB}(V_S, V_1)}{S_S} + \frac{F_{QB}(V_1, V_2)}{S_M} + \frac{F_{QB}(V_2, V_D)}{S_W} \right] \]  

(4.38)

As in the case of the drain current, to avoid numerical problems due to the smallness of \( S_W \), an alternative expression can be obtained for the contribution of the weakly inverted portions of the channel. Using (4.30) directly in (4.14), we obtain
\[ Q_{GW}^' = \frac{ZL}{I_D^*} \int \frac{Q_g^'}{q} \frac{kT}{(1-d\psi_S/dV)} dQ_n^' \]

\[ = - \frac{ZL}{I_D^*} \frac{kT/q}{(1-S_W)} \left[ \frac{Q_g^2}{2} + Q_g^'Q_b^' + \frac{2}{3} \frac{Q_b^3}{K^2} \right]_{V_2}^{V_D} \]  \hspace{1cm} (4.39)

Here we have used integration by parts with \( u = Q_g^' \) and \( dv = dQ_b^' \). A similar expression results for the bulk charge in weak inversion,

\[ Q_{BW}^' = \frac{ZL}{I_D^*} \int \frac{Q_b^'}{q} \frac{Q_b^'}{(1-d\psi_S/dV)} dQ_n^' \]

\[ = \frac{ZL}{I_D^*} \frac{kT/q}{(1-S_W)} \left[ \frac{2}{3} \frac{Q_b^3}{K^2} - \frac{Q_b^2}{2} \right]_{V_2}^{V_D} \]  \hspace{1cm} (4.40)

### 4.3.3 Limits for the Strong, Weak, and Moderately Inverted Portions of the Channel

The three-region piecewise-linear approximation employed in Sections 4.3.1 and 4.3.2 to obtain expressions for the current and charges uses two parameters: (1) the limits \( V_1 \) and \( V_2 \) that divide the strong, moderate, and weakly inverted portions of the channel; and (2) the approximate values at the slope \( d\psi_S/dV \) \( (S_S', S_M', S_W') \) in each of the three regions. These parameters will be now defined in terms of the applied external voltages.
In Section 4.2.3 we concluded that $\frac{d\psi_s}{dV}$ could be considered as a measure of the level of inversion along the channel. Here we will show that it is also the ratio of the contribution of the drift current to the total current. In (4.20) we indicated that

$$\frac{I_{\text{DIFF}}}{I_{\text{DRIFT}}} = \frac{1 - \frac{d\psi_s}{dV}}{\frac{d\psi_s}{dV}}$$

(4.20)

Thus, rearranging terms we obtain

$$\frac{I_{\text{DRIFT}}}{I_{\text{DRIFT}} + I_{\text{DIFF}}} = \frac{d\psi_s}{dV}$$

(4.41)

We will use this property of $\frac{d\psi_s}{dV}$ to define quantitatively the voltages $V_1$ and $V_2$ as follows.

In the strongly inverted regions we previously observed that $\frac{d\psi_s}{dV}$ is close to unity and drift dominates while in the weakly inverted regions diffusion dominates with $\frac{d\psi_s}{dV}$ being close to zero. Thus we will define the transition region corresponding to moderate inversion as the region in which both drift and diffusion are comparable. More specifically, we will define $V_1$ as the channel voltage at which the drift current constitutes 80% of the total current and $V_2$ as the channel voltage for which the drift component is 20% of the total current. This specification of $V_1$ and $V_2$ provides, approximately, the best least-squares fit between the piecewise linear approximation and the $\psi_s$ versus $V$ characteristic. Based on these definitions we can now obtain expressions for $V_1$ and $V_2$ by solving
\[
\frac{d\psi_S}{dV} = A \tag{4.42}
\]

where \( A \) has the value \( A = 0.80 \) when solving for \( V_1 \) and \( A = 0.20 \) when solving for \( V_2 \). Differentiation of both sides of (4.18) yields

\[
\frac{d\psi_S}{dV} = \frac{\beta(\psi_S - V - 2\phi_F)}{2\Omega g^2 + 1 + e^{\beta(\psi_S - V - 2\phi_F)}} = A \tag{4.43}
\]

Combining (4.43) and (4.18) and using the definition of \( \theta_g \), we find that

\[
V_1V_2 = V_G' - 2\phi_F - \frac{kT}{q} \ln \frac{2V_x}{A'K^2} + \frac{K^2}{2} - V_x
\]

where

\[
V_x = K \left[ V_G' - V_B + \frac{k^2}{4} + \frac{kT}{q} \left( \frac{1}{(A'K)^2} - 1 \right) \right]^{1/2} + \frac{kT/q}{A}
\]

Here, \( A' = (1-A)/A \). Hence, \( A' = 1/4 \) when calculating \( V_1 \) and \( A' = 4 \) when calculating \( V_2 \). Equation (4.44) applies only when \( V_S < V_1, V_2 < V_D \). The complete functional dependencies for \( V_1 \) and \( V_2 \) are given by

\[
V_1 = \begin{cases} 
V_1 & \text{from (4.44) if } V_S < V_1 < V_D \\
V_S & \text{if } V_1 < V_S \\
V_D & \text{if } V_1 > V_D 
\end{cases}
\]

\[
V_2 = \begin{cases} 
V_2 & \text{from (4.44) if } V_S < V_2 < V_D \\
V_S & \text{if } V_2 < V_S \\
V_D & \text{if } V_2 > V_D 
\end{cases}
\]
Using the functional dependencies for $V_1$ and $V_2$ given by (4.45), we now can solve (4.18) to obtain $\psi_S$ at the limits $V_1$ and $V_2$. The surface potential at those points can be used to define the approximate slopes $d\psi_S/dV$ in each region, which constitute the second parameter at our three-region piecewise linear approximation. They are,

$$S_S = \frac{\psi_S(V_S) - \psi_S(V_1)}{V_1 - V_S}$$

$$S_M = \frac{\psi_S(V_2) - \psi_S(V_1)}{V_2 - V_1}$$

$$S_W = \frac{\psi_S(V_D) - \psi_S(V_2)}{V_D - V_2}$$

(4.46)

Here, to calculate $\psi_S(V_S)$, $\psi_S(V_1)$, $\psi_S(V_2)$ and $\psi_S(V_D)$, one needs to solve (4.18) numerically. This process does not require much computer time. We used the Newton-Raphson method [35] to calculate the solution and found that less than five iterations were necessary to achieve convergence.

4.4 Results and Evaluation of the Model

Table 2 summarizes the results of the model merging weak, moderate and strong inversion. In Figures 4.4 through 4.7 we illustrate the characteristics for the drain current and the total charge components obtained from the proposed model. Notice that the curves in these characteristics and their slopes are continuous throughout the entire range of
Table 2 Drain Current and Total Charge Components

DRAIN CURRENT

\[
\frac{I_D}{I_{nC_{oL}}} = I_D^* = \frac{F_I(V_S', V_1)}{S_S} + \frac{F_I(V_1', V_2)}{S_M} + I_{DN}^*
\]

TOTAL CHARGE COMPONENTS

\[
\frac{Q_G}{Z_{LC_o}} = \frac{1}{I_D^*} \left[ \frac{F_QG(V_S', V_1)}{S_S} + \frac{F_QB(V_1', V_2)}{S_M} \right] + \frac{Q_{GN}}{Z_{LC_o}}
\]

\[
\frac{Q_B}{Z_{LC_o}} = \frac{1}{I_D^*} \left[ \frac{F_QB(V_S', V_1)}{S_S} + \frac{F_QB(V_1', V_2)}{S_M} \right] + \frac{Q_{BW}}{Z_{LC_o}}
\]

\[Q_N + Q_G + Q_B = 0\]
FOR THE CURRENT:

\[ F_I(V_a, V_b) = \left[ -\frac{Q_g'}{2} + \frac{2}{3} \frac{Q_b'}{K^2} \right]_{V_a}^{V_b} \]

\[ I_{DW}^* = \frac{kT/q}{1-S_W} \left[ \left. \frac{V_D}{\Omega'_{n1}} \right|_{V_2} \right. \]

FOR THE CHARGES:

\[ F_{QG}(V_a, V_b) = \left[ -\frac{Q_g'}{3} + \frac{2}{3K^2} (Q_g Q_b' + 2Q_b'^5/K^5) \right]_{V_a}^{V_b} \]

\[ F_{QB}(V_a, V_b) = \left[ \frac{Q_b'^4}{2K^2} + \frac{2}{3K^2} (Q_g Q_b' + 2Q_b'^5/K^5) \right]_{V_a}^{V_b} \]

\[ \frac{Q_{GW}}{Z_{LC_o}} = -\frac{1}{I_D^*} \frac{kT/q}{1-S_W} \left[ \frac{Q_g'}{2} + Q_g Q_b + \frac{2}{3} \frac{Q_b'}{K^2} \right]_{V_2}^{V_D} \]

\[ \frac{Q_{BW}}{Z_{LC_o}} = \frac{1}{I_D^*} \frac{kT/q}{1-S_W} \left[ \frac{2}{3} \frac{Q_b'}{K^2} - \frac{Q_b'}{2} \right]_{V_2}^{V_D} \]
Fig. 4.4 Calculated square-root dependence of the drain current on gate voltage.
Fig. 4.5 Calculated drain current as function of gate voltage for three doping concentrations ($x_0 = 2000\,\text{Å}$).
Fig. 4.6 Calculated drain current characteristics in weak and moderate inversion.
Fig. 4.7 Calculated charge components as function of gate voltage.
operation. This feature results from including the transition region for moderate inversion, which is not included in previous work treating weak [31] and strong [7] inversion.

Figure 4.4 shows that for strong inversion the functional relation between the drain current and the gate voltage follows a square law [20], while in weak inversion this relation is exponential, as shown in Figure 4.5. This behavior agrees qualitatively with previous models for the extremes of strong and weak inversions.

In Figure 4.6 the drain current is shown as a function of the drain voltage for weak and moderate inversion. The inclusion of drift and diffusion in our model has produced a smooth transition into saturation. The necessity of including diffusion to produce this smooth transition was first recognized by Pao and Sah [24].

The total charge components are shown in Figure 4.7 as functions of the gate voltage. Notice that the inversion charge increases exponentially at low gate voltages. The relationship between the charge components and the terminal voltages has apparently not been established previously for weak and moderate inversion. As is demonstrated in the next chapter, these relationships provide a basis for characterization of the device capacitances and the displacement currents.

In assessing the validity of our modeling approach and the accuracy of the expressions developed for the current and charges, we compare the results of our model against
results from previous theoretical treatments. Figure 4.8 shows experimental data for the square root of the drain current against gate voltage obtained in a commercial device (4007) having $N_{AA} = 3 \times 10^{15}$ cm$^{-3}$ and $x_0 = 1000\AA$. In this figure we also show the calculated characteristics obtained from the model just derived. Since $Q_{SS}$ and other fabrication parameters are not accurately known for this device, the calculated and the observed characteristics were matched using the value of the voltage and current at the extrapolated threshold voltage. Good agreement between experiment and theory is observed. We also show in Figure 4.8 theoretical characteristics obtained from a model using the strong inversion approximation [7]. The discrepancy at low gate voltages between this model and the experimental data arise because the strong inversion approximation assumes that an abrupt transition between depletion and inversion occurs when the surface potential at the source is equal to $2\phi_F$. This results in a discontinuity in the slope of the characteristics at the boundary between cut-off and saturation. A discrepancy also exists at high gate voltages. This arises because the surface potential, which in the strong inversion approximation is assumed independent of gate voltage, is in fact a logarithmic function of $V_G$. As one can show from equation (4.18) for $V_G > V_T$, this function can be approximated by

$$
\psi_S(0) = 2\phi_F + \frac{kT}{q} \log \left[ \frac{V_G^2 - V_G' - K^2(2\phi_F - V_B)}{k^2kT/q} \right]. 
$$

(4.47)
Experiment - Our model -- Strong inversion model [7]

\[ N_{AA} = 3 \times 10^{15} \text{ cm}^{-3} \]
\[ x_0 = 1000 \text{ Å} \]
\[ V_D = 2 \text{ v} \]
\[ V_B = 0 \text{ v} \]

Fig. 4.8 Experimental values for the drain current compared with values calculated using our model and using a model for strong inversion.
Fig. 4.9 Experimental values for the drain current compared with values calculated using our model, using a model for strong inversion and using a model for weak inversion.
Fig. 4.10 Experimental values for the drain current compared with values calculated using our model and using a previous model ($N_{AA} = 7 \times 10^{14}, x_o = 1470\mu$).

- Experiment [26]
- Our model
-- Previous model [26]
The proposed new model includes implicitly this dependence of $\psi_S$ in $V_G$.

Figures 4.9 and 4.10, which compare the predictions of our model with experimental data from the literature [26], show excellent agreement. Because information was available only for the doping concentration $N_{AA}$ and oxide thickness $x_o$ in this device, the calculated and the experimental characteristics were matched using the value of the gate voltage and drain current at the extrapolated threshold. In Figure 4.9 we show for comparison previous models obtained for weak inversion [31] and for strong inversion [7]. In Figure 4.10 we compare our model against a recently developed model for the entire range of operation [26]. Although this model shows good agreement with experiment in the weak and strong inversion limits, it fails for gate voltages near the transition region ($V_G \approx -0.1v$). Furthermore notice the discontinuities in the slope of the characteristics which our model avoids.

4.5 Conclusions

The major achievement of this chapter is the analytical description given in Table 2 that unifies weak, moderate and strong inversion and covers the cut-off, triode and saturation modes of operation. This description has the following properties:

(1) It includes the effects of substrate bias which enables the representation of four-terminal properties of the MOSFET.
(2) It includes the charges in the gate, channel and substrate regions as well as the drain current. These charges provide the basis for modeling capacitive effects.

(3) It consists of simple expressions having continuous derivatives with respect to the terminal voltages. This helps make the description useful for computer-aided circuit analysis.

The model developed here is subject to the limitations of the one-dimensional gradual channel approximation which become severe in MOSFET structures with short channel lengths. Other limitations arise from the idealizations used in Section 4.2: effective channel length, field independent mobility and effective charge in surface states. A number of publications in the technical literature deal with more detailed descriptions of these parameters and also with short-channel effects. As explained in Section 4.2, our model has enough flexibility to incorporate these descriptions.
5.1 Introduction

In Chapter III we developed an equivalent-circuit representation for the transient response of the MOSFET. By employing the results of Chapter IV, the functional dependencies of each element in this equivalent-circuit will be now derived in terms of the applied voltages and the fabrication parameters of the device. The main approximation used in deriving such dependencies is a quasi-static approximation through which, as discussed in Chapters II and III, one extends the knowledge of the dc steady-state behavior of the device to describe its large-signal transient response.

The equivalent-circuit for the intrinsic MOSFET derived in Chapter III is shown in Fig. 3.3. The definition for each element in the circuit is given in Table 1. Three types of elements are present: a current source between drain and source representing charge transport, and capacitors and transcapacitors connected between each node representing charge accumulation within the device. In Sections 5.2 through 5.4, the functional dependence of each of
these elements is derived. The resulting mathematical expressions are valid for the entire range of operation of the MOSFET, and include the effect of the substrate terminal. Such expressions are new.

This chapter also provides the first detailed discussion of the intrinsic capacitive effects of the substrate and the transcapacitive effects due to the non-symmetry of the four-terminal MOSFET. In Sections 5.3.4 and 5.4.2 we discuss the engineering importance of these two effects. Under certain conditions determined by the particular circuit environment in which the device is used the equivalent network representation can be simplified. An example is discussed in Section 5.4.3.

5.2 Source-Drain Current Source

Through the use of a quasi-static approximation, as discussed in Chapter III, the functional dependence of the nonlinear source-drain current source can be determined by extrapolating the static characteristics of the drain current found in Section 4.3.1. Thus,

\[
    i_{SD} = -I_D(v_S, v_D, v_G, v_B),
\]

which has the same functional dependencies on the terminal voltage as those describing the dc steady-state.
5.3 Capacitances

5.3.1 Expressions for the Capacitances

The capacitors in the equivalent-circuit are defined in Table 1 as the partial derivatives with respect to voltage of the time varying total charge components \( q_G, q_B, q_N \). As in the case of the transport current \( i_{SD} \) a quasi-static approximation allows us to write

\[
\begin{align*}
q_G &= Q_G(v_S, v_D, v_G, v_B) \\
q_B &= Q_N(v_S, v_D, v_G, v_B) \\
q_N &= -(q_G + q_B)
\end{align*}
\]  

(5.2)

One can anticipate that a partial differentiation of (5.2) with respect to the voltages would lead to very complicated expressions. But we will now show that because of the systematic approach used in Chapter III to define the circuit elements, one can find simple expressions for the functional dependencies of the capacitors.

From Table 1 the capacitors connected to the source are

\[ C_{SG} = -\frac{\partial Q_G}{\partial v_S} \]  

(5.3)

and

\[ C_{SB} = -\frac{\partial Q_B}{\partial v_S} \]  

(5.4)

We can use (4.14) to rewrite \( C_{SG} \)
\[
\frac{C_{SG}}{Z_{LC}} = \frac{\partial}{\partial V_S} \left( \frac{1}{I_D} \int_{V_S}^{V_D} Q' - Q'_{n} \, dv \right)
\]  

(5.5)

where \(Q'\) denotes a charge per unit area normalized by the oxide capacitance \(C_o\) (the dimensions of \(Q'\) are volts).

Using chain rule differentiation and the fundamental theorem of integral calculus,

\[
\frac{C_{SG}}{Z_{LC}} = \frac{1}{I_D} \left( Q'_G \frac{\partial I_D^*}{\partial V_S} - Q'_g(V_S) Q'_{n}(V_S) \right).
\]  

(5.6)

But since

\[
\frac{\partial I_D^*}{\partial V_S} = - \frac{\partial}{\partial V_S} \int_{V_S}^{V_D} Q'_n \, dv = Q'_n(V_S)
\]  

(5.7)

we finally obtain

\[
\frac{C_{SG}}{Z_{LC}} = \frac{Q'_n(0)}{I_D^*} \left( Q'_G - Q'_g(0) \right)
\]  

(5.8)

where \(Q'_n(0)\) and \(Q'_g(0)\) are the normalized and gate charge per unit area, given by (4.5) and (4.8), evaluated at the source end \((y=0)\). Similarly

\[
\frac{C_{SB}}{Z_{LC}} = \frac{Q'_n(0)}{I_D^*} \left( Q'_B - Q'_b(0) \right)
\]  

(5.9)

For the capacitances connected to the drain, the approach is the same except that
Thus, we obtain
\[
\frac{\partial I_D^*}{\partial V_D} = - \frac{\partial}{\partial V_D} \int_{V_S}^{V_D} Q_n' \, dV = -Q_n'(L).
\] (5.10)

\[
C_{DG} = \frac{Q_n'(L)}{I_D^*} \left[ Q_G - Q_g'(L) \right],
\] (5.11)

and
\[
C_{DB} = \frac{Q_n'(L)}{I_D^*} \left[ Q_B' - Q_b'(L) \right],
\] (5.12)

where \(Q_n'(L), Q_g'(L)\) and \(Q_b'(L)\) are the normalized channel, gate and substrate charge evaluated at the drain end \((y=L)\).

The gate-substrate capacitance is defined in Table 1 as
\[
C_{GB} = - \frac{\partial q_B}{\partial V_G} = - \frac{\partial Q_B}{\partial V_G}.
\] (5.13)

Substituting (4.14), which gives the functional relation for \(Q_B\), and applying the chain rule for differentiation yields
\[
C_{GB} = \frac{1}{I_D^*} \left[ Q_B \frac{\partial I_D^*}{\partial V_G} + \frac{\partial}{\partial V_G} \int_{V_S}^{V_D} Q_b Q_n' \, dV \right].
\] (5.14)

The expression for \(C_{GB}\) is more complicated than those for \(C_{SG}, C_{DG}, C_{SB}\), and \(C_{DB}\). To find this expression we take the partial derivatives, \(\partial I_D^*/\partial V_G\) and \(\partial/\partial V_G \left( \int_{V_S}^{V_D} Q_b Q_n' \, dV \right)\), using (4.27) and (4.36). The procedure is straightforward, and the results follow:
Here we have defined the functions $DF_I$ and $DF_{QB}$ as:

$$DF_I(v_a', v_b) = -\left[ Q'_g + \frac{Q'_n}{1 + \frac{K^2}{2Q'_g}} \left( 1 + e^{\frac{\beta(\psi_S - \psi_{-2} - 2\phi_F)}} \right) \right] v_a$$

$$DF_{QB}(v_a', v_b) = -\left[ \frac{2}{3} \frac{Q'_b}{K^2} - \frac{\frac{Q'_b}{Q'_n}}{1 + \frac{K^2}{2Q'_g}} \left( 1 + e^{\frac{\beta(\psi_S - \psi_{-2} - 2\phi_F)}} \right) \right] v_a$$

An alternative form for the contributions of the weakly inverted portions of the channel results from taking partial derivatives with respect to $v_G$ in (4.32) and (4.40). This yields,

$$\frac{\partial I^*_DW}{\partial v_G} = \frac{DF_I(v_2', v_D)}{S_w} = \frac{kT/q}{1-S_w} \left[ 1 + \frac{K^2/Q'_g}{1 + \frac{K^2}{2Q'_g}} \right] v_D$$

and

$$\frac{DF_{QB}(v_2', v_D)}{S_w} = -\frac{kT/q}{1-S_w} \left[ Q'_b \left[ \frac{1 + \frac{K^2/Q'_g}{1 + \frac{K^2}{2Q'_g}}} \right] v_2 \right].$$

The results for the capacitors are summarized in Table 3. Figure 5.1 illustrates the functional dependencies of the
Table 3  Functional dependencies for the capacitors.

\[
\frac{C_{SG}}{ZLC_o} = \frac{q_n'(0)}{I_D} \left( \frac{Q_g'}{ZL} - Q_g'(0) \right)
\]

\[
\frac{C_{DG}}{ZLC_o} = \frac{-q_n'(L)}{I_D} \left( \frac{Q_g'}{ZL} - Q_g'(L) \right)
\]

\[
\frac{C_{SB}}{ZLC_o} = \frac{q_n'(0)}{I_D} \left( \frac{Q_B'}{ZL} - Q_B'(0) \right)
\]

\[
\frac{C_{DB}}{ZLC_o} = \frac{-q_n'(L)}{I_D} \left( \frac{Q_B'}{ZL} - Q_B'(0) \right)
\]

\[
\frac{C_{GB}}{ZLC_o} = \frac{1}{I_D^2} \left[ \frac{3I_D'}{3V_G} + \frac{3}{3V_G} \int_{V(0)}^{V(L)} Q'_b Q'_n \, dv \right]^{(*)}
\]

(*) given by (5.15) through (5.17).
Fig. 5.1 Calculated capacitances.
capacitances in a specific device. In contrast with results obtained from models using the strong inversion approximation [7,36], these curves present smooth transitions between the different regions of operation: cut-off, saturation and nonsaturation. A physically based discussion about the main features of these characteristics is given in the next section.

5.3.2 Physical Interpretation of the Results for the Capacitances

Consider first the capacitances connected to the source and drain nodes in the equivalent-circuit. These capacitances are directly related to the apportionment between the currents charging the channel from the source island and from the drain island. To observe how this apportionment occurs, let us consider the total capacitance at the source \( C_{SS} \) given by

\[
C_{SS} = C_{SG} + C_{SB} = \frac{\partial Q_N}{\partial V_S}
\]

(5.20)

and the total capacitance at the drain \( C_{DD} \) given by

\[
C_{DD} = C_{DG} + C_{DB} = \frac{\partial Q_N}{\partial V_D}
\]

(5.21)

As we shall see, the functional dependence of these capacitances shown in Fig. 5.2 has the form to be expected from the discussion of the charge apportionment in Section 3.3. In cut-off there is no charging of the channel and both \( C_{SS} \) and \( C_{DD} \) are equal to zero. As the gate voltage is increased, the channel is turned on in an exponential form (see Fig. 4.5) causing an abrupt change in \( C_{SS} \). At higher gate voltages,
Fig. 5.2 Calculated total source and drain capacitance for the device described in Fig. 5.1.
while the device is in the saturation region, $Q_N$ increases almost linearly with gate voltage and hence $C_{SS}$ is nearly constant. In the saturation region, because there is no charging of the channel from the drain end, $C_{DD} = 0$. Further increase of the gate voltage drives the device into nonsaturation. Here the channel opens gradually into the drain allowing thereafter an increasing contribution of the drain end to the charging of the channel while the contribution from the source decreases. Thus in this region, as shown in Fig. 5.2, $C_{SS}$ decreases while $C_{DD}$ increases. For very large gate voltages the charging of the channel will tend to occur equally from the drain than from the source. When this happens the values of $C_{SS}$ and $C_{DD}$ tend to one another as shown in Fig. 5.2.

A measure of the apportionment of the contributions of the drain and source islands to the charging of the channel is given by the apportionment function $\lambda$ defined in Chapter III as

$$\lambda = \frac{1}{1 + \frac{C_{DG} + C_{DB}}{C_{SG} + C_{SB}}} = \frac{1}{1 + \frac{C_{DD}}{C_{SS}}}.$$  

This function is used in the next section to obtain expressions for the transcapacitances. Its functional dependence for a particular device is shown in Fig. 5.3. In saturation, $C_{DD} = 0$ and $\lambda = 1$, while in nonsaturation the values of $C_{DD}$ and $C_{SS}$ approach one another and $\lambda$ tends to 1/2.
Fig. 5.3 Apportionment function $\lambda$ for the device described in Fig. 5.1.
Note from Fig. 5.1 the similarity between the characteristics of the substrate capacitances \( C_{SB} \) and \( C_{DB} \) and the characteristics of the gate capacitances \( C_{SG} \) and \( C_{DG} \). This similarity, which also can be observed in the expressions defining these capacitances, will be used in the next section to obtain an engineering approximation for \( C_{SB} \) and \( C_{DB} \).

Consider now the gate substrate capacitance

\[
C_{GB} = -\frac{\partial Q_B}{\partial V_G}.
\]

This capacitance is related to the control of the gate over the substrate charge. In cut-off, where \( V_G \) is not large enough to turn on the channel, this capacitance is equal to the capacitance of a (two-terminal) MOS capacitance [37]. As \( V_G \) increases, an inversion channel starts forming at the surface of the semiconductor and more field lines emanating from the gate will terminate in the inversion channel. Thus, \( C_{GB} \) will decrease as shown in Fig. 5.1. For larger gate voltages, where a strong inverted channel is formed over the entire length of the intrinsic device, the gate will exert even less control over the substrate charge and \( C_{GB} \) decreases at a faster rate reaching eventually a zero value as illustrated in Fig. 5.1.

Figures 5.4 and 5.5 show the total gate capacitance \( C_{GG} \) and the total substrate capacitance \( C_{BB} \) together with their components,

\[
C_{GG} = \frac{\partial Q_G}{\partial V_G} = -\left(\frac{\partial Q_N}{\partial V_G} + \frac{\partial Q_B}{\partial V_G}\right) \quad (5.23)
\]

\[
C_{BB} = \frac{\partial Q_B}{\partial V_B} = -\left(\frac{\partial Q_N}{\partial V_B} + \frac{\partial Q_G}{\partial V_B}\right) \quad (5.24)
\]
Fig. 5.4 Calculated total gate capacitance and its components for the device described in Fig. 5.1.
Fig. 5.5 Calculated total substrate capacitance and its components for the device described in Fig. 5.1.
In the cut-off region there is no inversion channel and $C_{gg}$ and $C_{bb}$ are equal. Their functional dependency is that of an MOS capacitance [37]. In the saturation region, the gate charge depends almost linearly on the gate voltage (Fig. 4.7), and $C_{gg}$ shows a constant value of about $2/3 ZL C_0$ as predicted by strong inversion theory [20]. In this region, as $V_G$ is increased, the surface potential increases producing a widening of the depletion layer; consequently $C_{bb}$ decreases as shown in Fig. 5.5. At the onset of the nonsaturation region $C_{gg}$ abruptly rises due to the increase of electron concentration over the entire channel length. For even larger gate voltages $C_{gg}$ approaches the value of the total oxide capacitance. In this region, $C_{bb}$ attains a constant value because the substrate charge becomes independent of gate voltage. This constant value cannot be clearly determined from the expressions of the substrate capacitances just found. In the next section, however, we discuss an approximation for the substrate capacitance that permits a good estimation of their values for engineering purposes.

The main features of the functional dependencies for the gate capacitances in the MOSFET have been predicted by previous authors [7,20] using simplified models. Our results agree qualitatively with these predictions, giving additionally a detailed and continuous description for these capacitances and also for the substrate capacitances.
5.3.3 An Engineering Approximation for the Functional Dependencies of the Intrinsic Substrate Capacitances $C_{SB}$ and $C_{DB}$

The functional dependencies for the substrate capacitances $C_{SB}$ and $C_{DB}$ were derived in Section 5.3.1. Figure 5.1 shows these functional dependencies together with the functional dependencies for the gate capacitances and the gate-bulk capacitance. We pointed out previously the similarity between the functional dependencies of the gate and substrate capacitances appearing in this figure. From an engineering point of view, this similarity is advantageous because it suggests the existence of relations of the form:

$$\begin{align*}
C_{SB} &= \alpha_S C_{SG} \\
C_{DB} &= \alpha_D C_{DG}
\end{align*}$$

(5.25)

where $\alpha_S$ and $\alpha_D$ may be simple functions of the voltages. Such relations would allow considerable simplification in the computation of the substrate capacitances. In recent engineering applications [1], $C_{SB}$ and $C_{DB}$ are modeled to a first order approximation as

$$\begin{align*}
C_{SB} &= \alpha C_{SG} \\
C_{DB} &= \alpha C_{DG}
\end{align*}$$

(5.26)

with $\alpha$ being a constant. Because expressions for $C_{SB}$ and $C_{DB}$ were not previously available this approximation has not been verified. With the functional dependencies for $C_{SB}$ and $C_{DB}$ made available in the previous section we can now study
this engineering approximation. Figure 5.6 shows \( \alpha_S \) and \( \alpha_D \), defined in (5.25), as functions of the applied voltages. Notice that although in the nonsaturation region \( \alpha_S \) and \( \alpha_D \) are practically independent of the gate voltage they are in general not constant.

Using the functional dependencies for \( C_{SB} \) and \( C_{DB} \) given in Table 3 we will now derive an improved approximation for \( \alpha_S \) and \( \alpha_D \) that shows a better functional dependence on the applied voltages while remaining a simple function of the voltages.

Consider first

\[
\alpha_S = \frac{C_{SB}}{C_{SG}} = \frac{Q'_B / Z_L - Q'_b (0)}{Q'_G / Z_L - Q'_g (0)}
\]

(5.27)

Substituting the expression for \( Q'_B \), \( Q'_G \) and \( I_D \) given in (4.4) and (4.14) \( \alpha_S \) can be rewritten as

\[
\alpha_S = \frac{\int_{V_S}^{V_D} Q'_n (v) (Q'_b (v) - Q'_b (V_S)) \, dv}{\int_{V_S}^{V_D} Q'_n (v) (Q'_g (v) - Q'_g (V_S)) \, dv}.
\]

(5.28)

The integrals in (5.28) can be approximated by a series solution using the trapezoidal rule for the integration. A numerical comparison between the exact solution and the series solution shows that by taking only the first term in this series we can obtain an approximation that is both simple and accurate:
Fig. 5.6 Functional dependencies of $\alpha_S$ and $\alpha_D$ on the voltages of the terminals. The dashed line represents our approximation. ($N_{\Lambda\Lambda} = 1 \times 10^{15} \text{ cm}^{-3}$, $\chi_0 = 2000 \text{Å}$)
\[ \alpha_S = \frac{Q_b(v_D) - Q_b(v_S)}{Q_g(v_D) - Q_g(v_S)} \] (5.29)

If we repeat the procedure for \( \alpha_D \) we find that

\[ \alpha_D = \alpha_S = \alpha \] (5.30)

We can express \( \alpha_S \) and \( \alpha_D \) in terms of the external voltages by employing the expressions for \( Q_g \) and \( Q_b \) given in Section 4.2. If we use the strong inversion approximation \( \psi_S = v + 2\phi_F \) in (4.5) and (4.8) these expressions become

\[ Q_g = C_o(v'_G - v - 2\phi_F) \] (5.31)
\[ Q_b = -KC_o(v + 2\phi_F - v_B)^{1/2} \] (5.32)

Substituting in (5.30) we obtain

\[ \alpha = \frac{K}{v_{DS}} \left[ (v_{DS} + 2\phi_F - v_{BS})^{1/2} - (2\phi_F - v_{BS})^{1/2} \right] \] (5.33)

where \( K = (2q\epsilon_S N_{AA}/C_O^2)^{1/2} \). In saturation \( v_{DS} \) must be substituted by the saturation voltage: \( v_{DSS} = v'_G - 2\phi_F + K^2/2 - K(v'_G - v_B + K^2/4)^{1/2} \).

Figure 5.6 shows the approximation in (5.33) as a function of voltage. In Fig. 5.7 we compare the functional dependencies for \( C_{SB} \) and \( C_{DB} \) obtained from Section 5.3.1 with the one obtained by using (5.25) and (5.33). In both figures, good agreement is shown between the approximation and the more detailed expressions.
Fig. 5.7 $C_{SB}$ and $C_{DB}$ calculated using the detailed expressions of section 5.3.1 compared with $C_{SB}$ and $C_{DB}$ calculated using the approximation involving $\alpha_s$ and $\alpha_D$. 

$N_{AA} = 1 \times 10^{15} \text{cm}^{-3}$

$\chi_0 = 2000 \text{Å}$

$V_D = 1 \text{v}$

$V_B = 0 \text{v}$
5.3.4 Engineering Importance of the Intrinsic Substrate Capacitances $C_{SB}$ and $C_{DB}$

From a simplified theory of MOSFET operation [20] the approximated value of the gate capacitances $C_{SG}$ and $C_{DG}$ are

$$C_{SG} = \begin{cases} 2/3 \ Z L C_o & \text{in saturation} \\ 1/2 \ Z L C_o & \text{in nonsaturation} \end{cases}$$

(5.34)

$$C_{DG} = \begin{cases} 0 & \text{in saturation} \\ 1/2 \ Z L C_o & \text{in nonsaturation} \end{cases}$$

(5.35)

where $Z L C_o$ is the total oxide capacitance. The simplified theory does not include the intrinsic substrate capacitances. However, an excellent estimation of their value can be obtained by using the approximation discussed in Section 5.3.3.

From (5.25) we obtain

$$C_{SB} = \begin{cases} 2\alpha/3 \ Z L C_o & \\ \alpha/2 \ Z L C_o \end{cases}$$

(5.36)

$$C_{DB} = \begin{cases} 0 & \\ \alpha/2 \ Z L C_o \end{cases}$$

(5.37)

with $\alpha$ defined in (5.33) as

$$\alpha = \frac{1}{C_o} \left(2q\varepsilon_S N_{AA}\right)^{1/2} \left[\frac{(v_{DS} + 2\phi_F - v_{BS})^{1/2} - (2\phi_F - v_{BS})^{1/2}}{v_{DS}}\right]$$

(5.38)

To study the importance of the substrate capacitances we will consider a particular device with $N_{AA} = 1 \times 10^{15}$ cm$^{-3}$ and
$x_o = 2000\text{A}$. The functional dependency of $\alpha$ in the terminal voltages for this case is shown in Fig. 5.6. Notice that the value of $\alpha$ in this example is about one half which implies values of $C_{SB}$ and $C_{DB}$ of about one half the magnitude of $C_{SD}$ and $C_{DG}$. However, because the value of $\alpha$ is directly related to the square root of the substrate doping $N_{AA}$ and to the oxide thickness (equation (5.38)), the relative value of $C_{SB}$ and $C_{DB}$ will also depend on these parameters. For instance, if the doping concentration is changed from $1\times10^{15}$ to $4\times10^{15}$ in our example the functional dependence of $\alpha$ shown in Fig. 5.7 would be shifted upwards by a factor of two. This would result in values of $C_{SB}$ and $C_{DB}$ of about the same magnitude of $C_{SG}$ and $C_{DG}$. For even larger values of $N_{AA}$ or $x_o$ the values of $C_{SB}$ and $C_{DB}$ would exceed the values of $C_{SG}$ and $C_{DG}$. Thus, for the doping concentrations considered in this example, the values of $C_{SB}$ and $C_{DB}$ exceed the values of $C_{SG}$ and $C_{DG}$.

Notice from equations (5.36) through (5.37) that $C_{SB}$ and $C_{DB}$ are proportional to $\alpha/C_o$, but, because $\alpha$ is directly related to $1/C_o$, $C_{SB}$ and $C_{DB}$ are independent of $C_o$. This contrasts with $C_{SG}$ and $C_{DG}$ which directly depend on $C_o$.

5.4 Transcapacitors

5.4.1 Expressions for the Transcapacitors

In Chapter III the transcapacitances were defined (equations (3.20)-(3.24)) as:
\[
\begin{align*}
\frac{d_{SG}}{d_{SB}} &= \lambda \frac{\partial Q_N}{\partial v_G} + C_{SG} \\
\frac{d_{DG}}{d_{DB}} &= (1-\lambda) \frac{\partial Q_N}{\partial v_G} + C_{DG} \\
\frac{d_{GB}}{d_{SG}} &= d_{SG} + d_{DG}
\end{align*}
\]

(5.39)

To transform these definitions into functional dependencies of the voltages we require only an expression for \( \frac{\partial Q_N}{\partial v_G} \) in terms of the terminal voltages. Taking derivatives with respect to \( v_G \) in (4.14) which defines \( Q_N \), we obtain

\[
\frac{\partial Q_N'}{\partial v_G} = -\frac{1}{\Im_D} \left( Q_N' \frac{\partial \Im_D}{\partial v_G} + \frac{\partial}{\partial v_G} \int_{v_S}^{v_D} Q_n'^2 dv \right). 
\]

(5.40)

Here, \( \frac{\partial \Im_D}{\partial v_G} \) was previously evaluated by taking partial derivatives in (4.27) and is given by (5.16). The second term in (5.40) can be obtained in a form similar to the one employed to calculate the total charges in Section 4.3.2.

We obtain

\[
\frac{\partial}{\partial v_G} \int_{v_S}^{v_D} Q_n'^2 dv = 2\Im_D + \frac{\partial F_N(v_S, v_1)}{S_S} + \\
\frac{DF_N(v_1, v_2)}{S_M} + \frac{DF_N(v_2, v_D)}{S_W}
\]

(5.41)

where we have defined

\[
DF_N(v_a', v_b) = \left[ \frac{Q_n'^2}{1 + \frac{K^2}{2Q_g} \left[ 1 + e^{\beta(v_S - v - 2\phi_F)} \right]} \right]^{v_b}_{v_a}
\]

(5.42)
The complete expression for \( \frac{\partial Q_N}{\partial V_G} \) is then

\[
\frac{\partial Q_N}{\partial V_G} = \frac{1}{I_D} \left[ Q' \left( \frac{DF_I(v_s', v_1)}{S_S} + \frac{DF_I(v_1', v_2)}{S_M} + \frac{DF_I(v_2', v_D)}{S_N} \right) + 2I_D + \frac{DF_N(v_s', v_1)}{S_S} + \frac{DF_N(v_1', v_2)}{S_M} + \frac{DF_N(v_2', v_D)}{S_W} \right].
\] (5.43)

Here, the function \( DF_I(v_a', v_b') \) was previously defined by (5.16).

The functional dependencies for the transcapacitances are illustrated in Fig. 5.8 for a specific device. In this figure notice that the value of the transcapacitors is about one order of magnitude smaller than the value of the capacitors (Fig. 5.1) except when the device is operating in weak inversion (between cut-off and saturation). In contrast with the current flowing in the capacitors, the current flowing in the transcapacitors is not determined by the voltage across their terminals. Thus, the relative value of the transcapacitors with respect to the capacitors is not enough to assess their importance in the equivalent-circuit. In the next section we will consider the engineering importance of the transcapacitors in the overall equivalent-circuit.

### 5.4.2 Engineering Importance of the Transcapacitance Elements

The derivation of the equivalent-circuit for the MOSFET in Chapter III demonstrated the need for circuit elements in addition to capacitors for the network representation of charging currents. This need arises from the basic asymmetry of the physical structure of the device. The
Fig. 5.8 Calculated values for the transcapacitances.
additional circuit elements can be regarded as correction terms to an all-capacitive network representation. Thus, we can study the importance of the transcapacitance elements in the overall network representation by considering, at each terminal, the ratio of the charging currents flowing through transcapacitors to the charging currents flowing through capacitors.

From an analysis of the equivalent-circuit of Fig. 3.3 and using the properties of the indefinite admittance matrix describing this network, we obtain the following expressions for the relative importance of the transcapacitive currents:

\[
\frac{(i_S)^T}{(i_S)^C} = - \frac{d_{SG}(\dot{v}_G - \dot{v}_B)}{C_{SG} \dot{v}_G + C_{SB} \dot{v}_B - (C_{SG} + C_{SB}) \dot{v}_S} \quad (5.44)
\]

\[
\frac{(i_D)^T}{(i_D)^C} = - \frac{d_{DG}(\dot{v}_G - \dot{v}_B)}{C_{DG} \dot{v}_G + C_{DB} \dot{v}_B - (C_{DG} + C_{DB}) \dot{v}_D} \quad (5.45)
\]

\[
\frac{(i_G)^T}{(i_G)^C} = - \frac{d_{GB}(\dot{v}_G - \dot{v}_B)}{C_{SG} \dot{v}_S + C_{DG} \dot{v}_D + C_{GB} \dot{v}_B - (C_{SG} + C_{DG} + C_{GB}) \dot{v}_G} \quad (5.46)
\]

\[
\frac{(i_B)^T}{(i_B)^C} = 0 . \quad (5.47)
\]

This last expression indicates that, if we use an all-capacitive circuit model, with capacitors defined as in Table 3.1, the error involved in calculating the substrate current is zero. This results from the particular manner in which the
basic building block of Chapter II (Fig. 2.2) was used to construct the complete equivalent-circuit for the MOSFET.

In equations (5.44) through (5.47) notice that the relative importance of the transcapacitors depends on the particular circuit environment in which the device is used. For example, if the particular circuit environment is such that $\dot{v}_G = \dot{v}_B$ the contribution of the transcapacitors to the total charging currents at any terminal would be zero. A contrasting example is a circuit environment such that $\dot{v}_G \neq \dot{v}_B$ and

$$
\dot{v}_D = \frac{C_{DG} \dot{v}_G + C_{DB} \dot{v}_B}{C_{DG} + C_{DB}} = \frac{\dot{v}_G}{1 + \alpha} + \frac{\alpha}{1 + \alpha} \dot{v}_B \quad (5.48)
$$

where $\alpha = C_{DB}/C_{DG}$. From equation (5.45) one can see that in this case the charging currents flowing through the transcapacitors would be the main contribution to the total charging current at the drain.

To illustrate the practical importance of the transcapacitors, consider the input device of the inverter circuit shown in Fig. 5.9. In this figure there are also shown qualitative sketches of the waveforms of a particular applied excitation and the response of the circuit. If we apply (5.44)-(5.47),

$$
\frac{(i_S)_T}{(i_S)_C} = - \frac{d_{SG}}{C_{SG}} \quad (5.49)
$$
Fig. 5.9 Example to illustrate the engineering importance of the transcapacitances.
\[
\frac{(i_D)_T}{(i_D)_C} = - \frac{d_{DG}}{C_{DG} \left( \frac{V_{DD}}{V_I} + 1 \right) + C_{DB} \frac{V_{DD}}{V_I}} \quad (5.50)
\]

\[
\frac{(i_G)_T}{(i_G)_C} = \frac{d_{GB}}{C_{DG} \frac{V_{DD}}{V_I} + (C_{SG} + C_{DG} + C_{GB})} \quad (5.51)
\]

\[
\frac{(i_B)_T}{(i_B)_C} = 0. \quad (5.52)
\]

An inspection of Figs. 5.1 and 5.8 shows that the maximum values for these ratios will occur in the vicinity of the transition between cut-off and saturation and between saturation and nonsaturation. As an example consider \( V_{DD}/V_I = 1 \).

In the vicinity of the transition between cut-off and saturation \( (V'_G \approx 1.5v) \), we obtain

\[
\frac{(i_S)_T}{(i_S)_C} \approx -0.36 \quad \frac{(i_G)_T}{(i_G)_C} \approx -0.20, \quad (5.53)
\]

\[
\frac{(i_D)_T}{(i_D)_C} = \frac{(i_B)_T}{(i_B)_C} = 0, \quad (5.53)
\]

while, in the vicinity of the transition between saturation and nonsaturation \( (V'_G \approx 3v) \), we obtain

\[
\frac{(i_S)_T}{(i_S)_C} \approx 0.06 \quad \frac{(i_D)_T}{(i_D)_C} \approx 0.04 \quad (5.54)
\]

\[
\frac{(i_G)_T}{(i_G)_C} \approx 0.05, \quad \frac{(i_B)_T}{(i_B)_C} = 0. \quad (5.54)
\]
These numbers represent approximately the maximum percentage of error involved when evaluating the charging currents at each terminal using an all-capacitive network. Thus for this particular example the neglect of the transcapacitances would result in error of about 36% in calculating the charging current flowing in the source and errors of about 20% and 4% in calculating the charging currents flowing in the gate and in the drain. Although the example used for illustrating purposes is simple, it indicates the potential engineering importance of the transcapacitors for calculating rise times and other related behavior. The use of the transcapacitors adds complexity to the programs in computer aided design. Detailed studies involving computer simulations in large-scale integration are necessary to assess the practical engineering importance of these elements. This, however, lies beyond the scope of this work.

5.4.3. Transcapacitances in a Three-Terminal Equivalent-Circuit

Among the circuit environments in which the MOSFET is used, one in which the source and the substrate are connected together as shown in Fig. 5.10(a) is often found in practical applications. In this case we will show that transcapacitive effects are not necessary to model charging currents in the MOSFET and therefore the equivalent-circuit reduces to the result [11] for a three-terminal device shown in Fig. 5.10(b).

In Section 3.2.2 we proved that charging currents in the MOSFET can be represented by an indefinite admittance
Fig. 5.10  Reduction of the equivalent-circuit for the MOSFET when the device is used in a three-terminal application.
matrix (equation (3.7)). The interconnection of source and substrate imposes two constraints in the circuit: a constraint on the currents \(i_S + i_B\), and a constraint on the voltages \(\dot{v}_S = \dot{v}_B\). Consider these constraints applied to the indefinite admittance matrix in equation (3.7). The constraint on the currents requires adding the first and fourth rows. The constraint on the voltages requires adding the first and the fourth columns. The matrix resulting is then,

\[
\begin{align*}
\begin{bmatrix}
i_S + i_B \\
i_D \\
i_G 
\end{bmatrix} &= \begin{bmatrix}
\frac{\partial q_S}{\partial v_S} + \frac{\partial q_B}{\partial v_S} + \frac{\partial q_S}{\partial v_B} & \frac{\partial q_S}{\partial v_D} & \frac{\partial q_S}{\partial v_G} \\
\frac{\partial q_D}{\partial v_S} + \frac{\partial q_D}{\partial v_B} & \frac{\partial q_D}{\partial v_D} & \frac{\partial q_D}{\partial v_G} \\
\frac{\partial q_G}{\partial v_S} + \frac{\partial q_G}{\partial v_B} & \frac{\partial q_G}{\partial v_D} & \frac{\partial q_G}{\partial v_G} 
\end{bmatrix} \begin{bmatrix}
\cdot v_S \\
\cdot v_D \\
\cdot v_G 
\end{bmatrix} + \begin{bmatrix}
i_{SD} \\
-i_{SD} \\
0 
\end{bmatrix}
\end{align*}
\]

This matrix still satisfies the properties of the indefinite admittance matrix. Furthermore, if we neglect capacitive effects between source and drain \(\frac{\partial q_S}{\partial v_D} = \frac{\partial q_D}{\partial v_S} + \frac{\partial q_D}{\partial v_B} = 0\), one can see that this matrix is symmetric. That is,

\[
\begin{align*}
\frac{\partial q_S}{\partial v_G} &= \frac{\partial q_G}{\partial v_S} + \frac{\partial q_G}{\partial v_B} \\
\frac{\partial q_D}{\partial v_B} &= \frac{\partial q_G}{\partial v_D} 
\end{align*}
\]
Therefore, the network representation of this symmetric matrix consists only of capacitors [16]. One of these capacitors,

\[ C_{SG} = -\left( \frac{\partial q_G}{\partial V_S} + \frac{\partial q_G}{\partial V_B} \right), \] (5.57)

is connected between source and gate, and the other,

\[ C_{DG} = -\frac{\partial q_G}{\partial V_D}, \] (5.58)

is connected between drain and gate, as shown in Fig. 5.10(b).

5.5 Conclusions

In this chapter we derived the functional dependencies for the elements in the equivalent-circuit developed in Chapter III: a current-source representing charge transport, and capacitors and transcapacitors representing mobile charge accumulation. The main achievement of the chapter is the derivation of functional dependencies for these circuit elements that are valid and continuous in the entire range of operation of the device. This contrast with similar previous work [11,20,28] which applies only for the strong inversion operation of a three-terminal MOSFET.

Another result of this chapter is an approximation of engineering importance that permits the calculation of substrate capacitances directly from knowledge of the gate capacitances. The approximation is both simple and accurate.
The importance of the transcapacitances was discussed from an engineering point of view. If the transcapacitances are neglected, they can be a potential source of error in certain circuit applications.

Work is needed to assess the practical importance of the new equivalent-circuit model in computer-aided MOSFET circuit analysis. As a step in this direction Appendix C gives a computer subprogram that uses the results of this chapter to calculate the circuit elements as functions of the external voltages.
CHAPTER VI
SCOPE AND FUTURE WORK

In this study we have proposed a methodology for developing models for the large-signal transient response of n-terminal electronic devices, and have applied it to a particular device, the four-terminal MOSFET. The principal contributions achieved are summarized in this chapter together with recommendations for future research.

In Chapter II we presented a systematic modeling approach that applies to n-terminal electronic devices obeying the principle of charge control [2,3]. It is based on an extension of the indefinite admittance matrix from network theory. The approach is especially useful when modeling devices with three or more terminals. Its power is emphasized in Chapter III where we applied it to the four-terminal MOSFET.

The methodology considers the physical mechanisms commonly involved in the operation of electronic devices: the transport, net recombination, and accumulation of mobile charge. If analogous mechanisms occur in other systems, for example, in the chemical, physical, societal, and biological sciences [38], then to these systems the methodology could be also applied. Such extensions are beyond the objectives of the present work.
Chapter III presents a new large-signal equivalent-circuit model for the transient response of the MOSFET (Fig. 3.3) that differs in a fundamental sense from those developed by the more intuitive schemes previously used. It includes two effects not fully included in previous work: (1) the apportionment of charge amongst the four terminals, and (2) the properties resulting from the asymmetry of the physical structure of the MOSFET.

The equivalent-circuit model applies generally, subject only to the validity of the quasi-static approximation. In principle it can include two- and three-dimensional effects such as those occurring in short-channel MOSFETs. Static descriptions for these effects have been already advanced [8] and more can be expected in the future. Work toward applying such descriptions is in order.

Chapter IV describes the development of a dc static characterization for the behavior of the MOSFET. Its results offer advantages to computer-aided circuit design over previous static characterizations. The dc steady-state description covers the entire range of operation: weak to strong inversion, cut-off to saturation. It includes the effect of the substrate and the substrate terminal. It contains compact expressions for current and charges with continuous derivatives with respect to the terminal voltages.

The static characterization is used in Chapter V to obtain the functional dependencies for the equivalent-circuit elements. The results apply to the entire range of operation
of the MOSFET. Good agreement with some experimental results is shown; a more comprehensive experimental confirmation is in order, in particular, with respect to the capacitors and transcapacitors. Research is also necessary to investigate the practical value of the new equivalent-circuit in computer-aided circuit design.

A central approximation on which the results of our study depend is the quasi-static approximation (QSA). This approximation underlies not only our work but nearly all the models in common use for either bipolar transistors or MOSFETs in computer-aided circuit analysis. It is widely used because it yields compact network representations. Recent work by Frazer and Lindholm [39] has indicated, however, that under certain circumstances the QSA, for some of the MOSFET models now widely used in computer-aided design, fails to be consistent with itself. This lack of self-consistency can produce errors in computing system parameters such as the turn-off propagation delay time.

By including in our new model a self-consistency test proposed by Lindholm and Frazer [40], one can detect under which particular applications the QSA is violated. The violations are related to the transit time across various regions of the device, and therefore to the thickness of the device region under consideration. Thus an approach for eliminating the inadequacies of the model in these situations is to divide the intrinsic region into subregions and model each of them using the approach of Chapter II [41]. The
complete equivalent circuit is then the series connection of the models of each region. The size and number of these regions could be determined by the magnitude of the violation of the QSA. More research is needed to assess the suitability of this approach.

Another approach for eliminating the inadequacies of the QSA results from abandoning this approximation in the modeling process. Such an approach, based on approximating Shockley's six basic differential equations [42] by finite differences, has been proposed by Sah [43] but has never been applied to the MOSFET. Future work could include applying this approach to the MOSFET.
APPENDIX A

PROPERTIES OF QUASI-FERMI POTENTIALS

The objective of this appendix is to explore some properties of the quasi-Fermi potentials in semiconductor materials. These properties are used in the analysis of MOSFET behavior presented in Chapter IV.

Consider a long piece of semiconductor material and apply a voltage $V_a$ across its terminals as shown in Fig. A-1.

This system must satisfy Faraday's law expressed mathematically by

$$\oint \mathbf{E} \cdot d\mathbf{l} = 0$$  \hspace{1cm} (A-1)

or

$$\left\{ \begin{array}{c} \oint \mathbf{E} \cdot d\mathbf{l} \\
\text{external circuit} \\
\text{semiconductor} \end{array} \right. = 0 .$$  \hspace{1cm} (A-2)
In the semiconductor, \( E \approx E_x = -dV_I/dx \), where \( V_I \) is the electrostatic potential. Thus, Faraday's law requires

\[
V_a = - \int_0^L E \, dx = \int_0^L \frac{dV_I}{dx} \, dx = V_I(L) - V_I(0) .
\]

(A-3)

When studying semiconductor devices a very important tool of analysis results from defining electron and hole quasi-Fermi potentials \( V_N \) and \( V_P \) as follows [42]:

\[
\begin{align*}
V_N &= V_I - \frac{kT}{q} \ln \frac{N}{n_i} \\
V_P &= V_I - \frac{kT}{q} \ln \frac{P}{n_i}
\end{align*}
\]

(A-4)

Here, \( N \) and \( P \) are the electron and hole concentrations.

We wish to find the functional relation between the applied external voltage and the electron and hole quasi-Fermi potentials. From (A-4)

\[
V_I(x) = V_N(x) + \frac{kT}{q} \ln \frac{N(x)}{n_i}
\]

(A-5)

Substituting this equation in (A-3) we obtain for electrons

\[
V_a = V_N(L) - V_N(0) + \frac{kT}{q} \ln \frac{N(L)}{N(0)} .
\]

(A-6)

If we impose the constraint that the contacts between the semiconductor and the external circuit are ohmic, then at the boundaries \( x=0 \) and \( x=L \) the electron concentrations must have their equilibrium values.
\[
\begin{align*}
N(L) &= N_E(L) \\ N(0) &= N_E(0)
\end{align*}
\]  \hspace{1cm} (A-7)

Because \((A-6)\) must hold in equilibrium, when \(V_a = 0\) and the quasi-Fermi levels for electron and holes coincide with the Fermi level which is position independent, we obtain
\[
\frac{kT}{q} \ln \frac{N_E(L)}{N_E(0)} = 0
\]  \hspace{1cm} (A-8)

which implies
\[
N_E(L) = N_E(0) .
\]  \hspace{1cm} (A-9)

That is, the equilibrium concentration of electrons is equal at both ohmic contacts. Notice that this interesting result is not restricted to homogeneous material. For holes, one can show that a similar result is obtained.

If we now substitute \((A-7)\) and \((A-9)\) in \((A-6)\) we obtain
\[
V_a = V_N(L) - V_N(0)
\]  \hspace{1cm} (A-10)

Analogously, for holes
\[
V_a = V_P(L) - V_P(0) .
\]  \hspace{1cm} (A-11)

The relationships in \((A-10)\) and \((A-11)\) are very useful to establish boundary conditions when analyzing the physics of a semiconductor device as will be shown for the particular case of the MOSFET.

In the analysis presented in Chapter IV of the n-channel MOSFET we defined a "channel voltage" related to the quasi-
Fermi potentials by

\[ V = V_N - V_P \]  \hspace{1cm} (A-12)

Because there is no significant hole current in this device, \( V_P \) can be considered constant. If we denote the position of the ohmic contacts in the source and drain islands by \( 0' \) and \( L' \), then (A-10) requires

\[ V_D - V_S = V_N(L') - V_N(0') = V(L') - V(0') \]  \hspace{1cm} (A-13)

or

\[ V(0') = V_S + c \]

\[ V(L') = V_D + c \]  \hspace{1cm} (A-14)

where \( c \) is an arbitrary constant which for convenience we will set equal to zero. The source and drain islands are heavily doped and the quasi-Fermi levels there are independent of position. In these regions, therefore, \( V \) is also independent of position and at the boundaries of the channel, \( y=0 \) and \( y=L \) (Fig. 3.2), we have

\[ V(0) = V_S \]

\[ V(L) = V_D \]  \hspace{1cm} (A-15)

These are the boundary conditions for \( V \) used in Chapter IV.
APPENDIX B

APPROXIMATED EXPRESSION FOR THE DIFFUSION/DRIFT RATIO IN THE MOSFET

In this appendix we will discuss an analytical justification to the expression for the diffusion/drift ratio given in equation (4.20):

\[
\frac{D_n}{\mu_n N E_Y} \frac{dN/dy}{dN/dV} = \frac{1 - \psi_S/dV}{d\psi_S/dV} \approx \frac{N_{AA} - P_S}{N_S} \approx \frac{N_{AA}}{N_S}.
\]  

We will begin by proving the first part of this equation.

The ratio of the diffusion to the drift component along the channel in a MOSFET is

\[
\frac{J_{\text{DIFF}}}{J_{\text{DRIFT}}} = \frac{D_n}{\mu_n N E_Y} \frac{dN/dy}{dN/dV} \quad (B-1)
\]

We can express the electron concentration \(N\) in terms of the electrostatic potential \(V_I\) and the quasi-Fermi potential for electrons \(V_N\) by using

\[
N = n_i e^\beta (V_I - V_N).
\]  

Inspection of the energy band diagram in Fig. 4.1 shows that \(V_I - V_P = \psi - \phi_F\). Because \(V = V_N - V_P\), it follows that \(V_I - V_N = \psi - V - \phi_F\). Thus \(N\) can be rewritten as
\[
N = n_1 e^{\beta(\psi-V-\phi_F)}.
\]  
(B-3)

Taking derivatives with respect to \(y\) in this equation we obtain,

\[
\frac{dN}{dy} = \frac{N}{q/kT} \left[ \frac{d\psi}{dy} - \frac{dV}{dy} \right].
\]  
(B-4)

From the gradual channel approximation [19],

\[
\frac{dV_I}{dy} = \frac{d\psi(x,y)}{dy} = \frac{d\psi(0,y)}{dy} = \frac{d\psi_S}{dy}.
\]  
(B-5)

Substituting (B-5) and (B-4) in (B-1), using Einstein's relation \((D/\mu = kT/q)\), and simplifying common terms we finally obtain

\[
\frac{D_n N}{\psi_n N} \frac{dN/dy}{E_y} = \frac{dV/dy - d\psi_S/dy}{d\psi_S/dy} = \frac{1 - d\psi_S/dV}{d\psi_S/dV},
\]  
(B-6)

which proves the first part of (4.20).

To prove the second part of (4.20) we will use the relationship between the surface potential \(\psi_S\) and the applied voltages obtained from Gauss' law in Section 4.2.3:

\[
V'_G - \psi_S = K \left[ \psi_S - V_B + \frac{kT}{q} \left( e^{\beta(\psi_S-V-2\phi_F)} - 1 \right) \right]^{1/2}
\]  
(4.18)
Taking partial derivatives with respect to $y$ in both sides of (4.18) we obtain

$$\frac{d\psi_S}{dV} = \frac{e^{\beta(\psi_S-V-2\phi_F)}}{\frac{2Q_g}{K^2} + 1 + e^{\beta(\psi_S-V-2\phi_F)}} \quad (B-7)$$

Rearranging terms,

$$\frac{1 - \frac{d\psi_S}{dV} \frac{d\psi_S}{dV}}{\frac{d\psi_S}{dV}} = \frac{2Q_g/K^2 + 1}{e^{\beta(\psi_S-V-2\phi_F)}} \quad (B-8)$$

The exponential term in (B-9) is related, as we shall see, to the electron concentration at the surface. From (B-4),

$$N_S = n_1 e^{\beta(\psi_S-V-\phi_F)} \quad (B-9)$$

because $N_{AA}/n_1 = e^{\beta\phi_F}$, the exponential term in (B-9) is therefore equal to $N_S/N_{AA}$. We can then write

$$\frac{1 - \frac{d\psi_S}{dV} \frac{d\psi_S}{dV}}{\frac{d\psi_S}{dV}} = (2Q_g/K^2+1) \frac{N_{AA}}{N_S} \quad (B-10)$$

The term $(2Q_g/K^2+1) = (2(V_g-\psi_S)/K^2+1)$ varies along the channel within the same order of magnitude (Fig. 4.3) while $N_S$ changes by several orders of magnitude. Therefore, the behavior of the diffusion/drift ratio is strongly dominated by changes in $N_S$ and thus, for purposes of illustrating the main features of this ratio, we can take
\[
\frac{1 - \frac{d\psi_S}{dV}}{\frac{d\psi_S}{dV}} \simeq \frac{N_{AA}}{N_S}.
\] (B-11)

This justifies the second part of (4.20).
APPENDIX C

COMPUTER SUBPROGRAM TO CALCULATE THE VALUE OF THE ELEMENTS IN THE EQUIVALENT-CIRCUIT

The subroutine that we list in this appendix employs the results of Chapter V. The input to the subroutine is the voltage at the terminals: $V_S', V_D', V_G', V_B$. The output is the value of the elements in the equivalent circuit: the current source $I_D$, the capacitors $C_{SG}, C_{SB}, C_{DG}, C_{DB}, C_{GB}$ and the transcapacitors denoted by $T_{SG}, T_{SB}, T_{DG}, T_{DB}$ and $T_{GB}$. The parameters $BK = K$ (defined in equation (4.9)), $TK = kT/q$, and $TF1F = 2kT/q\ln N_{AA}/n_i$ must be entered in a COMMON block.
SUBROUTINE EQCKT(VS,VD,VG,VB,IE, *> +ID,CSG,CSB,CDG,CDB,CGB,TSG,TSB,TDG,TDB,TGB) COMMON BK,TK,TFIF REAL LAMBDA,LIM,LIM1,LIM2,IDS,IDM,IDW,ID

* * *

CHARGE PER UNIT AREA
QG(U)=VG-U
QB(U)=-BK*SQRT(U-VB-TK)
QN(U)=-(QG(U)+QB(U))

DEFINE FUNCTIONS USED IN SUBROUTINE
G(U)=-QG(U)**2/2.
B(U)=2*QB(U)**3/(3.*BK**2)
GG(U)=-QB(U)**3/3.
BB(U)=QB(U)**4/(2*BK**2)
GGG(U)=-QG(U)**3/3.
BGG(U)= QB(U)**4/(2*BK**2)
DSG(U,V)=1./(1+.5*BK**2*(1+EXP((U-V-TFIF)/TK))/QG(U))
DSI(ABK)=TK*ALOG((2./ABK)*(TK/ABK+SQR(T(VG-VB+BK**2/4.+TK*(TK/ABK**2-1.))))

VX(LIM)=VG-TFIF-LIM+BK**2/2.-
+ BK*SQRT(VG-VB+BK**2/4.+TK*(EXP(LIM/TK)-1))

* * *

MINIMUM VALUE OF GATE VOLTAGE FOR
ONSET OF INVERSION N=N1;PSIS-VB=FIF
IE=2
IF (VG-VB .LT. TFIF/2.+BK*SQRT(TFIF/2.)) RETURN

ESTABLISH LIMITS FOR WEAK MODERATE STRONG INVERSION
LIM1=DSI(BK/4.)
LIM2=DSI(4.*BK)
V1=VX(LIM1)
V2=VX(LIM2)
VDSS=VX(0.0)

* * *

COMPUTE SURFACE POTENTIAL U AT THE LIMITS AND
APPROXIMATE THE SLOPES IN THE THREE REGIONS
V=VS
ESTIMATION FOR US
TSI=V2+TFIF
FG=(VG**2-VG-BK**2*(TFIF-VB))/(BK**2*TK)
IF (FG .GT. 0.0) TSI=TFIF+TK*ALOG(FG)
ASSIGN 1 TO KK
GO TO 51
1
US=SION
IF (VD .GT. VS) GO TO 5
ID=0.0
GO TO 35
5
V=VD
ESTIMATION FOR UD
TSI=V2+TFIF
IF (V2 .GT. VD) TSI=VD+US
ASSIGN 10 TO KK
GO TO 51
10
UD=SION
SW=0.
SM=0.5
SS=1.
IF (V1 .GT. VS) GO TO 12
V1=VS
U1=US
GO TO 18
12 IF(V1 .LT. VD) GO TO 14
V1=VD
U1=UD
GO TO 16
14 V=V1
ASSIGN 15 TO KK
GO TO 50
15 U1=SIS
16 SS=(U1-US)/(V1-VS)
18 IF (V2 .GT. VS) GO TO 20
V2=VS
U2=US
GO TO 26
20 IF (V2 .LT. VD) GO TO 22
V2=VD
U2=UD
GO TO 28
22 V=V2
ASSIGN 25 TO KK
GO TO 50
25 U2=SIS
26 SW=(UD-U2)/(VD-V2)
28 IF (V2 .NE. V1) SM=(U2-U1)/(V2-V1)
*
* DRAIN CURRENT
IDS=(G(U1)+B(U1)-G(US)-B(US))/SS
IDM=(G(U2)+B(U2)-G(U1)-B(U1))/SM
IDW=TK*(QN(UD)-QN(U2))/(1-SW)
ID=IDS+IDM+IDW
IF (ID .GT. 0.0) GO TO 40
35 IE=1
DQGSI=BK**2*(1+EXP(US-VS-TFIF)/TK)-EXP((VB-US)/TK)/(2*QG(US))
CGB=DQGSI/(1.+DQGSI)
* ALL THE ELEMENTS IN EQCKT=0 BUT ID AND CGB
RETURN
40 IE=0
*
* TOTAL CHARGE COMPONENTS
SGNS=(GG(U1)+GB(U1)-GG(US)-GB(US))/SS
SGNM=(GG(U2)+GB(U2)-GG(U1)-GB(U1))/SM
SGNW=TK* (G(UD)-B(UD)-QG(UD)*QB(UD)
+ G(U2)+B(U2)+QB(U2)/QG(U2))/(1-SW)
QGG=(SGNS+SGNM+SGNW)/ID
SBNS=(BB(U1)+GB(U1)-BB(US)-GB(US))/SS
SBNM=(BB(U2)+GB(U2)-BB(U1)-GB(U1))/SM
SBNW=TK* (B(UD)-QG(UD)**2/2.-B(U2)+QB(U2)**2/2.)/(1-SW)
QBB=(SBNS+SBNM+SBNW)/ID
QNN=- (QGG+QBB)
IF (QGG .GT. QG(US) .OR. QBB .GT. QB(US)) GO TO 35
TRANSCONDUCTANCES
GS=QN(US)
GD=-QN(UD)
GMS=0 (QG(U1)+QN(U1)*DSG(U1,VL)-QG(US)-QN(US)*DSG(US,VS))/SS
GMM=(QG(U2)+QN(U2)*DSG(U2,V2)-QG(U1)-QN(U1)*DSG(U1,V1))/SM
GMW=TK* ((1-.5*BK**2/QB(UD))*DSG(UD,VD)
  +  -(1-.5*BK**2/QB(U2))*DSG(U2,V2))/(1-SW)
GM=GMS+GMM+GMW
GMB=-(GS+GD+GM)

INTEGRAL OF (D/DVG)QB*QN DV
DBNGS=-(B(U1)-QN(U1)*QB(U1)*DSG(U1,VL)
  +  -B(US)+QN(US)*QB(US)*DSG(US,VS))/SS
DBNGM=-(B(U2)-QN(U2)*QB(U2)*DSG(U2,V2)
  +  -B(U1)+QN(U1)*QB(U1)*DSG(U1,V1))/SM
DBNGW=-TK* ((QB(UD)-BK**2/2.)*DSG(UD,VD)
  +  -(QB(U2)-BK**2/2.)*DSG(U2,V2))/(1-SW)
DBNG=DBNGS+DBNGM+DBNGW

VARIATION OF TOTAL INVERSION CHARGE QNN WITH VG
DNNGS=2*IDS+(QN(U1)**2*DSG(U1,VL)
  -ON(US)**2*DSG(US,VS))/SS
DNNGM=2*IDM+(QN(U2)**2*DSG(U2,V2)
  -ON(U1)**2*DSG(U1,V1))/SM
DNNGW= TK* (QN(U1)*((1-(1-BK**2/(2*QB(UD))))*DSG(UD,VD)
  +  -QN(U2)*((1-(1-BK**2/(2*QB(U2))))*DSG(U2,V2))/(1-SW)
DNNG=DNNGS+DNNGM+DNNGW
DQNG=-(QNN*GM+DNNG)/ID

CAPACITANCES
CSG=QN(US)*(QG-QG(US))/ID
CSB=QN(US)*(QB-QB(US))/ID
CDG=-QN(UD)*(QG-QG(UD))/ID
CDB=-QN(UD)*(QB-QB(UD))/ID
CGB=(QBB*GM+DBNG)/ID

APPORTIONING FUNCTION LAMBDA
LAMBDA=1./(1+(CDG+CDB)/(CSG+CSB))

TRANSCAPACITANCES
TSG=LAMBDA*DQNG+CSG
TDG=(1-LAMBDA)*DQNG+CDG
TSB=-TSG
TDB=-TDG
TGB=TSG+TDG
RETURN

ROUTINE TO CALCULATE SURFACE POTENTIAL AT GIVEN V
50  TSI=V+US
51  DO 55 I=1,100
  E=(TSI-V-TFIF)/TK
  IF (E.LT.-100.) E=-100.
  QS=-BK*SQR(TSI-VB+TK*(EXP(E)-1))
  SI=TSI-(VG-TSI+QS)/((.5*BK**2*(1+EXP(E))/QS-1)
  IF (ABS(SI-TSI).LE.1.0E-06) GO TO 60
55 TSI = SI
60 SIS = SI
   GO TO KK, (1, 10, 15, 25)
* 
   END
LIST OF REFERENCES


   
   


BIOGRAPHICAL SKETCH

José Ignacio Arreola was born in Mexico City, Mexico, on January 10, 1950. He received the degree of Ingeniero Mecánico Electricista from the Universidad Iberoamericana in Mexico City in 1973. He worked for Instituto Nacional de Astrofísica, Optica y Electrónica located in Puebla, Mexico, for one year. José Ignacio has been a Fellow from Consejo Nacional de Ciencia y Tecnología at the University of Florida since June, 1974. He received the degree of Master of Science, major in Electrical Engineering, in August 1975.
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Fredrik A. Lindholm, Chairman
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Alan D. Sutherland
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

David R. MacQuigg
Assistant Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Earnest D. Adams
Professor of Physics
This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

March 1978

[Signature]

Dean, College of Engineering

[Signature]

Dean, Graduate School