INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International

300 N. ZEEB RD., ANN ARBOR, MI 48106

8217401

Chan, Tzo Yao

DESIGN OF MOS INTEGRATED CIRCUITS AT HIGH TEMPERATURE

The University of Arizona Ph.D. 1982

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

DESIGN OF MOS INTEGRATED CIRCUITS AT HIGH TEMPERATURE

by

Tzo Yao Chan

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 2

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Tzo Yao Chan

entitled DESIGN OF MOS INTEGRATED CIRCUITS AT HIGH TEMPERATURE

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of DOCTOR OF PHILOSOPHY.

D J Hamilton — Date 1/21/82

W J Kerwin — Date 1/21/82

J Fordemwalt — Date 1/21/82

J. J. Vuillemin — Date 1/21/82

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

D J Hamilton
Dissertation Director

1/21/82
Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: [signature]

ACKNOWLEDGMENTS

The author wishes to thank his academic advisor, Dr. Douglas J. Hamilton, for his valuable guidance and assistance throughout this work, and Dr. William J. Kerwin for his many helpful discussions. Technical discussions with Dr. James N. Fordemwalt are gratefully acknowledged. The invaluable assistance from Mr. James A. Homoki, Mr. Leonard S. Raymond, and other members of the Microelectronics Laboratory of The University of Arizona is greatly appreciated. Special thanks are due Mrs. Freida H. Long, who prepared, commented on, and typed the manuscript.

He also wishes to thank the Motorola Foundation, National Science Foundation and The University of Arizona for their support and financial assistance. In addition, the completion of this dissertation was made possible in part by an International Harvester Pre-Doctorate Fellowship.

Finally, he wants to express his personal thanks to his wife, Yuntien, for her many years of patience and understanding.

TABLE OF CONTENTS

		Page
LIST OF ILLUSTRATIONS		viii
LIST OF TABLES		xiv
ABSTRACT		xv
CHAPTER		
1	INTRODUCTION	1
2	THRESHOLD VOLTAGE AT HIGH TEMPERATURE	6
	2.1 Introduction	6
	2.2 Threshold Voltage at High Temperature	6
	2.3 Second Order Effects	10
	2.3.1 Physical Size Change and Permeability Change	10
	2.3.2 Thermal Carrier Generation	11
	2.3.3 Ionization of Surface States	13
	2.4 Short/Narrow Channel Effect	17
	2.4.1 Short-Channel Effect	17
	2.4.2 Narrow-Channel Effect	19
	2.5 Effect of Substrate Bias on the Threshold Voltage Sensitivity	23
	2.6 Substrate Feedback Stabilization of Threshold Voltage	25
	2.7 Threshold Voltage for Ion-Implanted Devices	28
	2.8 V_{th} Sensitivity versus Temperature of Ion-Implanted Enhancement MOS Devices	35
	2.9 Summary	38
3	LEAKAGE CURRENT AT HIGH TEMPERATURE	39
	3.1 Introduction	39
	3.2 Junction Leakage Current	39
	3.2.1 Generation-Recombination Leakage Current in the Space-Charge Region	39
	3.2.2 Diffusion Leakage Current	40
	3.3 Subthreshold Current	46
	3.4 Distributed Nature of the Channel Leakage Current	50
	3.5 Gate Leakage Current	65

TABLE OF CONTENTS--Continued

Page

3.6 Methods to Reduce Leakage Current Density at High Temperature 66
3.6.1 Guard Ring Collection of Minority Carriers 66
3.6.2 Epitaxial MOS 66
3.6.3 LOCOS Isolation 71
3.6.4 Silicon on Insulator Devices 71
3.6.5 Gettering Effect 73
3.7 Summary 74

4 ADDITIONAL FACTORS AFFECTING HIGH TEMPERATURE OPERATION OF MOS DEVICES 76

4.1 Introduction 76
4.2 Avalanche Multiplication and Breakdown Mechanism 77
4.2.1 Case (i): Junction Multiplication and Breakdown 78
4.2.2 Case (ii): Channel and Channel-to-Substrate Leakage Current Multiplication 82
4.3 Hot-Electron Effect 86
4.3.1 P-N Junction Avalanche Plasma 87
4.3.2 Channel Hot-Electron Effect 87
4.3.3 Substrate Hot-Electron Effect 88
4.4 Electromigration of Aluminum Metallization 91
4.5 Summary 94

5 MOS CIRCUITS AT HIGH TEMPERATURE 96

5.1 Introduction 96
5.2 Onset Temperature 96
5.3 Zero-TC Current 99
5.4 Inverter Behavior at High Temperatures 110
5.5 Improvement of Inverter Characteristics at High Temperatures 119
5.5.1 Improvement of Inverter Characteristics by Geometry Layout Change 119
5.5.2 Improvement of Inverter Characteristics by Charge Pumping 121
5.5.3 Choice of Inverter Circuit 123
5.5.4 Diode Compensation 126
5.6 Dynamic Logic: A Shift Register 130
5.7 Diode Protection in Logic Circuits 143
5.8 Summary 148

TABLE OF CONTENTS--Continued

Page

6 MAXIMUM OPERATING TEMPERATURE 149

6.1 Introduction 149
6.2 Substrate Concentration Considerations 149
6.2.1 Effect on Surface Channel Depletion Region 149
6.2.2 Effect on Threshold Voltages 150
6.3 Temperature Limit Due to Leakage Current Density 157
6.3.1 Channel to Leakage Current Ratio Consideration 157
6.3.2 Leakage Current Effect on the Source Junction 160
6.3.3 Leakage Current Effect on Inverter Characteristics 162
6.3.4 Maximum Fan-In Number in a Logic Circuit . . 165
6.3.5 Thermal/Radiation Latch-Up in CMOS Circuit 169
6.4 Ion Implanted Depletion-Load Devices 178
6.5 Summary 182

7 SCALING AT HIGH TEMPERATURES 185

7.1 Introduction 185
7.2 Conventional Scaling Rules 185
7.3 High-Temperature Oriented Scaling Rules 201
7.4 Summary 204

8 EXPERIMENTAL DATA FOR LEAKAGE CURRENT REDUCTION 205

8.1 Introduction 205
8.2 Background and Process Outline 205
8.3 Measured Leakage Currents for Epitaxial Diodes . . . 207

9 CONCLUSIONS AND RECOMMENDATIONS 210

9.1 Introduction 210
9.2 Original Contributions of the Dissertation 212
9.3 Recommendations for Future Work 214

APPENDIX A: THE USE OF DARE P PROGRAM 215

APPENDIX B: INVERTER SIMULATION PROGRAM 218

TABLE OF CONTENTS--Continued

Page

APPENDIX C: ION IMPLANTED DEPLETION TYPE MOS DEVICES 227

APPENDIX D: EPITAXIAL DIODE FABRICATION PROCEDURES 237

LIST OF REFERENCES 240

LIST OF ILLUSTRATIONS

Figure		Page
2.1a	Variation of V_{th} Terms with Temperatures	8
2.1b	Overall V_{th} Variation with Temperatures	8
2.2a	NMOS Device Showing Various Depletion Regions	18
2.2b	Surface Depletion Region with a Trapezoidal Shape for Short Channel Devices	18
2.2c	Schematic Diagram to Calculate V_{th}	18
2.3a	Schematic Diagram Showing the Surface Depletion Regions .	20
2.3b	Surface Depletion Region for a Narrow Channel MOS Device	20
2.4	V_{th} versus Temperatures for Different Substrate Bias Voltages and Substrate Concentrations Showing Reduced Threshold Voltage Sensitivity by Using Lower Substrate Concentration and Higher Substrate Bias Voltage	24
2.5	V_{th} versus Temperature at Different Substrate-to-Source Bias Voltages	26
2.6	Stabilization of the Inverter Transition Point by Temperature-Controlled Substrate Voltage Feedback	26
2.7a	Ion-Implant Profile for an Enhancement Type NMOS Device .	30
2.7b	Electron Potential Along x Direction	30
2.8a	A Shadow Threshold-Tailoring Implant with Decreased V_{th} Sensitivity versus Temperature	37
2.8b	A Deep Threshold-Tailoring Implant with Increased V_{th} Sensitivity versus Temperature	37
3.1	Measured and Theoretical Value of the Drain-Substrate Junction Leakage Current	44
3.2	Drain Current Characteristics at 300°C	45

LIST OF ILLUSTRATONS--Continued

Figure		Page
3.3	Subthreshold Characteristics of an NMOS Device at Three Different Temperatures	52
3.4	Distributed Nature of Substrate-To-Channel Leakage Current Showing the Interaction Between Channel and Channel-to-Substrate Leakage Current	54
3.5	A DC Model Showing Various Current Components by Taking Into Account the Interaction of Channel Current and Channel-to-Substrate Leakage Current	54
3.6	One-Dimensional DARE P Solution of Equation (3.27) with and without Taking into Account the Interaction of Channel Current and Channel-to-Substrate Leakage Current	59
3.7	Voltage, Electric Field, and Surface Charge Density Along the Channel at 400°C	60
3.8	Voltage, Electric Field and Surface Charge Density Along the Channel at 425°C	61
3.9	An NMOS at Very High Temperature Showing the Influence of Injected Electrons from the Substrate	64
3.10	Guard Ring Collection of Minority Carriers in the Substrate	67
3.11a	An Epitaxial MOS Device	69
3.11b	Improvement Factor of the Diffusion Leakage Current Density for Epitaxial MOS Devices	69
3.12	LOCOS Epi MOS Devices	72
4.1	Current Multiplication of NMOS Devices at High Temperature	79
4.2	Current Multiplication and Breakdown Characteristics of an NMOS Device at 23°C and 250°C	85
5.1	Experimentally Measured Drain Current as a Function of Temperature; Currents Are Matched at 25°C and 215°C by Varying Gate Bias Voltage	105

LIST OF ILLUSTRATIONS--Continued

Figure | Page

5.2 Experimental Measured Drain Current as a Function of Temperature; Currents Are Matched at 25°C and 215°C by Varying Substrate Bias Voltage 106

5.3 An NMOS Device Where Zero-TC Does Not Exist 107

5.4 Schematic Diagram Showing All the Current Components of an NMOS Inverter 111

5.5 Schematic Diagram Showing All the Current Components of a PMOS Inverter 113

5.6 Schematic Diagram Showing All the Current Components of a CMOS Inverter 113

5.7 Measured and Simulated Inverter Transfer Curves at 25°C, 150°C, and 300°C 117

5.8 Two Different Layout Geometries of a Saturated-Load Inverter 120

5.9 Charge-Pumping Technique to Boost the Output Voltage of an Inverter 122

5.10 Load Line Curve for Three Different Types of Inversters . 124

5.11 Schematic Diagram and Cross-Section View of a Diode-Compensated Saturated-Load Inverter 127

5.12a Small Signal Equivalent Circuit of a Diode-Compensated Saturated-Load Inverter 129

5.12b Pole-Zero Positions of the Saturated-Load Inverter; Arrow Shows the Shift of Pole Position Due to the Junction Depletion Capacitance of the Diode 129

5.13a Transfer Curve of a Saturated-Load Inverter at 25°C, With and Without Diode Compensation 131

5.13b Transfer Curve of a Saturated-Load Inverter at 300°C Showing the Effect of Diode Compensation 131

5.14 Transient Response of the Saturated-Load Inverter, Process Parameters as Before 132

LIST OF ILLUSTRATIONS--Continued

Figure | Page

5.15 A Saturated-Load Shift Register 134

5.16 Voltages at Node 3, 5, 7 and 8 of Above Shift Register at 25°C, Where Leakage Current is Negligible 136

5.17a Voltages at Node 3, 5, 7, 8 at 300°C. $J_L = 9.7301 \times 10^{-2}$ A/cm^2 is assumed 138

5.17b Voltages at Node 3, 5, 7, 8 at 300°C. $J_L = 9.7301 \times 10^{-2}$ A/cm^2 is assumed. All the frequencies are doubled from previous figure. 139

5.18a Voltages at Node 3, 5, 7, 8 at 300°C. $J_L = 9.7301 \times 10^{-1}$ A/cm^2 is assumed. Obvious logic error shown at V8 . . . 141

5.18b Voltage at Node 3, 5, 7, 8 at 300°C. $J_L = 9.7301 \times 10^{-1}$ A/cm^2 and all frequencies are ten times higher than previous figure. $V_{o_{max}}$ drop at Node 7 is the main reason for logic error shown at V8 142

5.19 Diode Protection Schemes in NMOS, PMOS and CMOS Logic Circuits 144

5.20 Thick Oxide Diode Connected Transistor Protection Scheme in CMOS Circuits 145

5.21 CMOS Characteristics and a New Protection Scheme 146

6.1 Maximum Operation Temperature versus Substrate Concentration and Resistivity at Ensure $N_a > n_i$ (T_{max}) 151

6.2 T_{max} versus Substrate Concentration to Assure V_{th} (T_{max}) > 0. Also shown is the threshold voltage at 25°C 153

6.3 T_{max} versus Substrate Concentration to Assume V_{th} (T_{max}) > 0. Also shown is the threshold voltage at 35°C 154

6.4a Band Bending Diagram for a PMOS Device 155

6.4b V_{th} versus Substrate Concentration at 25°C and 400°C for a P-Channel MOS Device 155

6.5 V_{th} versus Substrate Concentration at 25°C and 400°C for a P-Channel MOS Device 156

LIST OF ILLUSTRATIONS--Continued

Figure		Page
6.6	Load Line Characteristics of a Resistor-Loaded MOS Transistor at High Temperature	158
6.7	Channel-to-Leakage Current Ratio for a Typical NMOS Device	159
6.8a	MOS Device at High Temperature Showing the Leakage Current Components	161
6.8b	MOS Device with Parasitic Bipolar Transistor Action . . .	161
6.9	Transfer Characteristics of a Saturated-Load Inverter . .	164
6.10	A CMOS NAND Gate Circuit Showing All Important Leakage Components When All the Inputs are High	166
6.11	A CMOS NAND Gate Circuit Showing All Important Leakage Components When at Least One of the Inputs Are High . . .	168
6.12	Latch-Up in CMOS Due to Voltage Surge	170
6.13	Equivalent Circuit of the Latch	170
6.14	Small Signal Circuit Diagram of Fig. 6.13 Showing the Feedback Loop	171
6.15	Thermal Latch-Up in CMOS Circuit	175
6.16	Epi CMOS to Reduce Betas of the Parasitic Transistors and Also the Diffusion Leakage Current Density	177
6.17	Process Modification in CMOS to Eliminate R_1 and R_2 . . .	177
6.18a	Ion-Implanted Depletion Type NMOS with Box Profile Approximation	181
6.18b	Maximum Implant Depth versus Temperature to Ensure that the Device Could Be Pinched Off by Gate Bias; Also Shown Is the Corresponding Maximum Implant Dose	181
7.1a	A Polysilicon Gate NMOS Device	186
7.1b	Same Device after Scaling with K = 3.81	186

LIST OF ILLUSTRATIONS--*Continued*

Figure Page

8.1 Leakage Currents versus Temperature of Epitaxial (a) and Standard (b) Diodes 208

B.1 Flow Diagram of the Saturated-Load Inverter 219

C.1 Flow Diagram of the Ion-Implanted Depletion MOS Device Simulation Program 228

LIST OF TABLES

Table Page

2.1 Threshold Variation Due to Short/Narrow Channel Effect. . 22

3.1 Band Bending Factor b and Corresponding Surface Minority Carrier Concentration n_s at Different Temperatures 49

3.2 Values of b_1, V_{g1}, b_2, V_{g2}, and ΔV_g at Different Temperatures 51

4.1 Threshold Voltage Drift Due to Substrate Hot-Electron Effect 92

7.1 Value of $\phi'_{MS} + \phi_B$ for Aluminum Gate and N^+ Polysilicon Gate MOS Devices 190

ABSTRACT

Areas which require high-temperature MOS circuits are instrumentations for geothermal and petroleum well-logging, space exploration, aero-propulsion systems, and other hostile environments.

MOS digital circuits at high temperature are examineed as well as the maximum operating temperature of MOS devices. Factors affecting high-temperature operation of these devices, including threshold voltage sensitivity, mobility degradation, leakage current characterization and interactions, zero-TC current in analog applications and reliability considerations, are discussed. Methods to reduce threshold voltage sensitivities, process modifications to reduce leakage current density at high temperature, circuit techniques to eliminate the leakage current effects, diode compensation, CMOS thermal latch-up and MOS scaling rules at high temperature are investigated. Experimental results of epitaxial diodes to verify the leakage current reduction effect are discussed.

CHAPTER 1

INTRODUCTION

This dissertation research is directed toward the broad goal of providing electronic instrumentation for high-temperature environment applications. Examples of the need for such instrumentation are:

(i) Space exploration: where such electronic devices as transmitters, electromechanical devices, and ultrastable oscillators are expected to operate at temperatures as high as 350°C [Jurgen, 1981];

(ii) Fossil energy plants: where temperatures as high as 420°C are encountered for carbon steel piping [Managan, 1981];

(iii) Aero-propulsion systems: where the operating temperature requirements for operational aircraft engine monitoring systems are usually about 300°C when fuel cooling is not used [Nieberding and Powell, 1981];

(iv) Geothermal and well-logging industries: where temperatures as high as 250°C will be reached at a depth of only 18000 feet [Sanders, 1981].

Large scale integrated circuit (LSI) technology can provide the circuit complexity necessary in the small space available, but unfortunately, circuits that are presently available are only capable of operation up to 150°C. Some specially designed integrated circuits have operated at 260°C [Palkuti, 1979], while hybrid circuits have

been constructed for operation at 300°C. To use presently available high-complexity LSI circuits requires some special thermal insulation such as a Dewar flask.

The dominant LSI technology is the silicon MOS technology used for such sophisticated system as microprocessors. MOS devices are majority carrier devices and are therefore more suitable for high-temperature operation than bipolar devices. If the temperature range of silicon MOS circuits could be extended to above 300°C. much more instrumentation could be included in the hostile environments mentioned above, and a much higher degree of sophisticaltion in instrumentation systems could be employed. Work at The University of Arizona has demonstrated operation of N-channel enhancement saturated-load inverters at 300°C [Cosentino, 1980]. Recently, a dielectric isolation CMOS process which allows circuit operation at 350°C [Beasom, Moore, Mohammed, and Draper, 1981] is reported. What is now needed is to determine the ultimate temperature limitation for MOS circuits and to determine how LSI circuits must be designed and fabricated in order to operate at this temperature. Accordingly, the general objective of this dissertation is to determine the maximum temperature limitation for silicon MOS integrated circuits. The specific objectives are:

1. To evaluate the temperature dependence of all the factors which influence the terminal behavior of MOS devices.

2. To determine what processing modifications are necessary to fabricate devices which operate at the maximum temperature.

3. To develop device and circuit design guidelines for circuits operating at the maximum temperature.

4. To fabricate test devices to obtain experimental confirmation of theoretical results.

Briefly stated, the most important results of the dissertation are:

1. Methods to reduce leakage current density are developed. These include epitaxial MOS devices, guard-ring MOS devices, and others.

2. Maximum temperature limits for MOS device operation are investigated on both the device and the circuit levels. Guidelines to increase the maximum temperature limit are given.

3. Diode-compensated inverters are developed to improve the performance of the all N-channel inverter.

4. Design rules are developed for scaling device geometry and processing to achieve small size in a high-temperature environment.

5. Threshold voltage stabilization by a temperature-variable substrate bias voltage technique is developed.

6. All leakage current components are characterized and a complete model for single channel inverter circuits is developed.

The disseration proceeds as follows. In Chapter 2 factors influencing the threshold voltage at high temperatures are considered, including second order effects. A method to stabilize threshold voltage at high temperature with a temperature-controlled substrate feedback bias is introduced.

Various leakage current components at high temperature and their interaction with channel current are discussed in Chapter 3. Measurement results are illustrated. Methods to reduce the leakage current density at high temperature are also introduced and compared.

Chapter 4 is concerned with other second-order factors which affect high-temperature operation of MOS devices. These include multiplication and breakdown mechanisms, hot electron/hole effects and reliability considerations for metal conductor stripes. Measured and calculated results are given.

MOS circuit design for high-temperature applications is investigated in Chapter 5. Emphasis is on digital circuits. Measured and simulated inverter transfer characteristics are discussed. Methods to improve high-temperature inverter circuit performance are examined. Also a dynamic logic circuit at high temperature is simulated. A method is described for achieving zero-temperature-coefficient (ZTC) current for analog circuit applications.

The maximum temperature limitations for MOS circuit operation are discussed in Chapter 6. Both physical limitations and circuit performance limitations are considered. Parasitic bipolar transistor effects and the CMOS thermal latch-up mechanism are investigated.

One of the most important aspects of device design is the scaling of device size as the technology advances and achievable dimensions continue to become smaller. It is important to determine how the result of scaling influence device behavior at high temperature.

Chapter 7 deals with scaling rules at high temperature, and high-temperature-oriented scaling rules are derived.

Experimental results to verify the high-temperature design are presented in Chapter 8.

Finally, MOS device characteristics, circuit and process modifications, and experimental results for high-temperature applications are summarized in Chapter 9. Recommendations for future work are given.

CHAPTER 2

THRESHOLD VOLTAGE AT HIGH TEMPERATURE

2.1 Introduction

Threshold voltage is a strong function of temperature. Devices must be properly designed so that, for example, enhancement devices do not become depletion devices at high temperature. Threshold voltage must be correctly modeled in order to calculate the effect of temperature on DC device behavior. Threshold voltage is also important in determining the noise margin NM in digital circuitry.

In this chapter we discuss the threshold variation with temperature; this includes the variation of the Fermi energy ϕ_f , flat band voltage and depletion charge density. Second order effects such as geometry change, short/narrow channel effects and surface state ionization effects are discussed. Techniques to stabilize threshold voltage at high temperature are investigated. Finally, since threshold tailoring implant and E beam/laser annealing are frequently used in modern VLSI circuit design, we discuss threshold voltage calculation for an ion-implanted enhancement MOS device and its threshold sensitivity to temperature.

2.2 Threshold Voltage at High Temperature

Threshold voltage in its simplest form is

$$V_{th} = V_{fb} + 2\phi_f - \frac{Q_b}{C_{ox}} \tag{2.1}$$

where

$$V_{fb} = \phi'_{MS} - \frac{Q_{SS}}{C_{ox}} \tag{2.2}$$

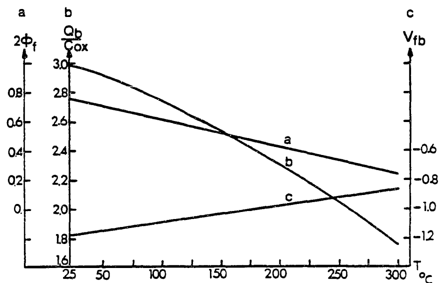
$$Q_b = \sqrt{2\varepsilon_s \, qN_a \, (2\phi_f - V_{sub})} \tag{2.3}$$

$$\phi'_{MS} = \phi_{Mos} - \phi_{Sox} - \frac{E_g}{2q} - \phi_f \tag{2.4}$$

where V_{fb} is the flat-band voltage, ϕ_f is the Fermi potential of the substrate, Q_b is the depletion charge density, C_{ox} is the gate oxide capacitance per unit area, V_{sub} is the substrate bias voltage relative to source, ϕ_{Mos} is the metal-silicon dioxide barrier energy, ϕ_{Sox} is the semiconductor-silicon dioxide barrier energy, E_g is the energy gap of silicon, Q_{SS} is the surface state density at the silicon-silicon dioxide interface and ϕ'_{MS} is the metal-silicon work function difference.

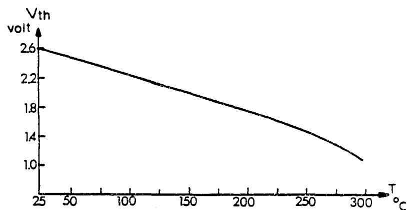
In addition to its explicit appearance in the $2\phi_f$ term, ϕ_f is implicitly involved in V_{fb} and Q_b/C_{ox} term. ϕ_f reduces the threshold voltage for N-channel devices and increases the threshold voltage for P-channel devices. This makes an enhancement-type device become a depletion-type device at high temperature. Physically speaking, the substrate becomes more intrinsic at high temperature and it is therefore easier to reach the strong inversion condition.

For a quantitative measure of the variation of each term in the threshold voltage expression we plot V_{fb}, $2\phi_f$ and Q_b/C_{ox} as functions of temperature in Fig. 2.1a for an aluminum gate N-channel device with



(Note the scale offset for these terms)

Fig. 2.1(a) Variation of V_{th} Terms with Temperatures



$N_a = 3.4 \times 10^{16}/cm^3$; $T_{ox} = 1100Å$; $\frac{Q_{SS}}{q} = 4 \times 10^{10}/cm^2$; $V_{sub} = 0$ volt; Aluminum gate NMOS

Fig. 2.1(b) Overall V_{th} Variation with Temperatures

N_a = 3.4 x $10^{16}/cm^3$, T_{ox} = 1,100 A, Q_{SS} = 4 x $10^{10}/cm^2$ and V_{sub} = 0 volt. $\frac{Q_b}{C_{ox}}$ is the dominant factor. The temperature variation of the threshold voltage of this device is shown in Fig. 2.1b. Note that $V_{th} \cong 1$ volt at 300°C; it has relatively constant slope up to 225°C and decreases faster beyond that point. Note that to ensure $V_{th} > 0$ at T = 300°C, N_a is chosen to be higher than conventional NMOS substrate doping concentration.

For more accurate calculations, an expression of the type $\frac{\Delta V_{th}}{\Delta T} = a + bT_c^c$ can be used to fit these simulated data. With this device we obtain

$$\frac{\Delta V_{th}}{\Delta T} = \left(-4.152 - 8.96 \times 10^{-5}\ T_c^{1.82}\right) mV/°C \tag{2.5}$$

where T_c is in °C. For comparsion, for the same device with V_{sub} = -2 volt, we obtain

$$\frac{\Delta V_{th}}{\Delta T} = \left(-2.537 - 1.0043 \times 10^{-4} T^{1.495}\right) mV/°C \tag{2.6}$$

In this case the substrate bias reduces the temperature dependence of V_{th}; this is discussed in a later section.

Note that in (2.5) the second term typically does not reach 50% of the first term until T_c increases to 250°C. Thus, for a reasonable approximation, we use

$$\frac{\Delta V_{th}}{\Delta T} = a = \text{constant} \tag{2.7}$$

2.3 Second Order Effects

2.3.1 Physical Size Change and Permeability Change

The oxide capacitance per unit area in the threshold voltage expression is

$$C_{ox} = \frac{\varepsilon_{ox}}{T_{ox}} \tag{2.8}$$

Variation with temperature comes from both ε_{ox} and T_{ox}.

$$\frac{1}{\Delta T}\frac{\Delta C_{ox}}{C_{ox}} = \left[\frac{\dfrac{\varepsilon_{ox} + \Delta\varepsilon_{ox}}{T_{ox} + \Delta T_{ox}} - \dfrac{\varepsilon_{ox}}{T_{ox}}}{(\varepsilon_{ox}/T_{ox})}\right]\frac{1}{\Delta T}$$

$$= \left(\frac{\Delta\varepsilon_{ox}}{\varepsilon_{ox}} - \frac{\Delta T_{ox}}{T_{ox}}\right)\frac{1}{\Delta T} \tag{2.9}$$

Using

$$\frac{1}{\Delta T}\frac{\Delta\varepsilon_{ox}}{\varepsilon_{ox}} \cong 80 \text{ ppm/°C} \qquad \text{[Raymond, 1980]}$$

and

$$\frac{1}{\Delta T}\frac{\Delta T_{ox}}{T_{ox}} \cong 9\text{-}14 \text{ ppm/°C} \qquad \text{[Kingery, Bowen and Uhlmann, 1976]}$$

we have

$$\frac{1}{\Delta T}\frac{\Delta C_{ox}}{C_{ox}} = 66\text{-}71 \text{ ppm/°C} \tag{2.10}$$

For temperature variation between 25°C and 300°C, the capacitance change is

$$\frac{\Delta C_{ox}}{C_{ox}} \cong 1.95\% \tag{2.11}$$

that is

$$C_{ox}\ (300°C) \cong C_{ox}\ (25°C) \times 1.02 \tag{2.12}$$

This variation will thus produce a small reduction of threshold voltage.

2.3.2 Thermal Carrier Generation

The thermally generated hole-electron pairs under the gate will influence the high temperature device characteristics. Electrons generated there will change the effective substrate doping concentration and contribute to the channel current; holes will contribute to the substrate current.

In thermal equilibrium

$$n_o p_o = \text{constant} = n_i^2\ (T) \tag{2.13}$$

where

$n_i(T) \triangleq$ intrinsic carrier concentration at temperature T

n_o = electron concentration at equilibrium

p_o = hole concentration at equilibrium

so

$$n_o p_o = n_i^2 \, (300°K) \tag{2.14}$$

$$(n_o + \Delta n)(p_o + \Delta p) = n_i^2 \, (T) \tag{2.15}$$

$$\Delta n = \Delta p \qquad \text{at temperature } T$$

Then

$$(\Delta n)^2 + (p_o + n_o)\Delta n = n_i^2 \, (T) - n_i^2 \, (300°K) \tag{2.16}$$

from which we obtain

$$\Delta n = -\frac{p_o}{2} + \sqrt{\frac{p_o^2}{4} + \left[n_i^2 \, (T) - n_i^2 (300°K)\right]} \tag{2.17}$$

For example, if $n_i(300°K) = 1.45 \times 10^{10}/cm^3$ and $n_i(673°K) = 2.1 \times 10^{15}/cm^3$ and $p_o = N_a = 3.4 \times 10^{16}/cm^3$

$$\Delta n = 1.292 \times 10^{14}/cm^3 = 0.38\% \text{ of } N_a \tag{2.18}$$

this will change ϕ_f and V_{th} as follows

$$\phi_f' = \frac{kT}{q} \ln \frac{N_a - \Delta n}{n_i} < \phi_f \tag{2.19}$$

$$V_{th}' = V_{fb}' - \frac{Q_b'}{C_{ox}} + 2\phi_f' < V_{th} \tag{2.20}$$

It will lower the threshold voltage by a small amount; for the previous device example above, we find $\Delta V_{th} = 3.34$ mV at 300°C.

2.3.3 Ionization of Surface States

Surface state charge density at the silicon-silicon dioxide interface is an important item for the threshold voltage calculation (where charge neutrality condition has to be preserved). These surface states, located at the interface and with energy level distributed over the energy gap, also function as trap/emission centers for electrons or holes, depending on the relative locations of their energy level and the Fermi energy level. At equilibrium with temperature T the number of occupied surface states is fixed; when the temperature changes a new equilibrium condition will be reached with a different number of occupied surface states resulting in different charge density Q_{SS} in the surface states and different V_{th}.

Note that the charge neutrality condition must be satisfied at all temperatures

$$Q_G + Q_{SS} + Q_{SD} + Q_n = 0 \qquad (2.21)$$

where $Q_G \triangleq$ charge density on the gate electrode per unit area

$\triangleq C_{ox} V_{ox}$

$Q_{SS} \triangleq$ occupied surface state density per unit area

$Q_{SD} \triangleq$ surface depletion charge density per unit area

$Q_n \triangleq$ charge density per unit area in the inversion channel

For a p-type substrate, only acceptor states are considered because these states can interact with the conduction band electrons in the inversion layer and the acceptor states have energy near the conduction band edge.

Q_G is found to be

$$Q_G \triangleq C_{ox}V_{ox} = [V_g - V_{sub} - \phi'_{MS}(\phi_f) - \phi_s(\phi_f)]C_{ox} \tag{2.22}$$

The charge density in the surface states is equal to

$$Q_{SS} = -q \int \frac{N_{SS}(E)}{1 + \exp\left(\frac{E-E_f}{kT}\right)} dE \tag{2.23}$$

where $N_{SS}(E)$ is the total number of acceptor states with energy E per unit area. The total charge in silicon is found from integrating Poisson's Equation [Richman, 1969]

$$Q_{SD} + Q_n = -2qn_i \left[\frac{\varepsilon_s kT}{2n_i q^2}\right]^{1/2} \left\{ 2\frac{q(\phi_s - V_{sub})}{kT} \sinh\frac{q\phi_f}{kT} + e^{\frac{q}{kT}(\phi_f - \phi_s + V_{sub})} - e^{\frac{q}{kT}\phi_f} + e^{\frac{q}{kT}(-\phi_f + \phi_s)} - e^{\frac{q}{kT}(-\phi_f + V_{sub})} \right\}^{1/2} \tag{2.24}$$

where

V_{sub} is the substrate bias voltage

ϕ_f is the Fermi energy potential

ϕ_s is the surface potential ($\phi_f = E_f/kT$)

Equation (2.21) becomes

$$Q_G(\phi_f) + [Q_{SD} + Q_n](\phi_f) = - Q_{SS}(\phi_f) \tag{2.25}$$

Note that at the threshold point $\phi(0) = 2\phi_f$. As temperature increases, the above transcendental equation must be solved to find ϕ_f and Q_{SS}. The change in surface charge density when temperature increases from T_1 to T_2 is

$$\Delta Q_{SS} = -q \int \frac{N_{SS}(E)}{1 + \exp \frac{E - E_{f2}}{kT_2}} dE + q \int \frac{N_{SS}(E)}{1 + \exp \frac{E - E_{f1}}{kT_1}} dE > 0 \tag{2.26}$$

where $E_{f1} = E_i - q\phi_{f1}$, $E_{f2} = E_i - q\phi_{f2}$ and E_i is the intrinsic energy level of silicon and the resultant threshold voltage change is

$$\Delta V_{th} = \frac{-\Delta Q_{SS}}{\varepsilon_{ox}} T_{ox} < 0 \tag{2.27}$$

It is often assumed that the distribution of $N_{SS}(E)$ around the maximum state density $N_{SS}(E_{sm})$ is Gaussian:

$$N_{SS}(E) = N_{SS}^{o} \frac{1}{\sqrt{2\pi}\,\delta} \exp \left[\frac{-(E-E_{sm})}{2\Delta^2}\right]^2 \tag{2.28}$$

where N_{SS}^{o} is the total number of acceptor states per unit area, δ is the standard deviation of surface state density. Recently deep level transient spectroscopy (DLTS) has been used to measure the interface-state density [Wang, 1980; Johnson, Bartlink and Schulz, 1978]. It has been found that the distribution of surface states across the energy gap depends on the orientation of the substrate, process control and annealing conditions.

Normally $(Q_{SD} + Q_n) >> Q_{SS}$. The introduction of surface charge would not change the Fermi energy level. ϕ_f is determined mainly by the substrate doping concentration; however, this change of ϕ_f becomes important when a large variation of temperature is encountered. Equations (2.26) and (2.27) are used to calculate ΔV_{th}.

For a quantitative measure of this effect, we consider an N channel device with T_{ox} = 1000Å, $Q_{SS}/q = 4 \times 10^{10}/cm^2$ at 25°C, N_a = $3.4 \times 10^{16}/cm^3$, ϕ_f(25°C) = 0.39 volt, then V_{th}(25°C) = 2.209 volts. If we make simplifying assumptions as follows:

(i) surface states are uniformly distributed in the energy gap

(ii) the introduction of surface states will not change ϕ_f

then at 300°C, we obtain ϕ_f (300°C) = 0.132 volt

$$\Delta Q_{SS} = 4.0 \times 10^{10} \times \frac{0.39 - 0.132}{(0.56 - 0.39)} \times 1.6\ 10^{-19} = 9.71 \times 10^{-5} \text{ coul} \tag{2.29}$$

and

$$V_{th} = \frac{-\Delta Q_{SS}}{C_{ox}} = -0.28 \text{ volt.} \tag{2.30}$$

2.4 Short/Narrow Channel Effect

When the dimensions of MOS devices are reduced as technologies advance, the conventional depletion approximation does not hold true, and the effect of geometry must be taken into account. This occurs when the channel dimensions are comparable to the maximum surface depletion width or the source and drain junction depletion width. For an accurate calculation of V_{th}, we have to solve Poisson's Equation in three dimensions with appropriate boundary conditions. If channel width W and channel length L of the device are larger than 2 μm, the following simple analysis can be used.

2.4.1 Short-Channel Effect [Yau, 1974]

As shown in Fig. 2.2a, because of the presence of the source and drain-to-substrate junctions, the surface depletion region has approximately a trapezoidal shape instead of rectangular. This effect is more obvious when short-channel devices are considered (Fig. 2.2b). Using the trapezoidal approximation, we have (Fig. 2.2c)

Fig. 2.2(a) NMOS Device Showing Various Depletion Regions

Fig. 2.2(b) Surface Depletion Region with a Trapezoidal Shape for Short Channel Devices

Fig. 2.2(c) Schematic Diagram to Calculate V_{th}

$$Q_b = qN_a\, x_{d_{max}} \frac{1}{L'} \left[\frac{1}{2} (L' + L'')\right] \tag{2.31}$$

$$r = r_j + x_1 \tag{2.32}$$

$$r^2 = x_{d_{max}}^2 + (\Delta L + r_j)^2 \tag{2.33}$$

$$L'' = L' - 2\Delta L \tag{2.34}$$

Solving for L" and substitute into threshold voltage expression, we obtain

$$V_{th} = \phi'_{MS} - \frac{Q_{SS}}{C_{ox}} + 2\phi_f + \frac{1}{C_{ox}} \sqrt{2\varepsilon_s\, qN_a\, (2\phi_f - V_{sub})} \cdot f \tag{2.35}$$

where

$$f = 1 - \frac{r_j}{L'} \left[\left(1 + \frac{2x_{d_{max}}}{r_j} + \frac{x_1^2 - x_{d_{max}}^2}{r_j^2}\right)^{\frac{1}{2}} - 1\right] \tag{2.36}$$

and

$$f(L \to \infty) = 1 . \tag{2.37}$$

2.4.2 Narrow-Channel Effect

Threshold voltage variation with W comes from the fact that the surface depletion region expands horizontally as well as vertically. Using the shape in Fig. 2.3a, we find

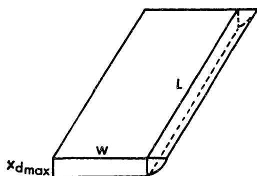


Fig. 2.3(a) Schematic Diagram Showing the Surface Depletion Regions

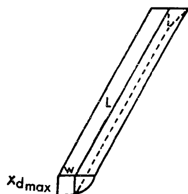


Fig. 2.3(b) Surface Depletion Region for a Narrow Channel MOS Device

$$V_{th} = \phi'_{MS} - \frac{Q_{SS}}{C_{ox}} + 2\phi_f + \frac{1}{C_{ox}} \sqrt{2\varepsilon_s qN_a(2\phi_f - V_{sub})} \cdot g \quad (2.38)$$

where

$$g = \frac{qN_a WL x_{d_{max}} + qN_a \frac{\pi x_{d_{max}}^2}{2} \cdot L}{qN_a WL x_{d_{max}}} \quad (2.39)$$

$$= 1 + \frac{\pi}{2W} x_{d_{max}} \quad (2.39)$$

and

$$g(w = \infty) = 1 . \quad (2.40)$$

Combining these effects, we obtain

$$V_{th} = \phi'_{MS} - \frac{Q_{SS}}{C_{ox}} + 2\phi_f + \frac{1}{C_{ox}} \sqrt{2\varepsilon_s qN_a(2\phi_f - V_{sub})} \cdot f \cdot g \quad (2.41)$$

As temperature increases, $2\phi_f$ decreases and variation of threshold voltage due to the channel dimension effects is less. However, it is still an important factor since V_{th} itself is reduced at high temperature.

Table 2.1 shows the results of short and narrow channel effects for an enhancement type MOS device with $N_a = 4.6 \times 10^{16}/cm^3$, $x_j = 1.5\ \mu m$, $T_{ox} = 700Å$, $Q_{SS}/q = 1.0 \times 10^{11}/cm^2$ and Al gate material. It should be noted that this device has thinner oxide than previous examples. As the oxide thickness increases, the short and narrow channel effects on threshold voltage are larger. For example, if this

Table 2.1 Threshold Variation Due to Short/Narrow Channel Effect

T(°C)	$\frac{W}{L}$ μm = $\frac{40}{10}$		$\frac{W}{L}$ μm = $\frac{2}{10}$		$\frac{W}{L}$ μm = $\frac{40}{2}$		$\frac{W}{L}$ μm = $\frac{2}{2}$	
	V_{th}	ΔV_{th}	V_{th}	ΔV_{th}	V_{th}	ΔV_{th}	V_{th}	ΔV_{th}
25	1.70	0	1.95	+0.25	1.57	-0.13	1.80	+0.10
80	1.52	0	1.72	+0.20	1.41	-0.11	1.61	+0.09
135	1.33	0	1.52	+0.19	1.23	-0.10	1.41	+0.08
190	1.12	0	1.28	+0.16	1.04	-0.08	1.19	+0.07
245	0.89	0	1.01	+0.12	0.83	-0.06	0.94	+0.05
300	0.63	0	0.72	+0.09	0.58	-0.05	0.67	+0.04

$N_a = 4.6 \times 10^{16}/cm^3$

$N_{S/D} = 5 \times 10^{19}/cm^3$

$x_j = 1.5$ μm

$T_{ox} = 700Å$

$Q_{SS}/q = 1.0 \times 10^{11}/cm^2$

Al gate

same device had T_{ox} = 1400Å, with W/L = 40 μm/10 μm, we find V_{th} = 3.60v at 25°C, V_{th}=1.66v at 300°C and with W/L = 2 μm/2 μm, we obtain V_{th} = 3.80v at 25°C and V_{th} = 1.73v at 300°C.

2.5 Effect of Substrate Bias on the Threshold Voltage Sensitivity

In Section 2.2 we learned that for an N-channel device, Q_b contributes greatly to the threshold voltage, where

$$Q_b = [2\varepsilon_s \, qN_a(2\phi_f - V_{sub})]^{\frac{1}{2}} \tag{2.42}$$

when V_{sub} = 0, Q_b has large temperature dependence because of ϕ_f, but it is less dependent on temperature if a nonzero substrate bias is used. The threshold voltage variation with temperature is

$$\frac{dV_{th}}{dT} = \frac{d(\phi'_{MS})}{dT} + 2\frac{d\phi_f}{dT} - \frac{1}{C_{ox}}\frac{dQ_b}{dT} \tag{2.43}$$

Since

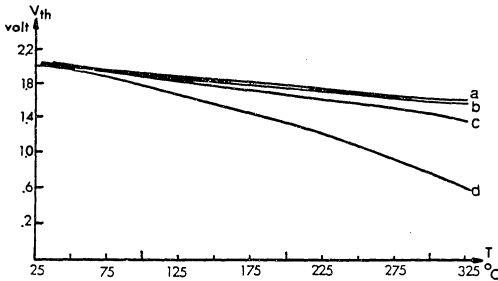
$$\frac{d(\phi'_{MS})}{dT} = \frac{1}{2q}\frac{dE_g}{dT} - \frac{d\phi_f}{dT} \simeq \frac{-d\phi_f}{dT} \tag{2.44}$$

we obtain

$$\frac{dV_{th}}{dT} \simeq \frac{d\phi_f}{dT}\left[\frac{-(4\varepsilon_s \, qN_a \, \phi_f)^{\frac{1}{2}}}{2C_{ox}\phi_f \, [1-(V_{sub}/2\phi_f)]^{\frac{1}{2}}} + 1\right] \tag{2.45}$$

Thus we see that dV_{th}/dT is less when nonzero V_{sub} is used.

Figure 2.4 shows this effect for four different substrate bias voltages. In this figure, for each V_{sub}, N_a has been chosen to produce



(a) $N_a = 3.0 \times 10^{15}/cm^3$; $V_{sub} = -6.44$ volts
(b) $N_a = 5.0 \times 10^{15}/cm^3$; $V_{sub} = -3.6$ volts
(c) $N_a = 1.0 \times 10^{16}/cm^3$; $V_{sub} = -1.4$ volts
(d) $N_a = 2.8 \times 10^{16}/cm^3$; $V_{sub} = 0$ volt

In all cases:

$T_{ox} = 1000\text{Å}$, $\frac{Q_{SS}}{q} = 4 \times 10^{10}/cm^2$, and aluminum gate.

Fig. 2.4 V_{th} versus Temperatures for Different Substrate Bias Voltages and Substrate Concentrations Showing Reduced Threshold Voltage Sensitivity by Using Lower Substrate Concentration and Higher Substrate Bias Volcage.

a room-temperature V_{th} of 2.0 volts. One disadvantage of using substrate bias is that high substrate bias voltage must be accompanied by lightly doped substrate to get a reasonable threshold voltage at room temperature. This results in high value of diffusion leakage current at high temperatures. One way to avoid this problem is to use an epitaxial MOS process as will be shown in Section 3.6.

2.6 Substrate Feedback Stabilization of Threshold Voltage

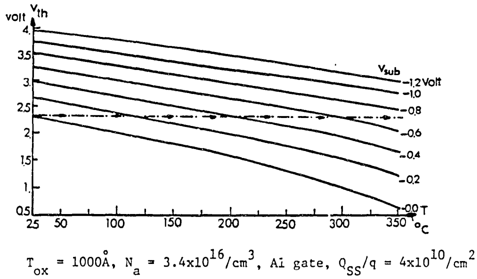
The above discussion indicates that one can control threshold voltage variation by appropriate variation of the substrate voltage. For a given substrate doping concentration, V_{th} increases with increasing substrate bias voltage and decreases with increasing temperature as shown in Fig. 2.5. However, if one increases the substrate bias as temperature increases (i.e., moving along the dashed line in Fig. 2.5), threshold variation may remain very small.

This idea is illustrated in Fig. 2.6. Here, a temperature sensing transistor Q_1 is used to change the substrate bias voltage of all other transistors. At room temperature, V_R and R are adjusted to give $V_{sub} = 0$. As temperature increases, the current in Q_1 drops, resulting in more negative substrate voltage which in turn will restore the threshold voltage of other transistors.

When $T = T_1$ = room temperature

$$V_{sub} = V_f^- + \frac{\beta}{2}\,[(V_R - V_{th_1})^2]R \cong 0 \qquad (2.46)$$

at $T = T_2$



T_{ox} = 1000Å, N_a = $3.4\times10^{16}/cm^3$, Al gate, Q_{SS}/q = $4\times10^{10}/cm^2$

Fig. 2.5 V_{th} versus Temperature at Different Substrate-to-Source Bias Voltages

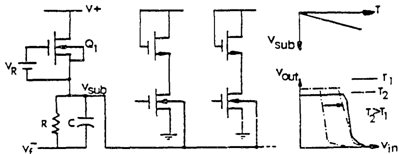


Fig. 2.6 Stabilization of the Inverter Transition Point by Temperature-Controlled Substrate Voltage Feedback

$$V'_{sub} = V_f^- + \frac{\beta'}{2} [(V_R - V_{th_2})^2] (R + \Delta R) \qquad \Delta R << R \tag{2.47}$$

where

$$V_{th_2} \cong V_{th_1} + a(T_2 - T_1)$$

$$\beta' \cong \beta \left(\frac{T_2}{T_1} \right)^{-1.5}$$

$$a \cong \left. \frac{dV_{th}}{dT} \right|_{T_1} \tag{2.48}$$

To obtain V'_{sub} at temperature T_2

$$R = \frac{2\Delta V_{sub}}{\beta'(V_R - V_{th_2})^2 - \beta(V_R - V_{th_1})^2} \tag{2.49}$$

where ΔV_{sub} is the required substrate voltage change to keep threshold voltage unchanged at $T = T_1$ and $T = T_2$. Note that temperature dependences of V_R and R are not included here.

In a practical circuit, use of a resistor would be undesirable. Rather, a more complex circuit would have to be designed to provide the proper variation of V_{sub} with temperature. For example, one might start with a voltage reference circuit [Blauschild, Tucci, Muller, and Meyer, 1978], and make appropriate modifications to obtain the required temperature behavior.

There is a disadvantage of this technique. Because of the non-zero substrate bias for the devices, body effect will be increased. For

PLEASE NOTE

Page(s) not included with original material
and unavailable from author or university.
Filmed as received.

UMI

Lately laser, ion beam and heat pulses had been applied [White and Peercy, 1979] to anneal the implant profile. Short period exposure by the laser or ion beam or heat pulses will leave the implanted profile almost unchanged, with the dopant carriers totally activated. This results in better profile control but makes the calculation of threshold voltage more difficult.

Consider the ion-implant profile for an enhancement type MOS device as shown in Fig. 2.7a. The dopant distribution is

$$P(x) = \frac{Q}{\sqrt{2\pi}\ \delta R_p} \exp\left[-\frac{(x - R_p)^2}{(\sqrt{2}\ \delta R_p)^2}\right] + N_a \tag{2.50}$$

where

$Q \triangleq$ total implant charge

$R_p \triangleq$ (Gaussian) average implant depth

$\delta R_p \triangleq$ standard deviation of the implant profile

$N_a \triangleq$ substrate doping concentration before implant

From Poisson's equation

$$\frac{d^2\phi}{dx^2} = \frac{qP(x)}{\varepsilon_s} \tag{2.51}$$

where ϕ is the electron potential in x direction as shown in Fig. 2.7b. The basic electrostatic equations of MOS devices are

$$\varepsilon_s[-\phi'(0)] - \varepsilon_{ox}E_{ox} = Q_{SS} \tag{2.52}$$

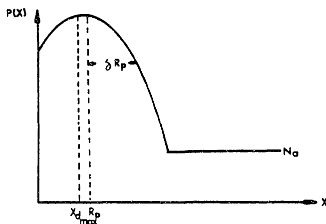


Fig. 2.7(a) Ion-Implant Profile for an Enhancement Type NMOS Device

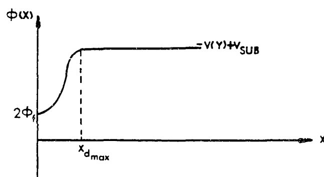


Fig. 2.7(b) Electron Potential Along x Direction

$$V_{ox} = E_{ox}T_{ox} \tag{2.53}$$

$$V_{gs} - \phi'_{MS} = V_{ox} + \phi(0) \tag{2.54}$$

The boundary conditions of the Poisson's equations are therefore obtained as follows:

(a) From Equations (2.52), (2.53), and (2.54), we obtain

$$\phi(0) = V_{gs} - \phi'_{MS} - \frac{Q_{SS}}{C_{ox}} - \frac{\varepsilon_s[-\phi'(0)]}{C_{ox}}$$

$$= V_{gs} - V_{fb} - \frac{\varepsilon_s[-\phi'(0)]}{C_{ox}} \tag{2.55}$$

(b) The field in the silicon at the edge of the depletion region is zero, so

$$\phi'(x_{dmax}) = 0 \tag{2.56}$$

(c) Electron potential at the maximum depletion width is $V(y) + V_{sub}$:

$$\phi(x_{dmax}) = -V(y) + V_{sub} \tag{2.57}$$

In the above equations

$V_{gs} \triangleq$ gate to substrate voltage

$V_{fb} \triangleq$ flat band voltage

$\phi'_{MS} \triangleq$ metal-silicon dioxide work function difference

$x_{d_{max}} \triangleq$ maximum depletion width

$V(y) \triangleq$ voltage along the channel (y) direction

$V_{sub} \triangleq$ substrate bias voltage realative to source

$Q_{SS} \triangleq$ surface state density per unit area

$C_{ox} \triangleq$ oxide capacitance per unit area

We choose the equilibrium Fermi energy level of the substrate as the potential reference.

Integrating Eq. (2.51) with Condition (b), we obtain

$$\phi'(x) = \frac{qQ}{2\varepsilon_s}\left[\operatorname{erf}\frac{x-R_p}{\sqrt{2}\,\delta R_p} - \operatorname{erf}\frac{(x_{d_{max}} - R_p)}{\sqrt{2}\,\delta R_p}\right] + \frac{q}{\varepsilon_s} N_a (x - x_{d_{max}}) \tag{2.58}$$

Using the following property of the error function

$$\int_0^x \operatorname{erf}(x)\,dx = x \operatorname{erf}(x)\Big|_0^x + \frac{1}{\sqrt{\pi}} e^{-x^2}\Big|_0^x \tag{2.59}$$

we can integrated Eq. (2.58) with Condition (c) to obtain

$$\phi(x) = \frac{qQ}{2\varepsilon_s}(x-R_p)\left[\operatorname{erf}\frac{(x-R_p)}{\sqrt{2}\,\delta R_p} - \operatorname{erf}\frac{(x_{d_{max}}-R_p)}{\sqrt{2}\,\delta R_p}\right] + \frac{qN_a}{2\varepsilon_s}(x-x_{d_{max}})^2$$

$$+ \frac{qQ\,\delta R_p}{\sqrt{2\pi}\,\varepsilon_s}\left[e^{-\frac{(x-R_p)^2}{2\,\delta R_p^2}} - e^{-\frac{(x_{d_{max}}-R_p)^2}{2\,\delta R_p^2}}\right] - V(y) + V_{sub} \tag{2.60}$$

At the threshold point $V(y) = 0$, $\phi(0) = 2\phi_f$. From Eq. (2.60)

$$\phi(0) = \frac{-qQ}{2\varepsilon_s} R_p \left[-\operatorname{erf} \frac{R_p}{\sqrt{2}\,\delta R_p} - \operatorname{erf} \frac{(x_{d_{max}} - R_p)}{\sqrt{2}\,\delta R_p} \right] + \frac{qN_a}{2\varepsilon_s} x_{d_{max}}^2$$

$$+ \frac{qQ\,\delta R_p}{\sqrt{\pi}\,\varepsilon_s} \left[\exp\left(\frac{-R_p^2}{2\delta R_p^2} \right) - \exp\left(\frac{-(x_{d_{max}} - R_p)^2}{2\delta R_p^2} \right) \right] + V_{sub} \tag{2.61}$$

and from Eq. (2.55)

$$V_{th} \triangleq V_{gs} \Big|_{\phi(0) = 2\phi_f} = 2\phi_f + V_{fb} + \frac{\varepsilon_s[-\phi'(0)]}{C_{ox}} \tag{2.62}$$

The algorithm to calculate V_{th} is as follows:

(i) Using Eq. (2.61) we obtain the second order equation

$$U = \frac{qN_a\, x_{d_{max}}^2}{2\varepsilon_s} + V \operatorname{erf} \frac{(x_{d_{max}} - R_p)}{\sqrt{2}\,\delta R_p} + \frac{2V}{\sqrt{2\pi}\,R_p}\, e^{-\frac{(x_{d_{max}} - R_p)^2}{2(\delta R_p^2)}} \tag{2.63}$$

where

$$U \triangleq 2\phi_f - \frac{qQR_p}{2\varepsilon_s} \operatorname{erf} \frac{R_p}{\sqrt{2}\,\delta R_p} - \frac{qQ\,\delta R_p}{\sqrt{2\pi}\,\varepsilon_s}\, e^{-\frac{R_p^2}{2\,\delta R_p^2}} \tag{2.64}$$

$$V \triangleq \frac{qQ\,R_p}{2\varepsilon_s} \tag{2.65}$$

This can be used to find $x_{d_{max}}$.

(ii) Using $x_{d_{max}}$ from (i), we find $\phi'(0)$ from Eq. (2.58)

$$\phi'(0) = \frac{qQ}{2\varepsilon_s}\left[\operatorname{erf}\frac{-R_p}{\sqrt{2}\,\delta R_p} - \operatorname{erf}\frac{(x_{d_{max}}-R_p)}{\sqrt{2}\,\delta R_p}\right] + \frac{qN_a}{\varepsilon_s}(x - x_{d_{max}}) \quad (2.66)$$

(iii) Using $\phi'(0)$ from (ii), we obtain V_{th}

$$V_{th} \triangleq V_{gs}\Big|_{\phi(0) = 2\phi_f} = 2\phi_f + V_{fb} + \frac{\varepsilon_s[-\phi'(0)]}{C_{ox}} \quad (2.67)$$

Note that we define threshold voltage as the gate voltage necessary to invert the surface by producing free carrier concentration equal to but of opposite type from that in the substrate at the surface. Unfortunately, the above equations are transcendental and numerical methods must be used to solve the problems. We list the algorithm here for further study, but check the validity of the equation by its default ($Q = 0$, or no implant) case.

If $Q = 0$, then all error functions disappear, and we obtain from Eq. (2.58)

$$\phi'(0) = \frac{-q}{\varepsilon_s} N_a x_{d_{max}} \quad (2.68)$$

and from Eq. (2.67)

$$V_{th} = 2\phi_f + V_{fb} + \frac{qN_a x_{d_{max}}}{C_{ox}} \tag{2.69}$$

where $x_{d_{max}}$ is found from Eq. (2.61)

$$\phi(0) = 2\phi_f = \frac{qN_a}{2\varepsilon_s} x_{d_{max}}^2 + V_{sub} \tag{2.70}$$

Combining Equations (2.68), (2.69), and (2.70), we find

$$V_{th} \triangleq V_{gs}\Big|_{\phi(0) = 2\phi_f} = V_{fb} + 2\phi_f + \frac{\sqrt{2\varepsilon_s qN_a(2\phi_f - V_{sub})}}{C_{ox}} \tag{2.71}$$

which is the conventional expression.

2.8 V_{th} Sensitivity versus Temperature of Ion-Implanted Enhancement MOS Devices

For a constant substrate concentration, $dV_T/dT < 0$ for N-channel devices, and $dV_T/dT > 0$ for P-channel devices. That is, devices are more "depletion type" with increasing temperature. The threshold voltage of a uniformly doped substrate is defined by (for N-channel devices)

$$V_{th} = \phi'_{MS} + 2\phi_f - \frac{Q_{SS}}{C_o} - \frac{1}{C_o}\sqrt{4\phi_f q\varepsilon_s N_a} \tag{2.72}$$

The last term, $-\frac{1}{C_o}\sqrt{4\phi_f\, q\varepsilon_s\, N_a}$, contributes a large amount to V_{th}. For an ion-implanted enhancement MOS device, N_a is not a constant but varies with the maximum depletion width $x_{d_{max}}$. Thus, it has different characteristics. An equivalent substrate concentration could be defined by a unity-weighted average

$$N_{eq} \cdot x_{d_{max}} = \int_{o}^{x_{d_{max}}} [P(x) + N_B]\, dx \qquad (2.73)$$

where P(x) is the distribution of the implant dose and N_B is the background concentration. This can be seen graphically in Fig. 2.8. When the two dashed regions have the same area, the above equation is satisfied.

In Fig. 2.8a, when the implant is shallow and $x_{d_{max}}$ penetrates deep into the implanted region, as temperature increases $2\phi_f$ decreases and $x_{d_{max}}$ decreases, this results in a higher N_{eq} as shown and a lower threshold sensitivity.

In Fig. 2.8b, the implant is deep, and $x_{d_{max}}$ only penetrates a small portion of the implanted region. Because $2\phi_f$ and $x_{d_{max}}$ decrease at high temperature, N_{eq} is lower and this results in higher threshold sensitivity.

Note that in above cases, the total depleted charge decreases at high temperature, and in all cases, V_{th} has a negative temperature coefficient. There is only minor difference among the threshold voltage sensitivities in all three cases (deep implant, shallow implant and uniform doping).

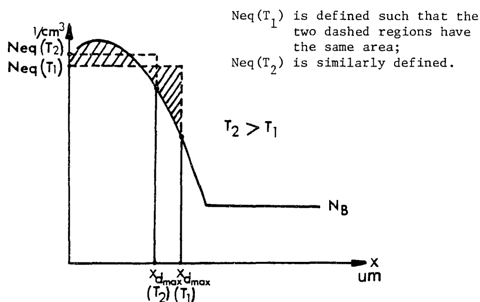


Fig. 2.8(a) A Shadow Threshold-Tailoring Implant with Decreased V_{th} Sensitivity versus Temperature

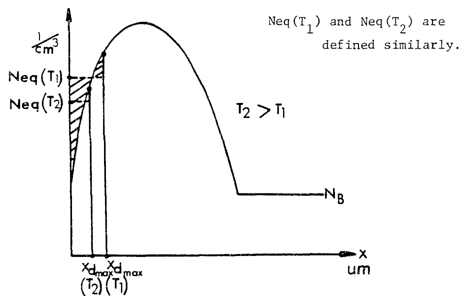


Fig. 2.8(b) A Deep Threshold-Tailoring Implant with Increased V_{th} Sensitivity versus Temperature

2.9 Summary

We have investigated various changes of threshold voltage as temperature increases including second order effects. For temperatures below 225°C, dV_{th}/dT remains nearly constant. The most important factors are the change of ϕ_f and surface state ionization effect. Threshold voltage variation for short/narrow channel devices, ion-implanted enhancement type devices have been studied. The geometry effect is less important at high temperature, but the percentage change of threshold voltage due to the geometry effect remains nearly constant at every temperature. Finally, methods to reduce threshold voltage sensitivity have been examined; these will increase the noise margin for digital circuits at high temperature. All of these are important considerations for high-temperature MOS circuit operation.

CHAPTER 3

LEAKAGE CURRENT AT HIGH TEMPERATURE

3.1 Introduction

Leakage current is a very important factor for high temperature MOS circuit operation. It will have an adverse effect on the operation of memory cells, gate protection diodes, and many other circuits. In this chapter we study different leakage current components as functions of temperature and the interaction of leakage current with channel current. Also, different ways to reduce leakage current density are examined.

3.2 Junction Leakage Current

Junction leakage current consists of two parts: generation-recombination (g-r) leakage current in the space-charge region, and diffusion leakage current [Grove, 1967].

3.2.1 Generation-Recombination Leakage Current in the Space-Charge Region

Electron-hole pairs generated in the space-charge region at drain-substrate, source-substrate, or P well-substrate (in CMOS) junctions will be swept out of the region by the high field existing there. This current is

$$I_{g-r} = J_{g-r} \cdot A_J = \frac{qn_i W_j}{2\tau_o} A_J \qquad (3.1)$$

where

τ_o = effective carrier lifetime within a reverse-biased space charge region (It is a function of temperature.)

W_j = junction depletion width

A_J = total junction area

This component will depend on the amount of reverse-bias voltage.

3.2.2 Diffusion Leakage Current

In addition to g-r leakage current in the space-charge region. electron-hole pairs generated in the neutral material within the average diffusion length of the space-charge region will be able to reach the edge of the space-charge region by diffusion and been swept toward P and N region for holes and electrons, respectively. The diffusion leakage current is given by

$$I_{diff} = J_{diff} A_J = \left[q \frac{n_{po}}{\tau_n} L_n + q \frac{p_{no}}{\tau_p} L_p \right] A_J \tag{3.2}$$

where

n_{po} = minority carrier in P region

p_{no} = minority carrier in N region

τ_n = average electron lifetime

τ_p = average hole lifetime

L_n = average diffusion leangth for electrons

L_p = average diffusion length for holes

A_J = total junction area

where the average diffusion lengths are

$$L_n = \sqrt{D_n \tau_n}$$

$$L_p = \sqrt{D_p \tau_p} \tag{3.3}$$

where

D_n = average diffusion coefficient of electrons

D_p = average diffusion coefficient of holes

Using the Einstein relation

$$D = \frac{kT}{q} u \tag{3.4}$$

where u is the carrier mobility, we obtain

$$J_{diff} = n_i^2 \sqrt{kTq} \left\{ \frac{1}{N_a} \sqrt{\frac{u_n}{\tau_n}} + \frac{1}{N_d} \sqrt{\frac{u_h}{\tau_h}} \right\} \tag{3.5}$$

where

N_a is the acceptor concentration in the P region

N_d is the donor concentration in the N region

where τ is calculated as

$$\tau = \frac{1}{\sigma_T \, v_{th} \, N_T} \tag{3.6}$$

with

N_T = Trap density

σ_T = cross section area for collision

V_{th} = thermal velocity of carriers = $\sqrt{\frac{3kT}{m^*}}$

m^* = effective masses of carriers

The mobilities are

$$u_n = (u_{n_L}^{-1} + u_{n_i}^{-1})^{-1}$$

$$u_p = (u_{p_L}^{-1} + u_{p_i}^{-1})^{-1} \tag{3.7}$$

The lattice mobility of electrons and holes are

$$u_{n_L} = 2.1 \times 10^5 \times T^{-2.5} \ m^2/\text{volt}\cdot\text{sec}$$

$$u_{p_L} = 2.3 \times 10^5 \times T^{-2.7} \ m^2/\text{volt}\cdot\text{sec} \tag{3.8}$$

and the impurity mobilities of electrons and holes are [Conwell and Weisskopf, 1950]

$$u_{n_i} = \frac{2^{7/2}\, \varepsilon_s^2\, (kT)^{3/2}}{\pi^{3/2}\, q^3\, m_n^{*1/2}\, N_I\, \ln\left[1 + \left(\frac{3\varepsilon_s\, kT}{q^2\, N_I^{1/3}}\right)^2\right]}$$

$$u_{pi} = \frac{2^{7/2}\ \varepsilon_s\ (kT)^{3/2}}{\pi^{3/2}\ q^3\ m_h^{*1/2}\ N_I\ \ell n \left[1 + \left(\frac{3\varepsilon_s\ kT}{q^2\ N_I^{1/3}}\right)^2\right]} \qquad (3.9)$$

m_n^* and m_h^* are effective masses of electrons and holes, respectively. N_I is the total impurity concentration within the average diffusion length of the junction. For example, in NMOS devices, the n^+P junction will have $N_I = N_a$ for electrons and $N_I = N_a + N_{S/D} = N_{S/D}$ for holes where $N_{S/D}$ is the source/drain impurity concentration.

Because of the n_i^2 dependence of diffusion leakage current, its effect will dominate g-r leakage current at high temperature. The latter has only n_i dependence.

. This is further illustrated in Fig. 3.1, where measured and calculated drain-substrate junction leakage current is plotted as a function of temperature, with the channel at cutoff. It is shown that the g-r component dominates at low temperature (region a) until the temperature reaches 135°C, above which diffusion leakage current dominates (region b).

Figure 3.2 again shows this effect where a test device is measured before and after the channel is turned on at 300°C. It is obvious that the drain leakage current is independent of drain bias voltage and is mainly due to the diffusion leakage current contribution.

Fig. 3.1 Measured and Theoretical Value of the Drain-Substrate Junction Leakage Current

$T_{ox} = 1000\text{Å}$ $Q_{SS}/q = 4 \times 10^{10}/cm^2$

$N_a = 3.4 \times 10^{16}/cm^3$ $A_j = 2.55 \times 10^{-5}/cm^2$

$V_D = 0$ volt $V_{sub} = -5$ volt

Channel cutoff

Simulation parameters are:

$\sigma_T = 1 \times 10^{-20}/m^2$ $N_T = 3.124 \times 10^{20}/m^3$

$\tau_n = \tau_n = 10 - 5\ (300/T)^{\frac{1}{2}}$ μsec in the depletion region

Also shows:

g-r dominant (a) and diffusion current dominant (b) region

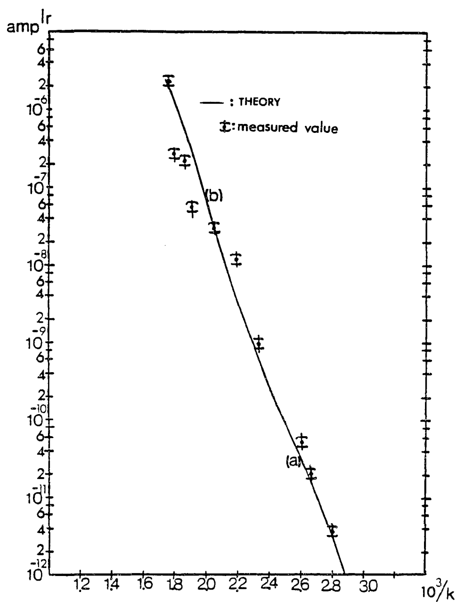
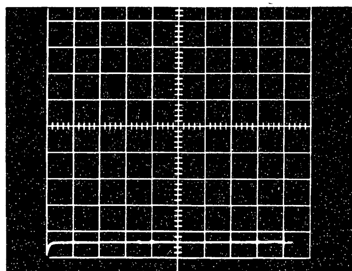


Fig. 3.1 Measured and Theoretical Value of the Drain-Substrate Junction Leakage Current



T_{ox} = 1000Å

$Q_{SS}/q = 4x10^{10}/cm^2$

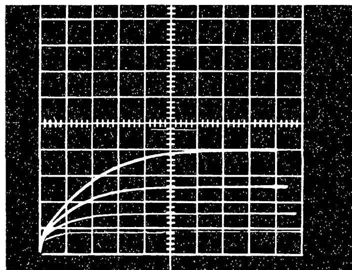
$N_a = 3.4x10^{16}/cm^3$

Al gate

H = 1 V/div

V = 0.1 mA/div

(a) Before channel is turned on,
showing the diffusion leakage current contribution



(b) With channel conducting V_g = 2, 4, 6 and 8 volts

Fig. 3.2 Drain Current Characteristics at 300°C

3.3 Subthreshold Current

In MOS devices, threshold voltage is usually taken to be the gate to substrate voltage at which the surface minority carrier concentration is equal to bulk impurity concentration. However, the drain current is not zero for $V_G < V_{th}$. Drain current flows along the channel whenever enough minority carriers exist at the surface channel. At room temperature, it is often assumed that subthreshold current starts when surface minority carrier concentration is equal to the intrinsic concentration ($\cong 1.45 \times 10^{10}/cm^3$ at room temperature). This is the weak inversion condition. Therefore, the subthreshold conduction region is defined as the region where the device is between the "flat band" condition and the "strong inversion" conditions. At very high temperature, because of the rapid increase of n_i even a small amount of band bending will result in a large number of surface minority carriers. The subthreshold conduction region is widened by this effect.

For a P-substrate material, band bending factor b is defined as: (before strong inversion)

$$\phi_s = v(y) + b\phi_f \tag{3.10}$$

where ϕ_s is the surface potential, v(y) is the bias voltage at point y along the channel, ϕ_f is the Fermi potential, b specifies the amount of bending. Three cases are important:

(i) b = 0 surface depletion begins

$$n = n_{p_o} \cong \frac{n_i^2}{N_a} - n_i \exp\left(\frac{-q\phi_f}{kT}\right)$$

(ii) $b = 1$ weak inversion begins

$$n \cong n_i$$

(iii) $b = 2$ strong inversion begins

$$n \cong N_a = n_i \exp \frac{q\phi_f}{kT} \tag{3.11}$$

or the minority carrier concentration at the surface is

$$n = n_i \exp \left[(b-1) \frac{q\phi_f}{KT} \right] \tag{3.12}$$

From the band bending factor b, the corresponding V_g can be obtained as

$$V_g = V_{fb} + b\phi_f + \frac{1}{C_{ox}} \sqrt{2\varepsilon_s q N_a \left(\frac{kT}{q} e^{\frac{-b\phi_f}{kT}} + b\phi_f - \frac{kT}{q} \right)} \qquad 0 \leq b \leq 2$$

[Troutman and Chakravarti, 1973]

and

$$V_g = V_{fb} + b\phi_f \qquad 0 > b \tag{3.13}$$

where $b < 0$ corresponds to the surface accumulation condition. The subthreshold current is [Troutman and Chakravarti, 1973]

$$I_{sub} = \left(\frac{W}{L}\right)\left(\frac{\varepsilon_s kT}{q^2 N_a}\right)^{\frac{1}{2}}\left(\frac{kT\mu}{q}\right) qn_i \left[1 - \exp\left(\frac{-qV_{DS}}{kT}\right)\right]$$
$$\cdot \left\{ \frac{\exp\left[(b-1)\frac{q\phi_f}{kT}\right]}{\left[2\left(\frac{q(V_S - V_{sub})}{kT} + b\frac{q\phi_f}{kT} - 1\right)\right]^{\frac{1}{2}}} - \frac{1}{\left[2\left(\frac{q(V_S - V_{sub})}{kT} + \frac{q\phi_f}{kT} - 1\right)\right]^{\frac{1}{2}}} \right\} \quad (3.14)$$

where W, L are channel width and length, V_{DS} is the drain source bias voltage, N_a is the substrate doping concentration, T is temperature in degrees Kelvin, μ is the carrier mobility along the channel.

At high temperature major concerns about the subthreshold current conduction are as follows:

(a) Because of the n_i dependence of surface minority carrier concentration I_{sub} increases rapidly with increasing temperature. Calculated data show that $I_{sub_{max}}$ (where b = 2) at 300°C is comparable to the diffusion leakage current component.

(b) Subthreshold current scales with W/L, not the area.

(c) Usually the subthreshold region can be defined as the region where $1 \leq b \leq 2$. This statement is true only for low or room temperature, where surface minority concentration is n_i at b = 1 (n_i = $1.45 \times 10^{10}/cm^3$ at 25°C). At very high temperature, surface minority carrier concentration exceeds $10^{12}/cm^3$ even in accumulation region as shown in Table 3.1. As a result, subthreshold conduction starts with b much less than 1. For a quantitative analysis, we define the start of

Table 3.1 Band Bending Factor b and Corresponding Surface Minority Carrier Concentration n_s at Different Temperatures

T °C	$n_s (1/cm^3)$				
	$b_1 = 1$	$b_1 = 0.5$	$b_1 = 0$	$b_1 = -1$	$b_1 = -2$
25	1.45×10^{10}	6.9×10^{6}	4.06×10^{3}	$\cong 0$	$\cong 0$
100	1.87×10^{12}	1.39×10^{10}	1.03×10^{8}	5.66×10^{3}	$\cong 0$
150	2.09×10^{13}	5.17×10^{11}	1.28×10^{10}	7.88×10^{6}	4.8×10^{3}
250	6.85×10^{14}	9.72×10^{13}	1.38×10^{13}	2.78×10^{11}	5.6×10^{9}
300	2.53×10^{15}	6.9×10^{14}	1.88×10^{14}	1.39×10^{13}	1.03×10^{12}

Device Parameters:

$T_{ox} = 1000Å$

$Q_{SS}/q = 4 \times 10^{10}/cm^2$

$N_a = 3.4 \times 10^{16}/cm^3$

Al Gate

subthreshold conduction as the point where the surface minority carrier concentration is $10^{12}/cm^3$, with band bending factor b_1 and corresponding gate voltage V_{g1}, and it ends when the channel is strongly inverted with $b_2 = 2$ and gate voltage $V_{g2} = V_{th}$. Calculated data of b_1, V_{g1}, and V_{g2} are shown in Table 3.2 for a particular device. It is obvious that the subthreshold conduction region ΔV_g has been increased with temperature until T equals 250°C and decreases thereafter, due to the nonlinear V_{th} drop at that temperature as discussed in Section 2.2.

This is shown in Fig. 3.3, where NMOS device subthreshold conduction characteristics are measured at 28°C, 157°C, and 217°C. The magnitude of the subthreshold current as well as the subthreshold conduction range ΔV_g increase with temperature. At 217°C, subthreshold current is much lower than the junction leakage current at the beginning of the subthreshold conduction; at the end of the subthreshold region (b = 2), the subthreshold current is not negligible compared to the junction leakage current. Increase of the subthreshold current is a serious problem for high temperature MOS circuit applications such as dynamic RAM, sampling circuits and switched capacitor filter circuit. Widening of the subthreshold conduction region reduces the noise margin NM of a logic circuit. In Section 5.4, it is shown that an NMOS inverter has a broader transition when it enters the subthreshold conduction region at 300°C.

3.4 Distributed Nature of the Channel Leakage Current

To this point, we have treated the leakage current as confined to the drain and source regions. But in the channel region, a diffusion

Table 3.2 Values of b_1, V_{g1}, b_2, V_{g2}, and ΔV_g at Different Temperatures

T (°C)	n_i 1/cm^3	Start of Subthreshold Region			End of Subthreshold Region			ΔV_g (volt)
		b_1	V_{g1} (volt)	n_{s1} (1/cm^3)	b_2	V_{g2} (volt)	n_{s2} (1/cm^3)	
25	1.45×10^{10}	1.30	1.48	10^{12}	2	2.33	3.4×10^{16}	0.85
100	1.87×10^{12}	0.94	0.82	10^{12}	2	2.03	3.4×10^{16}	1.21
250	6.85×10^{14}	-0.67	-1.03	10^{12}	2	1.29	3.4×10^{16}	2.32
300	2.53×10^{15}	-2.01	-1.11	10^{12}	2	0.98	3.4×10^{16}	2.09

Device Parameters:

$T_{ox} = 1000\text{Å}$

$Q_{SS}/q = 4 \times 10^{10}/\text{cm}^2$

$N_a = 3.4 \times 10^{16}/\text{cm}^3$

Al Gate

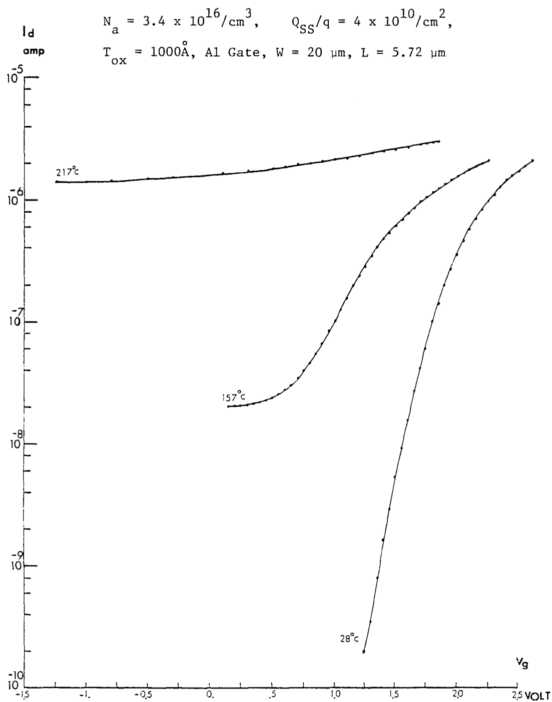


Fig. 3.3 Subthreshold Characteristics of an NMOS Device at Three Different Temperatures

leakage current flows out of the channel along its length, interacting with the normal channel current and causing an increase of drain current. It also alters the channel potential, and thus affects the device characteristics. In this section we analyze this interaction and its influence on the electric behavior of the device.

The reverse leakage current density is

$$J_L(T) = n_i^2(T)\sqrt{kTq}\left\{\frac{1}{N_a}\sqrt{\frac{u_n(T)}{\tau_n}} + \frac{1}{N_d}\sqrt{\frac{u_h(T)}{\tau_h}}\right\} + \frac{qn_i W_j}{2\tau_o} \tag{3.15}$$

At temperatures higher than 135°C, the g-r component of the junction leakage current can be neglected and a leakage current density independent of junction bias voltage exists.

$$J_L(T) = n_i^2(T)\sqrt{kTq}\left\{\frac{1}{N_a}\sqrt{\frac{u_n(T)}{\tau_n}} + \frac{1}{N_d}\sqrt{\frac{u_h(T)}{\tau_h}}\right\} \tag{3.16}$$

As shown in Fig. 3.4, for an N channel MOS device, the differential resistance of an element with length dy and width W is

$$dR = \frac{-dy}{Q_n(y)u_n W} \tag{3.17}$$

where $Q_n(y)$ is the charge density in the inversion layer per unit area along y. Rewriting Eq. (3.17), we obtain

$$\frac{dR}{dy} = \frac{-1}{Q_n(y)u_n W} \tag{3.18}$$

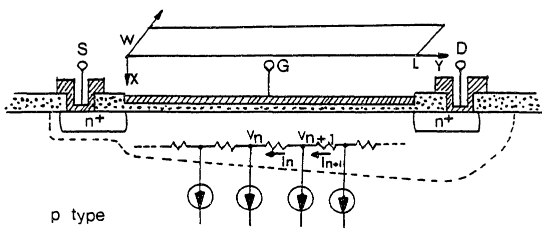


Fig. 3.4 Distributed Nature of Substrate-to-Channel Leakage Current Showing the Interaction Between Channel and Channel-to-Substrate Leakage Current

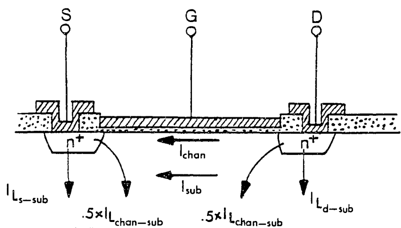


Fig. 3.5 A DC Model Showing Various Current Components by Taking Into Account the Interaction of Channel Current and Channel-to-Substrate Leakage Current

as the resistance per unit channel length at point y.

We now divide the region into infinitesimally small subregions as shown in Fig. 3.4. Then

$$I_{n+1} = I_n + A\Delta y$$

$$V_n = V_{n+1} - \left[\left(\frac{dR}{dy}\right)\Delta y\right] I_n \tag{3.19}$$

where $A = \frac{J_L(T)\ W\ L}{L} = J_L(T)\ W$ is the leakage current per unit channel length. If we let Δy be infinitesimally small, Equations (3.19) can be approximated by the following differential equations:

$$\frac{dI}{dy} = A$$

$$\frac{dV}{dy} = \left(\frac{dR}{dy}\right) I(y) \tag{3.20}$$

where we let $\Delta y \triangleq dy$ and $I_{n+1} - I_n \triangleq dI$. The boundary conditions are

$$V(y=L) = V_D$$

$$V(y=0) = V_S$$

$$I_d = \left(\frac{dV/dy}{dR/dy}\right)\Bigg|_{y=L} = I(y=L)$$

$$I_S = \left(\frac{dV/dy}{dR/dy}\right)\Bigg|_{y=0} = I(y=0) \tag{3.21}$$

From Eq. (3.20), we have

$$I = Ay + B \tag{3.22}$$

from Equations (3.20) and (3.21)

$$B = I_s$$

$$I_d = AL + I_s$$

and

$$I = Ay + I_s \tag{3.23}$$

Substituting this into Eq. (3.20), we obtain

$$\frac{dV}{dy} = \frac{dR}{dy} (Ay + I_s) \tag{3.24}$$

Substituting Eq. (3.18) into (3.24), we obtain

$$\frac{dV}{dy} = - \frac{1}{Q_n(y)\mu_n W} (Ay + I_s) \tag{3.25}$$

It is well known that in the inversion layer, the charge density is

$$Q_n(y) = (-C_{ox})[V_g - V_{fb} - 2\phi_f - V(y)]$$

$$+ [2\varepsilon_s qN_a(2\phi_f - V_{sub} + V(y)]^{\frac{1}{2}} \tag{3.26}$$

so

$$- \frac{dV}{dy} = \frac{Ay + I_s}{\mu_n W(- C_{ox})[V_g - V_{fb} - 2\phi_f - V(y)] + \mu_n W[2\varepsilon_s q N_a (2\phi_f - V_{sub} + V(y)]^{\frac{1}{2}}} \tag{3.27}$$

and the boundary conditions are

$$V(0) = V_S$$

$$V(L) = V_D \tag{3.28}$$

Separating the variables y and V, we find

$$\int_0^L (Ay + I_s)dy = \int_{V_S}^{V_D} \left\{ \mu_n W C_{ox}[V_g - V_{fb} - 2\phi_f - V(y)] - \mu_n W[2\varepsilon_s q N_a (2\phi_f - V_{sub} + V(y))]^{\frac{1}{2}} \right\} dV \tag{3.29}$$

The integration at the right side is the channel current expression without taking into account the interaction. We define it as $I_{chan} \cdot L$ Then, we have

$$I_{chan} = \frac{\mu_n W}{L} C_{ox} \left\{ (V_D - V_S) \left[V_g - V_{fb} - 2\phi_f - \frac{(V_D - V_S)}{2} \right] - \frac{2}{3} (2\varepsilon_s q N_a)^{\frac{1}{2}} \frac{1}{C_{ox}} \left[(V_D - V_{sub} + 2\phi_f)^{3/2} - (V_S - V_{sub} + 2\phi_f)^{3/2} \right] \right\} \tag{3.30}$$

$$I_s = I_{chan} - \frac{1}{2} A L$$

$$I_d = I_{chan} + \frac{1}{2} A L \tag{3.31}$$

From the above analysis we note the following:

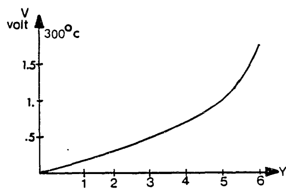
(i) The channel-substrate leakage current increases the drain current and decreases the source current. We can divide this channel-to-substrate leakage current into two equal parts; one adds to the drain current and one subtracts from the source current, as shown in Fig. 3.5.

(ii) The above analysis can be applied to P channel MOS devices, but the directions of currents I_n, I_{n+1}, and J_L have to be reversed.

(iii) In the above analysis we neglect the short-channel effect, but it is not difficult to incorporate this effect if a trapezoidal depletion region is used.

(iv) In this analysis, we neglect the interaction between the channel-substrate leakage current and subthreshold current. Since the magnitude of the subthreshold current is less than I_{chan} for most of the temperature range of interest, the interaction between subthreshold current and channel-to-substrate leakage current is less important than that between the channel current and channel-to-substrate leakage current.

Solving the differential Equation (3.27) by the DAREP simulation program developed at The University of Arizona, we can plot the electrical potential, longitudinal electric field and the surface charge density along the channel. Figures 3.6, 3.7, and 3.8 show the results



(a) Voltage Distribution Along the Channel

$N_a = 3.4 \times 10^{16}/cm^3$

Al gate

$\frac{W}{L} = \frac{20\ \mu m}{6\ \mu m}$

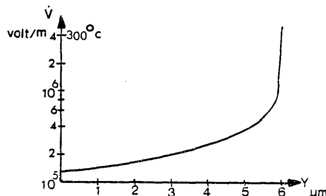
$V_S = 0$ volt

$V_g = 6$ volt

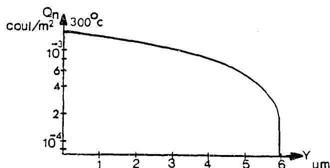
$V_D = 1.8$ volt $= V_{Dsat}$

T = 300°C

$J_L = 9.7301 \times 10^{-1}$ A/cm^2



(b) Electric Field Distribution Along the Channel



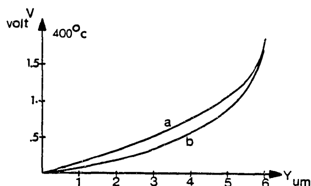
(c) Surface Charge Density Along the Channel

Fig. 3.6 One-Dimensional DARE P Solution of Equation (3.27) with and without Taking into Account the Interaction of Channel Current and Channel-to-Substrate Leakage Current

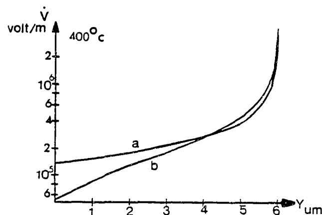
Device parameters as before except:

$V_D = V_{D_{sat}} = 2$ volt

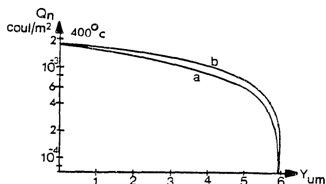
$J_L = 5.7942 \times 10^1$ A/cm^2



(a) Voltage Distribution Along the Channel



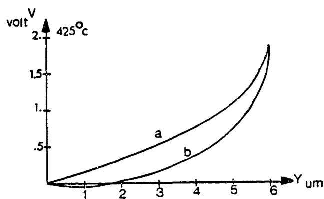
(b) Electric Field Distribution Along the Channel



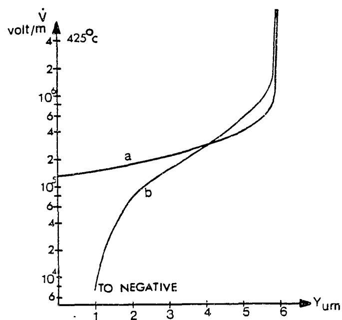
(c) Surface Charge Density Along the Channel

a - without taking into account leakage contribution
b - taking into account leakage contribution

Fig. 3.7 Voltage, Electric Field, and Surface Charge Density Along the Channel at 400°C.

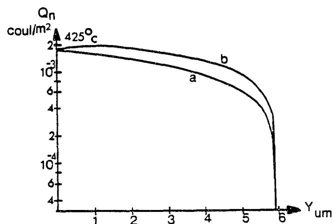


(a) Voltage Distribution Along the Channel



(b) Electric Field Distribution Along the Channel

Fig. 3.8 Voltage, Electric Field and Surface Charge Density Along the Channel at 425°C



(c) Surface Charge Density Along the Channel

Device parameters are as before except:

$V_D = V_{D_{sat}} = 2.2$ volt

$J_L = 1.35 \times 10^2$ A/cm^2

a - without taking into account leakage contribution

b - taking into account leakage contribution

Fig. 3.8 (Continued)

at three different temperatures: 300°C, 400°C, and 425°C. The leakage current density data are extrapolated from the leakage current measurement in Section 3.2. The following should be noticed:

(i) At 300°C, J_L W L < I_{chan}; the curves with and without taking into account of the interaction almost coincide.

(ii) At 400°C, J_L W L < I_{chan}, but the two components are comparable. Then:

(a) Due to the injection of electrons from the substrate, the electric potential is lowered.

(b) The electric field along the channel is modulated by the injected electrons. Its magnitude is lowered at the source end and increased at the drain end. This is because the electric field lines have to terminate at negatively charged particles. As shown in Fig. 3.9, field lines terminated at those extra injected electrons will cause an increase of the electric field at the drain end and a decrease of the field at the source end.

(c) As a result of the electron injection, the surface charge density per unit area increases under the channel.

(iii) At 425°C, J_L W L > I_{chan}. Due to the large amount of injected electrons, the potential will be negative and the electric field reverses its direction near the source end. Surface charge density per unit area is further increased.

The preceding analysis shows that for operation at temperatures much lower than 400°C, it is a reasonable assumption that electric

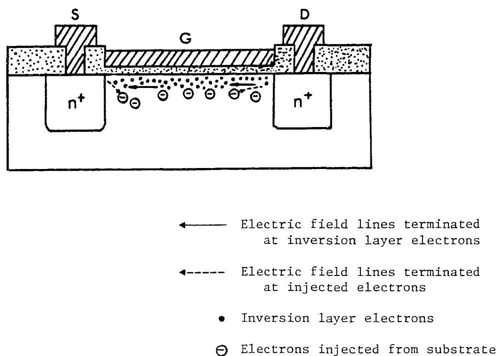


Fig. 3.9 An NMOS at Very High Temperature Showing the Influence of Injected Electrons from the Substrate

potential, electric field and surface charge density do not vary very much along the channel. As we will see in Chapter 7, this is the basis for high temperature scaling analysis.

3.5 Gate Leakage Current

The oxide conductivity as function of temperature can be approximated from Kingery [Kingery, Bowen and Uhlmann, 1976] as follows:

$$\log \sigma = -15.4 + 4.48 \left(2.9 - \frac{1000}{T}\right) \qquad T \geq 416°K$$

$$\log \sigma = -18 + 3.71 \left(3.1 - \frac{1000}{T}\right) \qquad T \leq 416°K$$

(3.32)

calculated value of σ at 25°C and 300°C are $1.115 \times 10^{-19} (\Omega\text{-cm})^{-1}$ and $1.26 \times 10^{-12} (\Omega\text{-cm})^{-1}$, respectively.

The potential across the oxide varies from $V_g + \phi'_{MS} - (2\phi_f + V_S)$ at the source to $V_g + \phi'_{MS} - (2\phi_f + V_D)$ at the drain. In the worst case ($V_D = V_S = 0$), the current density across the oxide is

$$J_g = \sigma \frac{1}{T_{ox}} (V_g + \phi'_{MS} - 2\phi_f) \qquad (3.33)$$

For a gate oxide field of 10^5 volt/cm, J_g at 25°C and 300°C are $1.15 \times 10^{-14} A/cm^2$ and $1.26 \times 10^{-7} A/cm^2$, respectively.

Gate leakage current at high temperature reduces the input impedance of MOS devices. It has an adverse effect if the gate is connected to a high impedance node or a capacitor.

3.6 Methods to Reduce Leakage Current Density at High Temperature

3.6.1 Guard Ring Collection of Minority Carriers

Diffusion leakage current originates from thermally generated minority carriers in the substrate of MOS devices. The amount of minority carriers diffused into source, drain and channel region can be reduced if proper guard ring structures are implemented in the design. One example is shown in Fig. 3.10. If V_c is connected to the most positive power supply, minority carriers (electrons) generated in the P substrate are collected at the n^+ guard ring and the buried-layer diffusion. As a consequence, only minority carriers (electrons) generated in between source/drain diffusion and n^+ guard ring can be collected at source, drain or channel regions of the device. The magnitude of diffusion leakage current component can be much reduced. For example, if L_n (300°C) = 60μ, and the distance between source/drain and n^+ buried layer is 2μ, more than 60 times reduction of diffusion leakage current density is expected. However, three more steps are involved in the fabrication process: buried layer, epitaxial growth and an extra n^+ diffusion. Also, the area per device is increased. For a device with (W/L) = 1/2 and 5μ design rule, the area increase is approximately 5 times if every device is surrounded by the guard ring.

3.6.2 Epitaxial MOS

At temperatures above 135°C, the diffusion component is the major component of leakage currents, and it is inversely proportional to substrate doping concentration. We therefore can reduce the

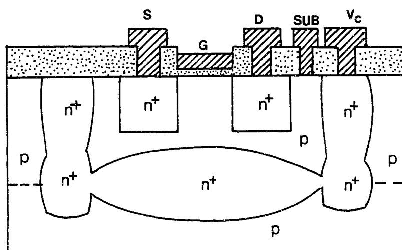


Fig. 3.10 Guard Ring Collection of Minority Carriers in the Substrate

diffusion leakage current density by increasing the substrate doping concentration. However, we will obtain a high threshold voltage which is impractical in circuit design. To compromise, we use a highly doped substrate (P^+ for leakage current reduction) and grow a thin epitaxial layer (lightly doped for threshold voltage control). This process, epitaxial MOS, has been used in the fabrication of MOS memory circuits. Its low resistance in the substrate provides an electrically short path for radiation damage, and therefore reduces the soft error rate in memory circuits. In this section we will show that epitaxial MOS is also desirable in high temperature MOS circuits.

The diffusion component of leakage current is

$$J_{diff} = \frac{qn_i^2}{N_a \tau_n} L_n \tag{3.34}$$

The diffusion length L_n for minority carriers is in the range of 60μ to 150μ depending on wafer quality, temperature, and process control. From our leakage current measurement, L_n of 140μ and 60μ are obtained at 25°C and 300°C, respectively. Since minority carriers generated in the neutral material within a diffusion length of space charge region will be collected at source/drain node, if L_n is much larger than the epitaxial layer thickness, the diffusion leakage current will be reduced if P^+ substrate is used.

As shown in Fig. 3.11a, we assume lateral diffusion leakage current is negligible and L_n is unchanged by the epitaxial growth. Then

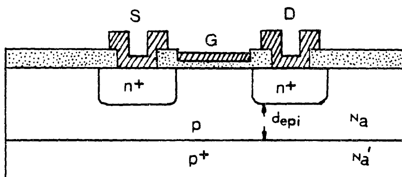
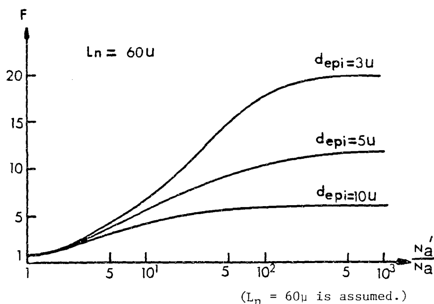


Fig. 3.11(a) An Epitaxial MOS Device



(L_n = 60μ is assumed.)

Fig. 3.11(b) Improvement Factor of the Diffusion Leakage Current Density for Epitaxial MOS Devices

$$J_{diff}\text{ (without epi)} = \frac{qn_i^2 L_n}{N_a \tau_n} \tag{3.35}$$

$$J_{diff}\text{ (with epi)} = \frac{qn_i^2 L_n}{N_{a_{eq}} \tau_n}$$

$$= \frac{qn_i^2 d_{epi}}{N_a \tau_{n_{epi}}} + \frac{qn_i^2 (L_n - d_{epi})}{N'_a \tau_{n_{sub}}} \tag{3.36}$$

If it is assumed that τ_n is the same in substrate and epitaxial layer, then Eq. (3.34) becomes

$$J_{diff}\text{ (with epi)} = \frac{qn_i^2 d_{epi}}{N_a \tau_n} \left[1 + \frac{N_a}{N'_a} \frac{(L_n - d_{epi})}{d_{epi}} \right]$$

$$= \frac{qn_i^2 d_{epi}}{N_a \tau_n} \left[1 - \frac{N_a}{N'_a} + \frac{N_a}{N'_a} \frac{L_n}{d_{epi}} \right] \tag{3.37}$$

The improvement ratio is

$$F \triangleq \frac{J_{diff}\text{ (without epi)}}{J_{diff}\text{ (with epi)}} = \frac{(L_n/d_{epi})}{1 - \frac{N_a}{N'_a} + \frac{N_a}{N'_a} \frac{L_n}{d_{epi}}} \tag{3.38}$$

Figure 3.11b shows the relationship between F and N'_a/N_a ratio with d_{epi} as parameters (L_n = 60μ is assumed). For better improvement of the diffusion leakage current, higher N'_a and lower d_{epi} should be used.

Note that in the above analysis, we assumed abrupt boundary between P^+ and P layers (neglect the up diffusion of P^+ dopant in epitaxial growth and later processes). In reality, a nonuniform doping concentration along the depth will result, and a slightly higher value of F will be obtained.

The epi MOS technique reduces the minority carrier concentration in the P^+ region by the ratio of N_a'/N_a, and therefore reduces the diffusion leakage current density. For example, assume $d_{epi} = 2\mu$, $N_a'/N_a = 100$; the resulting reduction ratio of diffusion leakage current density is 23.3. This is not as effective as guard ring technique, but only one extra epi growth process is needed and no additional area is required. It is therefore more cost-effective.

3.6.3 LOCOS Isolation

This process uses a silicon nitride layer to mask against thermal oxidation of silicon in a protected silicon area. The oxide is then etched and another oxide layer is grown to the level of the protected silicon. In between the oxide regions, MOS devices can be fabricated. This method eliminates side-edge collection of minority carriers for source and drain junctions, but leaves the bottom side unchanged. Little improvement can be expected from this process. Figure 3.12 shows the structure of a LOCOS isolated epi MOS device.

3.6.4 Silicon on Insulator Devices

Examples of these devices are silicon on sapphire devices (SOS), SIMOS devices [Izumi, Doken, and Ariyoshi, 1978] and

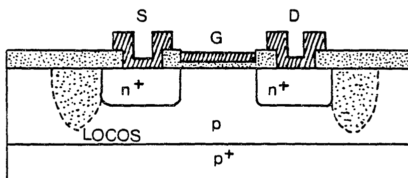


Fig. 3.12 LOCOS Epi MOS Devices

lateral-seeding epitaxial MOS devices [Lam, Sobezak, Pinizzotto, and Tasch, 1980]. Theoretically, if a thin layer of epitaxial silicon is grown on top of an insulator, the effective length of minority carriers collected in the vertical direction is much reduced ($L_{n_{eff}} \leq L_{epi}$). In practice, because of the difference in crystal lattice spacing and thermal expansion coefficient between the silicon and insulator, this process results in poorer quality of epitaxial layer with high leakage current and low minority carrier lifetime compared to bulk silicon devices. Leakage current ratio of 10^3 to 1 at room temperature for SOS and bulk MOS device has been reported [Micheletti, 1980]. Also, a large number of interface states exists between the silicon and insulator boundary due to the stress in epitaxial growth. Problems such as back-gate bias effect must be considered; this effect is more serious at high temperature. Unless epi quality can be improved dramatically in the future, these devices are not suitable for very high temperature operation.

Recently, CMOS logic circuits based on a dielectric isolation (DI) process which allows circuit operation up to 350°C have been reported [Beasom, Moore, Mohammed and Draper, 1981], but the many steps involved in this process make mass production extremely costly.

3.6.5 Gettering Effect

This process, usually the last step in the fabrication of devices, introduces strains into silicon lattice where segregation of impurities (metallic, oxygen, carbon, etc.) will occur. Different types of gettering process could be used. They include:

(a) back surface mechanical damage;

(b) back surface laser induced damage;

(c) intrinsic gettering by good control of the depth of precipitation in the silicon such that impurities reside in the deep non-active region of the silicon wafer, serving as gettering sites for impurities in the near surface active regions;

(d) back surface ion implanted gettering.

Among them, ion-implant gettering is easier to incorporate in the fabrication processes as the last step. The effect of impurity gettering depends on the ion type, with $^{40}Ar^{+}$ being the most favorable, the dosage and the subsequent annealing temperature. Results yielding an order of magnitude reduction of $n^{+}P$ junction leakage current at room temperature [Tice and Geipel, 1979], and 50 times increase of minority carrier lifetime [Ryssel, Schmiedt, and Kranz, 1977], are reported. However, relations between gettering parameters like ion dose, annealing temperature, implant energy and the optimum carrier lifetime must be studied.

3.7 Summary

In this chapter, we have examined different leakage current components encountered in high temperature MOS devices. At room or medium temperature g-r and subthreshold leakage currents that have n_i dependence dominate. At temperatures higher than 135°C (typically), diffusion leakage current which has n_i^2 dependence is the major concern.

The subthreshold conduction region is widened at high temperature and this affects the inverter characteristics. The interaction of channel current and channel-to-substrate leakage current increases the current flowing out of the drain diffusion and decreases the current flowing into the source diffusion. It also affects the electric field and charge density distribution of the device at very high temperature ($\geq$ 400°C). Finally, methods that can be used to reduce leakage current density have been suggested. Among them, guard-ring collection, epitaxial MOS and ion-implant gettering processes seem to be most promising.

CHAPTER 4

ADDITIONAL FACTORS AFFECTING HIGH TEMPERATURE OPERATION OF MOS DEVICES

4.1 Introduction

In this chapter we investigate factors which are less important than threshold voltage and leakage current, but which also influence high temperature behavior of MOS devices. Avalanche multiplication of leakage current and channel current is studied first. We show that although the thermally generated leakage current is large, multiplication factor decreases at high temperature. This is due to the lattice vibration (phonon emission) effect; the ionization rate of electrons and holes decrease as temperature increases. We then study the hot-electron effect. At large applied voltages, electrons flowing from the source to the drain of an NMOS device may gain sufficient energy from the high field region near the drain to be emitted into the gate insulator near the drain junction. A non-equilibrium condition exists where electron equivalent temperature far exceeds the lattice temperature. These so-called hot electrons can also come from the channel-substrate depletion region if high substrate bias voltage and high leakage current are present. As a result of the phonon emission effect, the probability of an electron or hole in the silicon being trapped at the silicon-silicon dioxide decreases with increasing temperature.

Finally, reliability concerns about the aluminum metallization at high temperature and high current density conditions are examined.

4.2 Avalanche Multiplication and Breakdown Mechanism

Avalanche multiplication and breakdown phenomena in MOS devices can be divided into three categories:

(i) thermally generated carrier multiplication due to the high field present in the source/drain to substrate junction;

(ii) channel current multiplication due to the large electric field in the vicinity of the drain junction under the gate; and

(iii) gate control surface breakdown due to extremely high gate bias.

Cause (iii) is only important when unusually high gate bias is used, and will not be discussed here.

In general, at high temperature, electrons or holes lose their energy to lattice vibrations; the optical phonon mean free path is decreased; more energy is absorbed from the field by phonon emission and less energy is available to cause impact ionization. But the seed current which is the current across the junction before avalanche multiplication takes place is higher at high temperature.

4.2.1 Case (i): Junction Multiplication and Breakdown

This occurs when the channel is off or when the channel current is very small. The seed current of avalanche multiplication is the reverse saturation current (g-r leakage plus diffusion leakage current).

$$I_{d_1} = MI_a \tag{4.1}$$

I_{d_1}, I_a are drain and seed current respectively, as shown in Fig. 4.1, and M is the multiplication factor

$$M^{-1} = 1 - \int_0^W \alpha_n \exp\left[-\int_0^x (\alpha_n - \alpha_p)dx'\right]dx \tag{4.2}$$

α_n, α_p are the ionization rates of electrons and holes, respectively, and W is the junction depletion width. Using abrupt P-N junction approximation and making the following change of variable

$$x = \frac{\varepsilon_s}{qN_a} E \tag{4.3}$$

we obtain

$$M^{-1} = 1 - \int_0^{E_{max}} \alpha_n \exp\left[-\int_0^E (\alpha_n - \alpha_p)dE'\right]dE \tag{4.4}$$

where x is the distance measured from the depletion edge at the substrate side. E is the electric field in the depletion region. The maximum field for this type of multiplication and breakdown is

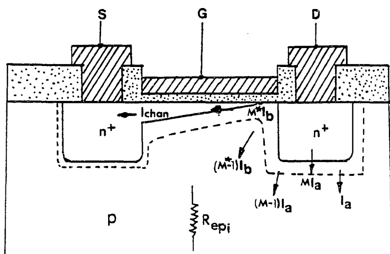


Fig. 4.1 Current Multiplication of NMOS Devices at High Temperature

$$E_{max} = \left[\frac{2qN_a(V_D + V_{bi} - V_{sub})}{\varepsilon_s}\right]^{1/2} \qquad (4.5)$$

V_{bi} is the built-in voltage of drain-substrate junction. At breakdown point $M = \infty$ and

$$\int_o^{E_{max}} \alpha_n \exp\left[-\int_o^E (\alpha_n - \alpha_p)dE'\right] dE = 1 \qquad (4.6)$$

From this, breakdown voltage is found to be

$$V_{BV} = \frac{\varepsilon_s E_{max}^2}{2qN_a} \qquad (4.7)$$

The dependence of α_n and α_p with electric field and temperature are related through the complicated relation [Crowell and Sze, 1966]

$$\alpha_n = \frac{1}{\lambda_n} \exp\left[a_o + a_1 \beta_n + a_2 \beta_n^2\right]$$

$$\alpha_p = \frac{1}{\lambda_p} \exp\left[b_o + b_1 \beta_p + b_2 \beta_p^2\right] \qquad (4.8)$$

where

$$\lambda_n \triangleq \lambda_{on} \tanh \frac{\varepsilon_r}{2kT}$$

$\triangleq$ mean free path of electron for optical phonon generation

$\lambda_{on} \triangleq$ limiting phonon mean free path as T approaches zero

$\varepsilon_r \triangleq$ Raman optical phonon energy

$k \triangleq$ Boltzman constant

$T \triangleq$ degree temperature in Kelvin

$\lambda_p \triangleq \lambda_{op} \tanh \frac{\varepsilon_r}{2kT}$

$\triangleq$ mean free path of hole for optical phonon generation

$\lambda_{op} \triangleq$ limiting phonon mean free path as T approaches zero

$\beta_n \triangleq \frac{\varepsilon_i}{qE\lambda_n}$

$\beta_p \triangleq \frac{\varepsilon_i}{qE\lambda_n}$

ε_i = the ionization energy = 1.5 Eg (Eg is the band gap energy)

$q \triangleq$ electron charge

$E \triangleq$ electric field

$a_o \triangleq -757\ r^2 - 75.5r + 1.92 = b_o$

$a_1 \triangleq 46\ r^2 - 11.9r - 1.75 \times 10^{-2} = b_1$

$a_2 \triangleq 11.5\ r^2 - 1.17r + 39 \times 10^{-4} = b_2$

$r \triangleq \frac{\langle\varepsilon_r\rangle}{\varepsilon_i}$

$\langle\varepsilon_r\rangle \triangleq$ average energy loss per collision with phonon

$$\triangleq \varepsilon_r \tanh \frac{\varepsilon r}{2kT} \tag{4.9}$$

Experimentally verified values of $\lambda_{on} = 76$ Å, $\lambda_{op} = 47$ Å, and ε_r = 0.063 eV are used.

Although I_a is higher as temperature increases, α_n and α_p drop rapidly with increasing temperature because of the decrease of mean free path of carriers for optical phonon generation. As a consequence, overall breakdown voltage is larger as temperature increases. From Sze's graphical data, the following relation is derived

$$\frac{V_{BV}(T)}{V_{BV}(300°K)} = 1 + 5.416 \times 10^{-4} \; (T-300)(N_a)^{-0.31} \tag{4.10}$$

where N_a is the substrate concentration in atom/cm^3. T is in degrees Kelvin.

4.2.2 Case (ii): Channel and Channel-to-Substrate Leakage Current Multiplication

Because of the high field present at the drain-substrate depletion region near the silicon surface, channel current which flows along the surface, as well as the channel-to-substrate leakage current flowing perpendicularly to surface into substrate, will be multiplied. These primary electrons can gain enough energy to cause impact ionization of hole-electron pairs. The secondary electrons are collected at the drain; secondary holes are prevented from entering the drain by the built-in voltage and substrate bias voltage, and are collected at the substrate.

The drain current is

$$I_{d2} = M^* \left(I_{chan} + I_{Lchan-sub}\Big|_{at\ drain} \right) = M^* I_b \tag{4.11}$$

where I_{chan} is the channel current and

$$I_{Lchan-sub}\Big|_{at\ drain} = \frac{1}{2} \cdot A_{gate} \cdot J_L \tag{4.12}$$

A_{gate} is the gate area; J_L is the leakage current density. The calculation of M^* is similar to Case (i) but the electric field expression is different:

$$E_t = E_1 + E_2 + E_3 \tag{4.13}$$

where E_t is the total electric field.

E_1 is the maximum electric field in the lateral direction (the device is assumed to be in the saturation region)

$$E_1 = \left[\frac{2qN_a (V_D - V_{Dsat})}{\varepsilon_s} \right]^{\frac{1}{2}} \tag{4.14}$$

V_{Dsat} is the drain saturation voltage

E_2 is the fringing field in the drain depletion region due to voltage between gate and drain

$$E_2 = 0.2 \frac{\varepsilon_{ox}}{\varepsilon_s} \frac{(V_D - V_g - Q_{SS}/C_{ox})}{T_{ox}} \tag{4.15}$$

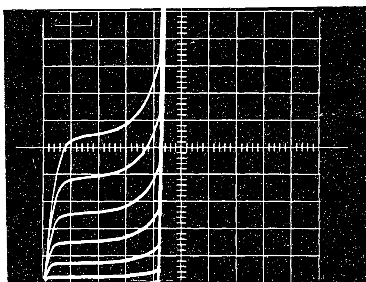
E_3 is the fringing field in the drain depletion region due to voltage between gate and the saturation point

$$E_3 = 0.6 \frac{\varepsilon_{ox}}{\varepsilon_s} \frac{(V_g + Q_{SS}/C_{ox} - V_{Dsat})}{T_{ox}} \tag{4.16}$$

V_g is the gate voltage, ε_s and ε_{ox} are permeability of silicon and silicon dioxide respectively; Q_{SS} is the surface charge density. Here, 0.2 and 0.6 are chosen for proper simulation of their contribution [Frohman-Bentchkowsky and Grove, 1969]. Knowing V_D, we can find M^*.

Except for devices with very thin gate oxide, breakdown voltage is determined by the drain-substrate junction and M is higher than M^*. (M^* is gate voltage dependent and M is not.) Figure 4.2 shows the multiplication and breakdown characteristics of an NMOS device at 23°C and 250°C. From this, it is seen that the multiplication factor decreases and breakdown voltage increases at high temperature.

Note that the above analyses assume that current flows directly beneath the surface, and the electric field in the depletion region under the surface will determine M^* and V_{BV}. This is not true for small-dimension scaled-down devices. For more accurate analysis, a two-dimensional numerical program must be used to solve Poisson's

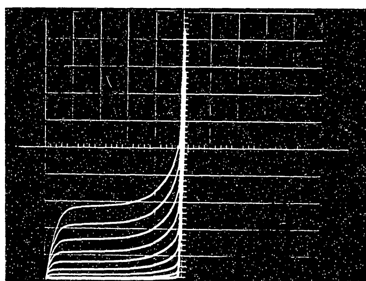


V_g: 0 to 8 volts in 1 volt/step

H: 5v/div

V: 0.1 mA/div

23°C



V_g: 0 to 8 volts in 1 volt/step

H: 5v/div

V: 0.1 mA/div

250°C

$N_a = 3.4 \times 10^{16}/cm^3$, $Q_{SS}/q = 4 \times 10^{10}/cm^2$, $T_{ox} = 1000Å$, Al gate, $W = 20\ \mu m$, $L = 5.72\ \mu m$.

Fig. 4.2 Current Multiplication and Breakdown Characteristics of an NMOS Device at 23°C and 250°C

equation and the current continuity equation with avalanche multiplication factor calculated along the high field path [Toyabe, Yamaguchi, Asai, and Moch, 1978].

For MOS devices operated at high temperatures, M, M^*, and V_{BV} are not of major importance. However, the leakage currents I_L and 0.5 $I_{L_{chan-sub}}$ increase rapidly with temperature. These currents, when multiplied, will have an effect on the source-substrate junction. This can be explained as follows. Holes will flow through the substrate, developing a voltage drop across the substrate resistance (of the order of 20K for typical NMOS devices), thereby lowering the source-substrate junction potential and causing more electrons to flow out of the source. This positive feedback can cause a "soft" or negative resistance breakdown having low sustaining voltage. Because α_n is higher than α_p, and substrate resistance is higher for NMOS than PMOS devices, the former are more susceptible to this soft breakdown than PMOS [Toyabe, et al., 1978].

4.3 Hot-Electron Effect

Advances in MOS VLSI integrated circuits require small devices, higher substrate concentration, and thinner gate oxide. These all increase the hot-electron effect for MOS devices. There are three sources of hot electrons: (1) P-N junction avalanche plasma, (2) channel hot electrons, and (3) substrate hot electrons. Among these three sources, No. 3 is important at high temperature.

4.3.1 P-N Junction Avalanche Plasma

When the substrate is reverse biased and the gate is biased to cause surface accumulation, avalanche multiplication will occur and a plasma will form in the junction region because of the field crowding effect near the surface at the corners of the source-substrate and drain-substrate junctions. A large enough substrate bias voltage will give electrons enough energy to be injected into the SiO_2 layer near source and drain junctions. This is not a problem for normal MOS circuit operation because we will usually have the silicon surface depleted or inverted.

4.3.2 Channel Hot-Electron Effect

Channel hot-electron effect occur when high electric field exists at the drain-substrate junction depletion region. Electrons which drift across the depletion region will have a high probability of experiencing a large energy change before lattice collisions occur. These electrons which obtain sufficient energy to overcome the Si-SiO_2 energy barrier will be injected into the gate material, creating gate current. Also, some of these electrons are trapped in the insulator, thus causing threshold instability and transconductance degradation of MOS devices. At high temperature, usually the channel current decreases; also, the probability of these electrons being trapped in the SiO_2 decreases because of the decrease of phonon mean free path length. Physically speaking, electrons will travel a shorter distance before they have lattice collision. Therefore, they gain less energy and have less chance to overcome the potential

barrier [Eitan and Frohman-Bentchkowsky, 1981]. Thus, the channel hot-electron effect diminishes at high temperature.

4.3.3 Substrate Hot-Electron Effect

This mechanism is similar to the channel hot-electron effect. In this case, electrons which are thermally generated in the depletion region (g-r leakage current) and electrons which diffuse from the bulk neutral region of the substrate (diffusion leakage current) are the source current. For high temperature MOS circuit operation, these leakage components increase rapidly. Although the emission probability decreases as temperature increases, the product of the leakage current (g-r plus diffusion) and the emission probability may be an important concern especially when high doping concentration, high substrate bias and thinner gate oxide are used.

The emission current density is the product of the leakage current density and emission probability.

$$J_{em} = [J_{L_{g-r}} + J_{L_{diff}}] \, P \tag{4.17}$$

To calculate P, the so-called "lucky electron model" can be used [Ning, Osburn, and Yu, 1977] to explain the measured results with good accuracy. It uses the following assumptions:

(a) An electron can be emitted if it reaches the interface with energy higher than the interface barrier energy

$$\Phi_{Ba} = 3.1 \text{ eV} - \beta \, E_{ox}^{1/2} - \alpha \, E_{ox}^{2/3} \tag{4.18}$$

with

$$\beta = 2.6 \times 10^{-4} \text{ e } (V \cdot cm)^{1/2}$$

$$\alpha = 1.0 \times 10^{-5} \text{ e } (V \cdot cm)^{1/3} \tag{4.19}$$

β and α are included for Schottky-lowering and tunnelling effect, respectively. E_{ox} is the oxide electric field.

(b) These electrons must be "lucky" enough to fall along a potential drop equal to Φ_{Ba}/q without any collision with the lattice before reaching the interface. This potential sets up a critical distance d_{cr} between the top of the interface barrier and the conduction band edge. For a uniformly doped substrate with N_a as the concentration, d_{cr} is found from

$$d_{cr} = \left[\sqrt{\frac{2\varepsilon_s}{qN_a}} \sqrt{\phi_s - V_{sub}} - \sqrt{\phi_s - V_{sub} - \Phi_{Ba}/q} \right] \tag{4.20}$$

where ϕ_s is the surface potential (= 2 ϕ_f at strong inversion). An electron must pass through this distance under the surface to pick up the energy Φ_{Ba}/q.

(c) An electron which has a collision within the distance d_{cr} would not be emitted into the interface.

This model then predicts the emission probability to be

$$P = A_{em} \exp(-d_{cr}/\lambda) \tag{4.21}$$

where A_{em} is a constant ($A_{em} = 2.9$) and λ is the mean free path of electron optical-phonon collision. This mean free path is defined in Section 4.2, and is given by

$$\lambda = \lambda_o \tan \frac{\varepsilon_r}{2kT} \tag{4.22}$$

where $\lambda_o = 108$ Å, $\varepsilon_r = 0.063$ eV, k is the Boltzman constant and T is the lattice temperature. From Eq. (4.22), $\lambda(25°C) = 91$ Å, $\lambda(300°C) = 61$ Å. It is obvious that emission probability is much smaller at high temperature. λ_o is larger than the value used in Eq. (4.9) to account for multiple-phonon emission.

The number of electrons injected into the interface per unit area in time Δt is

$$N_{inj} = \frac{J_{em}\,\Delta t}{q} = \frac{\Delta t}{q}\left(J_{L_{g-r}} + J_{L_{diff}}\right)A_{em}\, e^{-[d_{cr}/\lambda(T)]} \tag{4.23}$$

The rate of change of threshold voltage is

$$\frac{\Delta V_{th}}{\Delta_t} = \frac{N_{inj} \cdot q}{\Delta t \cdot C_{ox}} = \frac{J_{em}}{C_{ox}} \tag{4.24}$$

and the rate of transconductance change (in saturation region) is

$$\frac{\Delta g_m}{\Delta t} = \left(- u \frac{W}{L} C_{ox}\right) \frac{\Delta V_{th}}{\Delta t} = - u \frac{W}{L} J_{em} \quad (4.25)$$

This effect causes long-term threshold voltage shift and transconductance degradation.

Table 4.1 shows calculated parameters and threshold voltage drift rate due to the substrate hot-electron effect. An NMOS device with $N_a = 3.4 \times 10^{16}/cm^3$ and $T_{ox} = 600$ Å is used; $J_L = 10^{-8}$ A/m^2 at 25°C and $J_L = 10^{-1}$ A/m^2 at 300°C are assumed. It is shown that except for low substrate bias voltage and low oxide field, substrate hot-electron effect is more important at high temperature. Also, substrate bias is more important than the gate oxide electric field in determining the threshold voltage drift.

4.4 Electromigration of Aluminum Metallization

At high temperature and high current-density operation, aluminum metallization will pose a reliability problem. The so-called "electron wind" inside the conductor forces the metal ions to travel toward the positive end of the conductor while vacancies move toward the negative end. The vacancies condense to form voids while the metal ions condense at certain discontinuities to form crystals or hillocks. This mechanism results in an electrical open circuit due to the loss of conductor metal.

Black [1969] has experimentally characterized the mean time of failure (MTF) in terms of temperature, current density and dimensions

Table 4.1 Threshold Voltage Drift Due to Substrate Hot-Electron Effect

	V_{sub} = 5 volt, V_{ox} = 12 volt		V_{sub} = 5 volt, V_{ox} = 3 volt		V_{sub} = 3 volt, V_{ox} = 3 volt		V_{sub} = 3 volt, V_{ox} = 12 volt	
T_c	25°C	300°C	25°C	300°C	25°C	300°C	25°C	300°C
ϕ_{Ba} (eV)	2.574	2.574	3.097	3.097	3.097	3.097	2.574	2.574
ϕ_s (volt)	0.76	0.27	0.76	0.76	0.27	0.27	0.76	0.27
d_{cr} (μm)	$1.20\text{x}10^{-1}$	$1.27\text{x}10^{-1}$	$1.50\text{x}10^{-1}$	$1.6\text{x}10^{-1}$	$2.19\text{x}10^{-1}$	$2.71\text{x}10^{-1}$	$1.65\text{-}10^{-1}$	$1.90\text{x}10^{-1}$
A_{em}	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9
λ (Å)	91	61	91	61	91	61	91	61
P	$5.62\text{x}10^{-6}$	$2.52\text{x}10^{-9}$	$7.33\text{x}10^{-8}$	$7.95\text{x}10^{-12}$	$1.03\text{x}10^{-10}$	$1.47\text{x}10^{-19}$	$3.67\text{x}10^{-8}$	$9.05\text{x}10^{-14}$
J_L (A/cm^2)	10^{-8}	10^{-1}	10^{-8}	10^{-1}	10^{-8}	10^{-1}	10^{-8}	10^{-1}
$\frac{\Delta V_{th}}{\Delta t}$ (V/sec)	$9.8\text{x}10^{-7}$	$4.38\text{x}10^{-3}$	$1.28\text{x}10^{-8}$	$1.38\text{x}10^{-5}$	$1.79\text{x}10^{-9}$	$2.56\text{x}10^{-13}$	$6.39\text{x}10^{-9}$	$1.57\text{x}10^{-7}$

$N_a = 3.6\text{x}10^{16}/cm^3$; T_{ox} = 600Å

of aluminum metal film. His result is

$$\frac{MTF}{W_c \cdot d} = \left[A_m \, J^2 \exp\left(-\phi/kT\right)\right]^{-1} \tag{4.31}$$

where

$MTF \triangleq$ mean time of failure (hours)

$A_m \triangleq$ a constant

$J \triangleq$ current density on the alumimum stripes (A/cm^2)

$\phi \triangleq$ activation energy (electron volt)

$k \triangleq$ Boltzman constant

$T \triangleq$ film temperature in degree Kelvin

$W_c \triangleq$ conductor width (cm)

$d \triangleq$ conductor thickness (cm) (4.32)

For large crystallite film (grain size $\cong$ 8 μm), $A_m = 5 \times 10^{-13}$, $\phi =$ 0.84 eV; for small crystallite film (grain size $\cong$ 1.2 μm), $A_m =$ 2.43×10^{-16}, $\phi = 0.48$ eV. The mean time of failure decreases rapidly as temperature increases. This effect is important for power supply interconnection stripes, since these must carry the largest current in the circuit. For example, assume $W_c = 8 \times 10^{-4}$ cm, $d = 0.6 \times 10^{-4}$ cm, $J = 2.08 \times 10^5$ A/cm^2 ($I = J \cdot W_c d = 10$ mA) for an aluminum stripe.

The calculated values of MTF at 25°C, 150°C and 300°C are 1.16×10^9 hrs., 7.4×10^4 hrs., and 179 hrs., respectively.

For high-temperature, high-current operation, one has to choose wider and thicker metallization or use refractory metal or metal silicides; W and WSi_2 are some examples. It is believed that they have lower A and high ϕ than aluminum metallization. Continuous operation of 68 hours of a tungsten metal stripe with $J = 4.1 \times 10^6$ A/cm^2, W_c = 10 μm, d = 0.61 μm without failure at 300°C is reported in Blacke [1980]. At the same condition, an aluminum stripe has a mean time to failure of MTF = 3.6 hours. However, tungsten oxidizes above 300°C, and a passivation scheme should be considered.

4.5 Summary

Avalanche multiplication and breakdown, hot-electron, and electromigration effects have been investigated in this chapter. The avalanche multiplication factor decreases at high temperature, but the junction leakage seed current is higher at high temperature. The breakdown voltage, which is defined as the voltage where the multiplication factor is infinite, is not a major concern. However, the multiplied current could be higher at high temperature, and could cause soft breakdown in NMOS devices. At high temperature, the substrate hot-electron effect is the most important; channel hot-electron effect is only important at room or low temperature. Substrate hot-electron effects could be avoided by reducing substrate bias voltage and

substrate concentration, and increasing gate oxide thickness within the constraint limit imposed by the circuit design.

Finally, for high temperature MOS operation there exists a maximum current density (which decreases with increasing temperature) that aluminum metallization could safely carry. Refractory metals appear to be very promising, but more characterization of the mean time of failure is required.

CHAPTER 5

MOS CIRCUITS AT HIGH TEMPERATURE

5.1 Introduction

Having discussed threshold voltage variation with temperature, we are ready to study MOS circuits at high temperature. Emphasis is on digital circuits because they are more widely used. In this chapter we first discuss the onset temperature of MOS devices when temperature effects start to become important. We then discuss zero-temperature-coefficient current for MOS devices and methods to achieve zero-TC current. MOS inverter circuit at high temperature is investigated next, taking into consideration the high temperature characteristics of the inverter, different methods to improve its high temperature performance, and the choice of inverter type for better performance at high temperature. We then study a dynamic logic circuit, a shift register, its high temperature behavior and ways to improve its performance at high temperature. Finally, problems with conventional diode protection methods are discussed, and a new protection scheme with better performance at high temperature is devised.

5.2 Onset Temperature

Except for very small drain current, where $(V_g - V_{th})$ is small, I_d will decrease as temperature increases. For a quantitative measure of this current drop effect, we define the onset temperature T_{os} as

that temperature at which drain current decreases to 80% of its room temperature value.

At any temperature $T > T_o$, the drain current is given by

$$I_d = A\mu_o\left(\frac{1}{T_o^{-1.5}}\right) T^{-1.5} (V_g - V_{th})^2$$

$$= A\mu_o\left(\frac{1}{T_o^{-1.5}}\right) T^{-1.5} \Big\{V_g - [V_{tho} + a(T-T_o)]$$

$$- \gamma\sqrt{2\phi_{f_T} - V_{sub}} + \gamma\sqrt{2\phi_{f_T}}\Big\}^2 \tag{5.1}$$

where

$$a \triangleq \left.\frac{-dV_{th}}{dT}\right|_{V_{sub}=0} \qquad A \triangleq \frac{1}{2}\left(\frac{W}{L}\right) C_{ox}$$

$$\gamma \triangleq \frac{\sqrt{2\varepsilon_s q N_a}}{C_{ox}} \qquad T_o = 298°K \tag{5.2}$$

and

$$V_{tho} = V_{th}(T_o)$$

Note that a is the threshold voltage variation per degree Kelvin without substrate bias.

(i) Onset temperature for zero substrate bias or high gate bias voltage:

This case occurs when $V_{sub} = 0$ or $V_g - [V_{tho} - a(T-T_o)] >> \gamma\sqrt{2\phi_{f_T} - V_{sub}}$

$$I_d(25°C) = A\mu_o \frac{1}{298^{-1.5}} (298^{-1.5})(V_g - V_{tho})^2 \tag{5.3}$$

and

$$I_d(T_{os}) = 0.8\ I_d(25°C)$$

$$= A\mu_o \left(\frac{1}{298^{-1.5}}\right)\left(T_{os}^{-1.5}\right) [V_g - V_{tho} - a(T_{os} - 298)]^2 \tag{5.4}$$

Then T_{os} is found from

$$\left(\frac{298}{T_{os}}\right)^{-1.5} = \frac{[V_g - V_{tho} - a(T_{os} - 298)]^2}{(V_g - V_{tho})^2 \cdot 0.8} \tag{5.5}$$

As an example, we assume $V_{tho} = 2.49$ volt, $a = -5.09$ mV/°C and $V_g - V_{tho}$ = 2 volts; then $T_{os} = 432°K = 159°C$.

(ii) Onset temperature for non-zero substrate bias voltage:

This occurs when the value of $\gamma\sqrt{2\phi_{f_T} - V_{sub}}$ is important compared to other items. Onset temperature is found from

$$\left(\frac{298}{T_{os}}\right)^{-1.5} = \frac{\left[V_g - V_{tho} - a(T_{os} - 298) - \gamma\sqrt{2\phi_{f_{T_{os}}} - V_{sub}} + \gamma\sqrt{2\phi_{f_{T_{os}}}}\right]^2}{\left[V_g - V_{tho} - \gamma\sqrt{2\phi_{f_{T_o}} - V_{sub}} + \gamma\sqrt{2\phi_{f_{T_o}}}\right]^2 \cdot 0.8} \tag{5.6}$$

where

$$\phi_{f_{T_{os}}} = \frac{kT_{os}}{q} \ln \frac{N_a}{n_i(T_{os})}$$

$$n_i(T) = 3.86 \times 10^{16} \; T^{3/2} \; \exp\left(\frac{-E_{go}}{2kT}\right)$$

$$E_{go} = 1.21 \text{ eV}$$

Assuming $\gamma = 3.384v^{\frac{1}{2}}$, $\phi_{f_{T_0}} = 0.38$, $V_g - V_{tho} = 3$ volt, $V_{sub} =$ -1 volt and $a = -5.09$ mV/°C, then we obtain $T_{os} = 368°K = 95°C$. Note that because N_a is chosen high enough to obtain $V_{th}(300°C) > 0$, a and γ are larger than typical values.

Onset temperature in Case (ii) is lower than that in Case (i). This is because the decrease of V_{th} is less when substrate bias is applied. Therefore, the number of inverted electrons $C_{ox}(V_g - V_{th})$ is less in Case (ii) compared with Case (i).

In the above analysis, we assume that leakage currents are negligible at those temperatures. As is shown in Section 3.2, this is a reasonable assumption.

5.3 Zero-TC Current

In analog integrated circuits, device parameters such as threshold voltage, mobility and transconductance all vary with temperature. These variations can often be compensated by use of feedback techniques. However, it is important that the devices be properly biased so that they operated in the correct mode at all temperatures.

Accordingly, we investigate the feasibility of operating an MOS device at some bias condition for which the drain current exhibits nearly zero variation with temperature. It is shown that this is possible at temperatures below that at which the leakage current become dominant. Also this ZTC drain current can be achieved either by applying proper gate bias or by using an appropriate substrate bias voltage.

The drain current is given by

$$I_d = \mu \frac{W}{L} \int_{V_S}^{V_D} Q_n(y)dV = \mu \frac{W}{L} \theta \tag{5.7}$$

where

$$\theta = C_{ox} \left\{ (V_D - V_S) \left[V_g - V_{fb} - 2\phi_f - \frac{V_D}{2} + \frac{V_S}{2} \right] - \frac{2}{3} \sqrt{2\varepsilon_s q N_a} \frac{1}{C_{ox}} \left[(V_D + 2\phi_f - V_{sub})^{3/2} - (V_S + 2\phi_f - V_{sub})^{3/2} \right] \right\} \tag{5.8}$$

The percentage change of I_d with respect to temperature is

$$\frac{1}{I_d} \frac{dI_d}{dT} = \frac{1}{\mu} \frac{d\mu}{dT} + \frac{1}{\theta} \frac{d\theta}{dT}$$

$$\triangleq \frac{-1.5}{T} + \frac{1}{\theta} \frac{d\theta}{dT} \tag{5.9}$$

where $d\theta/dT$ is found to be

$$\frac{d\theta}{dT} = C_{ox} \left\{ (V_D - V_S) \left[- \frac{dV_{fb}}{dT} - 2 \frac{d\phi_f}{dT} \right] - \frac{2\sqrt{2\varepsilon_s q N_a}}{C_{ox}} \left[(V_D + 2\phi_f - V_{sub})^{\frac{1}{2}} - (V_S + 2\phi_f - V_{sub})^{\frac{1}{2}} \right] \frac{d\phi_f}{dT} \right\} \tag{5.10}$$

For normal doping concentration, the first part of the equation

$$C_{ox}(V_D - V_S) \left[- \frac{dV_{fb}}{dT} - 2 \frac{d\phi_f}{dT} \right]$$

dominates the rest.

From the above equations we note that

(i) when θ is large (high gate bias or small substrate bias)

$$\frac{1}{\mu}\frac{d\mu}{dT} + \frac{1}{\theta}\frac{d\theta}{dT} < 0 \tag{5.11}$$

and I_d has a negative temperature coefficient.

(ii) when θ is small (low gate bias or large substrate bias)

$$\frac{1}{\mu}\frac{d\mu}{dT} + \frac{1}{\theta}\frac{d\theta}{dT} > 0 \tag{5.12}$$

and I_d has a positive temperature coefficient.

(iii) between these extrema, there is an optimum θ which will give the lowest temperature coefficient.

(iv) if a zero TC point exists, then

$$\frac{1}{I_d}\frac{dI_d}{dT} = 0 \tag{5.13}$$

or

$$\frac{1}{\theta}\frac{d\theta}{dT} = \frac{1.5}{T} \tag{5.14}$$

$$\theta = \left(\frac{\theta_o}{T_o^{1.5}}\right)T^{1.5} \tag{5.15}$$

and

$$\frac{d\theta}{dT} = \frac{1.5\ \theta_o}{T_o^{1.5}}\ T^{0.5} \tag{5.16}$$

T_o and θ_o are the reference temperature and corresponding θ at that temperature. Note that since $d\theta/dT$ depends on dV_{fb}/dT and $d\phi_f/dT$, it cannot be constrained to have the above temperature dependence. Therefore, a true ZTC current does not exist.

However, the drain current can be made the same at two temperatures, T_o and T_1. This can be achieved either by V_g or V_{sub} (V_{sub} is more effective for higher substrate concentration N_a) for reasonably separated V_D and V_S voltages. From Eq. (5.15)

$$\frac{\theta_1}{\theta_o} = \left(\frac{T_1}{T_o}\right)^{1.5} = \frac{\left\{(V_D-V_S)\left[V_g-V_{fb_1}-2\phi_{f_1}-\frac{V_D}{2}+\frac{V_S}{2}\right]-\frac{2}{3}\sqrt{2\varepsilon_s qN_a}\cdot\frac{1}{C_{ox}}\right.}{\left\{(V_D-V_S)\left[V_g-V_{fb_o}-2\phi_{f_o}-\frac{V_D}{2}+\frac{V_S}{2}\right]-\frac{2}{3}\sqrt{2\varepsilon_s qN_a}\cdot\frac{1}{C_{ox}}\right.}$$

$$\frac{\left.\left[(V_D+2\phi_{f_1}-V_{sub})^{3/2}-(V_S+2\phi_{f_1}-V_{sub})^{3/2}\right]\right\}}{\left.\left[(V_D+2\phi_{f_o}-V_{sub})^{3/2}-(V_S+2\phi_{f_o}-V_{sub})^{3/2}\right]\right\}} \tag{5.17}$$

and the "quasi-zero-TC current per unit W/L ratio" $I_{dz}/(W/L)$ is

$$\frac{I_{dz}}{W/L} = \mu\theta = u_o \left(\frac{T}{T_o}\right)^{-1.5} \left(\frac{T}{T_o}\right)^{1.5} \theta_o = u_o\, \theta_o \qquad (5.18)$$

In the linear region, bias voltage V_D is used in Eq. (5.17), the "quasi-zero-TC current per unit W/L ratio" $[I_{dz}/(W/L)]$ is independent of the choice of V_g or V_{sub}. For design purposes, one can specify the quasi-zero-TC current I_{dz} and the two temperatures T_o and T_1; for given V_g (or V_{sub}) value, one can use Eq. (5.17) to find the necessary V_{sub} (or V_g) value, θ_o and θ_1 to achieve this current. From θ_o one can find the proper W/L ratio to obtain the specified I_{dz}.

Note that the linear region is not of any practical interest. For analog circuits, we would always want the device to operate in the saturated region.

In the saturation region, V_D in Eq. (5.17) is replaced by the drain saturation voltage $V_{D_{sat}}$ as

$$V_{sat} = V_g - V_{fb} - 2\phi_f - \frac{\varepsilon_s q N_a}{C_{ox}^2} \left[\sqrt{1 + \frac{2C_{ox}^2}{\varepsilon_s q N_a}(V_g - V_{fb} - V_{sub})} - 1\right] \qquad (5.19)$$

Because of the $V_{D_{sat}}$ dependence on V_g and V_{sub}, one has to use an iterative method to solve Eq. (5.17) to obtain V_g (or V_{sub}) for quasi-zero-TC current at temperatures T_o and T_1. Slightly different values of $[I_{dz}/(W/L)]$ result between V_g adjustment and V_{sub} adjustement for

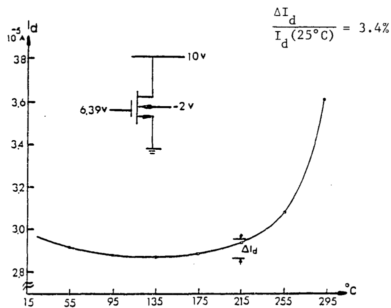
zero-TC currents; depending on the magnitude of $V_{D_{sat}}$. Higher $V_{D_{sat}}$ values will have higher I_{dz}.

Between these temperatures I_d will first decrease because of mobility degradation and then increase because of the decrease of ϕ_f. Figures 5.1 and 5.2 show experimental results for a case in which currents were matched at 25°C and 215°C with both gate voltage adjustment (Fig. 5.1) and substrate bias adjustment (Fig. 5.2). In the latter case, because of the high value of $|V_{sub}|$, $V_{D_{sat}}$ is smaller and smaller I_{dz} results compared with the former case. In future discussions, we refer to the current which has been matched at two temperatures as the zero-TC current.

(v) Zero-TC current scales with W/L ratio.

(vi) When the temperature is very high, leakage currents are comparable with channel current. Thus, zero-TC matching is impossible because leakage currents are not controlled by V_g or V_{sub}. Figures 5.1 and 5.2 shows this effect when T is higher than 215°C.

(vii) Zero-TC current might not exist depending on the bias conditions. If $|V_D - V_S|$ is too small or $|V_{sub}|$ too large, it may be impossible to find V_g to match I_d at two different temperatures. If $|V_D - V_S|$ and $|V_g|$ are too small, it may be impossible to find V_{sub} to match I_d at two temperatures. Figure 5.3 shows this effect. Note that if we extend the curve to positive V_{sub} region, they will meet at one point. This is not, of course, an allowable V_{sub} since it would cause source or drain junctions to be forward biased.



Device Parameters:

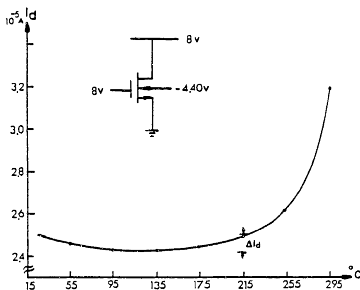
$N_a = 3.4x10^{16}/cm^3$

$T_{ox} = 1000Å$

$Q_{SS}/q = 4x10^{10}/cm^2$

Al gate

Fig. 5.1 Experimentally Measured Drain Current as a Function of Temperature; Currents Are Matched at 25°C and 215°C by Varying Gate Bias Voltage



$$\frac{\Delta I_d}{I_d(25°C)} = 4\%$$

Device parameters are the same as given in Fig. 5.1

Fig. 5.2 Experimental Measured Drain Current as a Function of Temperature; Currents Are Matched at 25°C and 215°C by Varying Substrate Bias Voltage

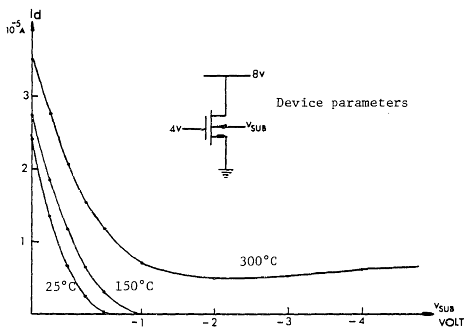


Fig. 5.3 An NMOS Device Where Zero-TC Does Not Exist

(viii) A simple solution could be obtained if we use the square law approximation (when in saturation region)

$$I_d = A\mu_o \left(\frac{T}{T_o}\right)^{-1.5} (V_g - V_{th})^2$$

$$= A\mu_o \left(\frac{T}{T_o}\right)^{-1.5} [V_g - V_{tho} + a(T-T_o) - \gamma\sqrt{2\phi_{f_T} - V_{sub}} + \gamma\sqrt{2\phi_{f_T}}\ ]^2 \tag{5.20}$$

(a) To match I_d at temperature T_o and T_1 with a given substrate voltage, we first write the current as

$$I_{d_o} = A\mu_o \left(\frac{T_o}{T_o}\right)^{-1.5} \left[V_g - V_{tho} + a(T_o - T_o) - \gamma\sqrt{2\phi_{f_{T_o}} - V_{sub}} + \gamma\sqrt{2\phi_{f_{T_o}}}\right]^2 \tag{5.21}$$

$$I_{d_1} = A\mu_o \left(\frac{T_1}{T_o}\right)^{-1.5} \left[V_g - V_{tho} + a(T_1 - T_o) - \gamma\sqrt{2\phi_{f_{T_1}} - V_{sub}} + \gamma\sqrt{2\phi_{f_{T_1}}}\right]^2$$

$$= I_{d_o} \tag{5.22}$$

From this, the necessary gate voltage is found to be

$$V_g = V_{tho} + a(T_o - T_1) + \frac{\left[a(T_o-T_1) - \gamma\left(\sqrt{2\phi f_{T_o} - V_{sub}} - \sqrt{2\phi f_{T_1} - V_{sub}}\right) + \gamma\left(\sqrt{2\phi f_{T_o}} - \sqrt{2\phi f_{T_1}}\right)\right]}{\left(\frac{T_o}{T_1}\right)^{3/4} - 1} \tag{5.23}$$

The value of current which must be used is found by substituting this V_g in the equation for I_{d_o}.

(b) To match I_d at temperatures T_o and T_1 with substrate voltage with a given gate voltage V_g, we find V_{sub} from the following equation

$$\frac{cc - bb\sqrt{aa - V_{sub}}}{dd - bb\sqrt{ee - V_{sub}}} = \left(\frac{T_1}{T_o}\right)^{3/4} \tag{5.24}$$

where

$$aa \triangleq 2\phi f_{T_o}$$

$$bb \triangleq \gamma$$

$$cc \triangleq V_g - V_{tho} + \gamma\sqrt{2\phi f_{T_o}}$$

$$dd \triangleq V_g - V_{tho} + a(T_1 - T_o) + \gamma\sqrt{2\phi f_{T_1}}$$

$$ee \triangleq 2\phi f_{T_1} \tag{5.25}$$

For a specified I_{dz} at two temperatures T_o and T_1, if $|V_D-V_S|$ and V_{sub} are in reasonable range, one can choose some V_g or V_{sub} and with the proper W/L ratio achieve zero-TC current at these temperatures. In the linear region, the ZTC value per W/L ratio $[I_{dz}/(W/L)]$ is a function of process only, independent of bias conditions. In the saturation region, $[I_{dz}/(W/L)]$ is a function of process and the $V_{D_{sat}}$ value. Higher values of $V_{D_{sat}}$ result in higher values of $[I_{dz}/(W/L)]$, but the magnitudes of $[I_{dz}/(W/L)]$ are very close in all cases.

Note that from a practical point of view, using substrate bias means that the device will have to be in its own well, since other devices will not use the same V_{sub}. Thus, this discussion is most important for the CMOS case, where it is possible to use separate P-wells for NMOS devices.

5.4 Inverter Behavior at High Temperatures

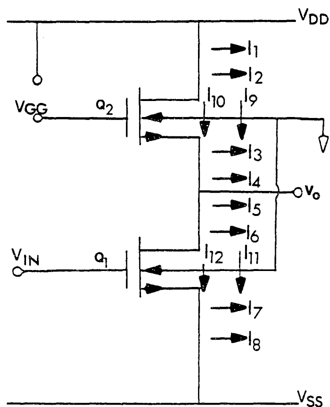
In this section we use the model shown in Fig. 3.4 to simulate a saturated-load NMOS inverter. The schematic diagram including all the leakage current components is shown in Fig. 5.4. In this figure

$I_1 = (Ad_2)\ J_L$ = Drain leakage current of load transistor Q_2 from power supply V_{DD} to substrate

$I_2 = \frac{1}{2}\ (WL)_2\ J_L$

$I_3 = \frac{1}{2}\ (WL)_2\ J_L$

$I_4 = (A_{s2})\ J_L$ = source leakage current of Q_2 from output node to substrate



$V_{GG} = V_{DD}$ for a saturated-load inverter

Fig. 5.4 Schematic Diagram Showing All the Current Components of an NMOS Inverter

$I_5 = (A_{d_1}) \quad J_L$ = drain leakage current of driver transistor Q_1 from output node to substrate

$I_6 = \frac{1}{2} (W\ L)_1 \quad J_L$

$I_7 = \frac{1}{2} (W\ L)_1 \quad J_L$

$I_8 = (A_{s_1}) \quad J_L$

$I_9 = I_{chan2}$

$I_{10} = I_{sub2}$

$I_{11} = I_{chan1}$

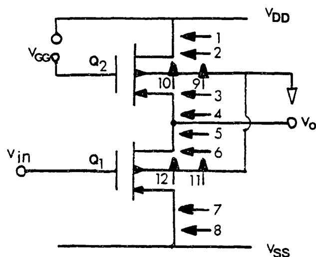
$$I_{12} = I_{sub2} \tag{5.26}$$

Note that for PMOS and CMOS inverters, directions of current components are different (Figures 5.5 and 5.6).

During room temperature operation V_o remains at the edge of the saturation point $\left(V_o \triangleq V_{o_{max}} = V_{DD} - V_{th}\Big|_{Q_2}\right)$ when V_{in} is less than $V_{th}\Big|_{Q_1}$. The transition slope is very high and is given by

$$A_o = \frac{-1}{1+\eta} \frac{g_{m_1}}{g_{m_2}} = \frac{-1}{1+\eta} \sqrt{\frac{K'(W/L)_1 I_{d_1}}{K'(W/L)_2 I_{d_2}}} = \frac{-1}{1+\eta} \sqrt{\frac{(W/L)_1}{(W/L)_2}} \tag{5.27}$$

where



$V_{GG} = V_{DD}$ for a saturated-load inverter

Fig. 5.5 Schematic Diagram Showing All the Current Components of a PMOS Inverter

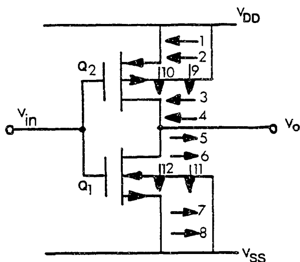


Fig. 5.6 Schematic Diagram Showing All the Current Components of a CMOS Inverter

$$\eta \triangleq \frac{\gamma}{2\sqrt{-V_{BS_2} + 2\phi_f}} \tag{5.28}$$

is the V_{BS}-to-I_d transconductance coefficient defined as

$$g_{m_b} \triangleq \frac{\partial I_d}{\partial V_{BS}} = \eta g_m \tag{5.29}$$

V_{BS} is body to source bias voltage and

$$\gamma \triangleq \frac{\sqrt{2\varepsilon_s q N_a}}{C_{ox}} \tag{5.30}$$

When V_{in} increases beyond $V_{th}\big|_{Q_1}$, V_o starts to drop. The source-body effect causes the transition slope to decrease further until Q_1 enters the non-saturated region.

As temperature slightly increases, $V_{th}\big|_{Q_2}$ is lowered; this results in a higher $V_{o_{max}}$, where

$$V_{o_{max}} = V_{DD} - V_{th2} = V_{DD} - (V_{tho} + \gamma\sqrt{-V_{BS_2} + 2\phi_f} - \gamma\sqrt{2\phi_f}) \tag{5.31}$$

V_{tho} is the threshold voltage of Q_2 without source-body effect. The transition slope is almost unchanged with respect to temperature because $-V_{BS_2} >> 2\phi_f$ in this transition region.

As temperature further increases, I_{sub1}, I_{sub2}, and various junction leakage currents cannot be neglected. Then:

(i) leakage current components flowing from the source of Q_2 and drain of Q_1 to the substrate will force Q_2 to conduct, causing $V_{o_{max}}$ to drop by the gate-source voltage V_{gs_2} necessary to sustain these leakage currents.

$$V_{o_{max}} = V_{DD} - (V_{tho} + \gamma\sqrt{-V_{BS_2} + 2\phi_f} - \gamma\sqrt{2\phi_f}) - V_{gs2} \tag{5.32}$$

and:

(ii) $I_{chan1} \neq I_{chan2}$, and Eq. (5.27) no longer hold true.

Instead, we apply KCL in the circuit of Fig. 5.4

$$I_9 + I_{10} - I_3 - I_4 - I_5 - I_6 = I_{11} + I_{12} \tag{5.33}$$

$$\begin{aligned} I_{chan2} \overset{\Delta}{=} I_9 &= I_{11} + I_{12} - I_{10} + I_3 + I_4 + I_5 + I_6 \\ &= I_{chan1} + (I_{sub2} - I_{sub1}) + \frac{1}{2}(W\ L)_2\ J_L \\ &+ (A_{s_2} + A_{d_1})\ J_L + \frac{1}{2}(W\ L)_1\ J_L \end{aligned} \tag{5.34}$$

and Eq. (5.27) becomes

$$\begin{aligned} A_o &= \frac{-1}{1+\lambda}\frac{g_{m_1}}{g_{m_2}} = \frac{-1}{1+\lambda}\sqrt{\frac{K'(W/L)_1 I_{chan1}}{K'(W/L)_2 I_{chan2}}} = \frac{-1}{1+\lambda}\sqrt{\frac{(W/L)_1}{(W/L)_2}}\ F \\ &= (A_o)_{low\ temperature}\ F \end{aligned} \tag{5.35}$$

Here, F is the slope-degradation factor due to the high-temperature leakage components.

$$F = \sqrt{\frac{I_{chan1}}{I_{chan2}}}$$

$$= \sqrt{1 - \frac{(I_{sub2}-I_{sub1})+\frac{1}{2}(W\ L)_2\ J_L+\frac{1}{2}(W\ L)_1\ J_L+(As_2+As_1)\ J_L}{I_{chan2}}} \quad (5.36)$$

In the above analysis, current components I_1, I_2, I_7, and I_8, have no effect in changing the high-temperature characteristics. They contribute only to the power dissipation.

Figure 5.7 shows the measured and simulated transfer curves of a saturated load inverter at 25°C, 150°C, and 300°C, respectively. It shows that:

(i) $V_{o_{max}}\ (150°C) > V_{o_{max}}\ (25°C) > V_{o_{max}}\ (300°C)$

(ii) The transition slopes are almost the same for T = 25°C and T = 150°C cases, but the slope at T = 300°C is much less.

(iii) The transition edge at T = 300°C is not as sharp as at T = 25°C and T = 150°C. This is because the subthreshold current increases greatly at T = 300°C and the range of gate-source voltages for which subthreshold conduction occurs has been increased.

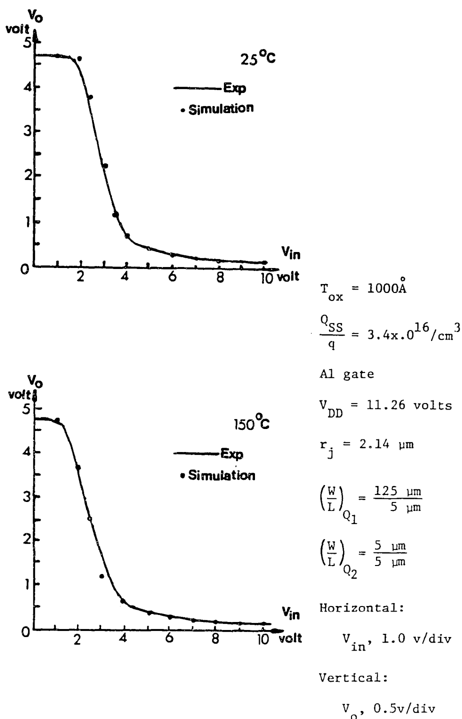


Fig. 5.7 Measured and Simulated Inverter Transfer Curves at 25°C, 150°C, and 300°C

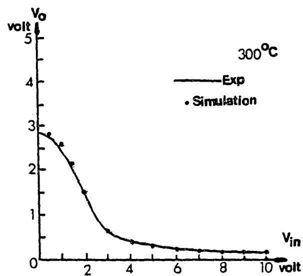


Fig. 5.7 (Continued)

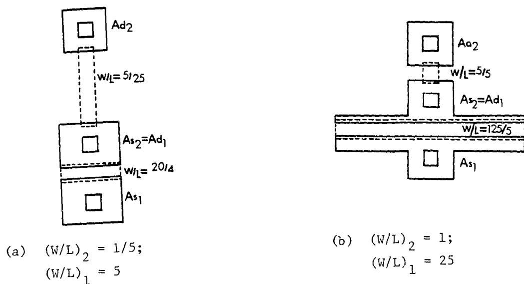
In this simulation, substrate leakage currents and subthreshold leakage currents are included. The program and its flow graph are listed in Appendix B.

5.5 Improvement of Inverter Characteristics at High Temperatures

5.5.1 Improvement of Inverter Characteristics by Geometry Layout Change

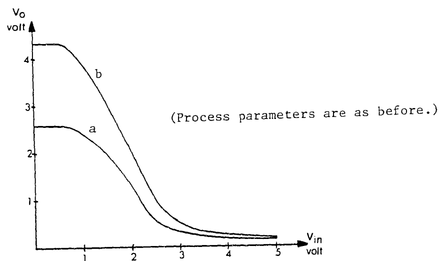
In an inverter operated at high temperature, the various leakage components at the source of the load device (drain of the driver device) can force the load device to conduct even when the driver device is cut off. These components cause additional voltage drop at output node of the inverter. One way to reduce this voltage drop is to increase the transconductance of the load device so that only a small V_{gs} for the load transistor enables the drain current of the device to supply all the leakage components. This also requires an increase of the transconductance of the driver.

Figure 5.8 shows this feature. Here a saturated-load inverter with V_{DD} = 11.26 volt and device parameters of $Q_{SS}/q = 4 \times 10^{10}/cm^2$, T_{ox} = 1000Å, $N_a = 3.4 \times 10^{16}/cm^3$, and Al gate is simulated. Both inverters have a β ratio of 5 where $\beta^2 = \frac{(W/L)_D}{(W/L)_L}$, but Case (b) has five times higher load transistor transconductance. The input-output characteristic at 300°C is shown in Fig. 5.8. Note that the maximum output voltage has increased almost 1.8 volt in Case (b). The price paid for this improvement is an increase of the total channel area used by the inverter; in the above example, the channel area increased by a factor of 3.2.



(a) $(W/L)_2 = 1/5$;
$(W/L)_1 = 5$

(b) $(W/L)_2 = 1$;
$(W/L)_1 = 25$



(c) Corresponding Transfer Curves of (a) and (b)

Fig. 5.8 Two Different Layout Geometries of a Saturated-Load Inverter

An alternate way to incorporate a high transconductance device for the load transistor is to use a depletion-load inverter; this is discussed in Section 5.5.3.

5.5.2 Improvement of Inverter Characteristics by Charge Pumping

Another technique that can be used to increase the output voltage of an inverter is to externally pump the charge into that node from a pulsed supply. As shown in Fig. 5.9, the circuit is similar to a shift register. The output of the first stage inverter is boosted by the pulse voltage v; that signal is then transferred to the next stage by keeping the clock signal ϕ high.

The amount of increase of V_o during pumping period is

$$\Delta V_o = \frac{v\ C_1}{C_1 + C_2 + C_3} \qquad (5.37)$$

where

C_1 is the gate-source capacitance of the charge pumping transistor Q_3

C_2 is the capacitance at the output node of the first inverter

C_3 is the source-substrate capacitance of Q_3

Note also that two gated switches, transistor Q_1 and Q_2, are used, so V_o will be pumped up when its value is less than V_{DD} and higher than $V_{o_{max}}$ (T).

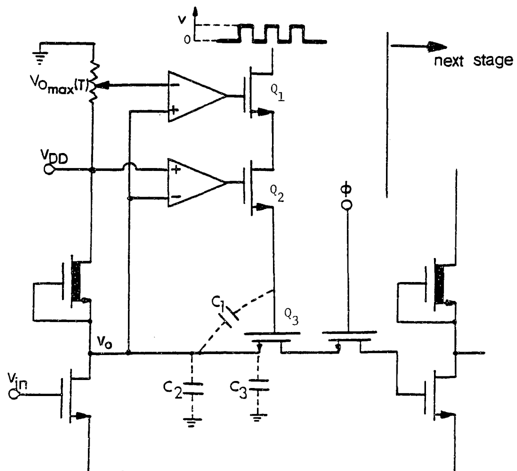


Fig. 5.9 Charge-Pumping Technique to Boost the Output Voltage of an Inverter

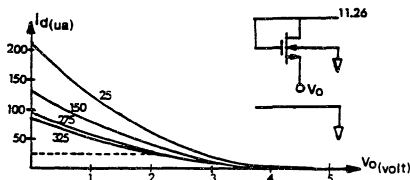
One disadvantage of this technique is that voltage at the first inverter output node (voltage across C_2) will be reduced by the source/substrate and drain/substrate leakage current associated with that node. Therefore, high frequency pumping voltage V should be used. Another disadvantage is that it requires a lot more silicon area, several switches plus gain blocks.

5.5.3 Choice of Inverter Circuit

For inverters using devices all having the same polarity of inversion channel, for example, n-type, three different types of load transistor configurations can be used. They are saturated-load inverter, unsaturated-load inverter, and depletion-load inverter. The depletion-load inverter has the highest switching speed and lowest power consumption. The unsaturated-load inverter has higher speed than the saturated-load case, but requires an extra gate bias.

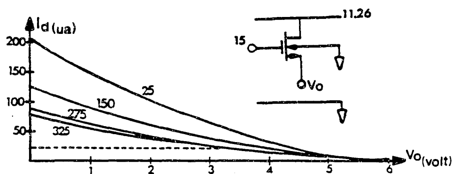
At high temperature, leakage current at the source of the load transistor (or the drain of the driver transistor) is the major concern. For a given amount of leakage current flowing at the output node, the depletion-load inverter has the highest transconductance and, therefore, the smallest decrease of $V_{o_{max}}$.

This effect is shown in Fig. 5.10. In this figure, we plot the load line characteristics of three different types of inverter. I_d at 25°C is chosen to be the same in all cases. At room temperature, $V_{o_{max}}$ for these inverters are 4.7 volts, 6.8 volts, 5 volts for Cases (a), (b), and (c), respectively. For a total of 25 μA leakage current at output node at 325°C, the maximum output voltage will drop to 2 volts, 3.2



(a) Saturated-Load Inverter

Parameters are the same as given below except
$(W/L) = \frac{7.45\ \mu m}{10\ \mu m}$



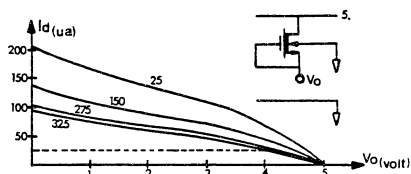
(b) Unsaturated-Load Inverter

$(W/L) = \frac{3.5\ \mu m}{10\ \mu m}$; $V_{sub} = 0$; $N_a = 3.4x10^{16}/cm^3$;

$x_j = 2.14\mu$; $T_{ox} = 1000\mathring{A}$; $\frac{Q_{SS}}{q} = 4x10^{10}/cm^2$;

μ_o (25°C) = 423 cm^2/v.sec; Al gate

Fig. 5.10 Load Line Curve for Three Different Types of Inverters



(c) Depletion-Load Inverter

$V_{sub} = 0$; $N_a = 1.9\times10^{15}/cm^3$; $N_{imp} = 2.0\times10^{16}/cm^3$;

$x_{I_{imp}} = 0.2\ \mu m$; $x_j = 2.14\mu$; $(W/L) = \frac{20\ \mu m}{8\ \mu m}$;

$T_{ox} = 1000\text{Å}$; $\frac{Q_{SS}}{q} = 4\times10^{11}/cm^2$; Al gate;

$\mu_s\ (25°C) = 400\ cm^2/v.sec$; $\mu_b\ (25°C) = 800\ cm^2/v.sec$

Fig. 5.10 (Continued)

volts and 4 volts in Cases (a), (b), and (c), respectively. Therefore, the depletion-load inverter, owing to its high transconductance, is least susceptible to the leakage current's effect. The saturated-load inverter, on the other hand, has the lowest value of $V_{O_{max}}$ at 325°C.

5.5.4 Diode Compensation

One of the circuit techniques that can improve the high temperature performance of an inverter is to use a diode to cancel (i) the P-N junction leakage current at the drain of the driver (source of the load) transistor, plus (ii) the distributed substrate-to-channel leakage current of both driver and load devices. For example, in a saturated-load inverter, if complete cancellation occurs, the load device will remain cut off even at high temperature; therefore, the maximum output voltage $V_{O_{max}}$ will stay at one threshold voltage drop below the power supply voltage. This condition is met when

$$J_d A_d = J_L \left[Ad_1 + As_2 + \tfrac{1}{2}(WL)_1 + \tfrac{1}{2}(WL)_2\right] \tag{5.38}$$

where

$J_d \triangleq$ leakage current density of the compensation diode

$A_d \triangleq$ junction area of the diode

$J_L \triangleq$ leakage current density of n^+P junction of MOS devices

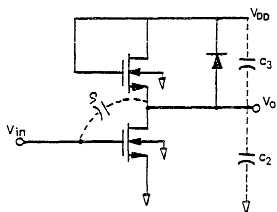
$Ad_1 \triangleq$ drain area of driver transistor

$As_2 \triangleq$ source area of load transistor

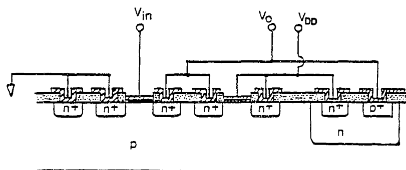
$(WL)_1 \triangleq$ gate dimension of driver transistor

$(WL)_2 \triangleq$ gate dimension of load transistor

Figure 5.11 shows the schematic diagram and cross-section view of this diode-compensated inverter.



(a) Schematic Diagram



(b) Cross-Section View

Fig. 5.11 Schematic Diagram and Cross-Section View of a Diode-Compensated Saturated-Load Inverter

It is, of course, expected that addition of the diode with its attendant capacitance will degrade the transient response of the inverter. Since the circuit is nonlinear and large signals are applied, analysis of the problem is complicated. However, for our purposes we are interested in a comparison of the response with and without the diode. An indication of relative performance can be obtained by examining the frequency response of the circuits from a small-signal point of view. Figure 5.12 shows the small signal equivalent circuit of an inverter with the compensation diode. Since the diode is always reverse biased, it is represented as C_3, the average depletion capacitance of the diode. Let

C_1 = capacitance between input and output nodes ($\cong C_{gd_1}$)

C_2 = capacitance between ground and output node ($\cong C_{d-sub_1} + C_{s-sub_2}$)

C_3 = average voltage-dependent junction capacitance of the diode

$$= \frac{1}{V_{o_{max}} - V_{o_{min}}} \int_{V_{o_{min}}}^{V_{o_{max}}} \frac{C_{j_o}}{\left(1 - \frac{V}{V_{bi}}\right)^m} \tag{5.39}$$

Applying Kirchoff's Current Law at the output node, we obtain

$$\frac{V_o}{r_{ds_1}} + g_{m_1}V_{in} + V_oSC_2 + (V_o - V_{in})SC_1 = \frac{-V_o}{r_{ds_2}} - g_{m_2}V_o - ng_{m_2}V_o - V_oSC_3 \tag{5.40}$$

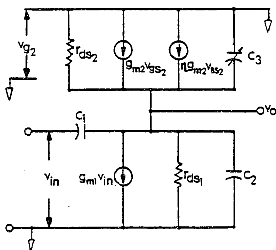


Fig. 5.12(a) Small Signal Equivalent Circuit of a Diode-Compensated Saturated-Load Inverter

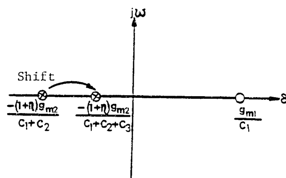


Fig. 5.12(b) Pole-Zero Positions of the Saturated-Load Inverter; Arrow Shows the Shift of Pole Position Due to the Junction Depletion Capacitance of the Diode

If it is assumed that $r_{ds_1} \to \infty$, $r_{ds_2} \to \infty$, Eq. (5.40) becomes

$$\frac{V_o}{V_{in}} = \frac{-g_{m_1} + SC_1}{g_{m_2}(1+\eta) + S(C_1 + C_2 + C_3)} \tag{5.41}$$

Comparing this with the transfer function without diode compensation, we see that

$$\left(\frac{V_o}{V_{in}}\right)_{\text{without diode}} = \frac{-g_{m_1} + SC_1}{g_{m_2}(1+\eta) + S(C_1 + C_2)} \tag{5.42}$$

The pole has been shifted toward the real axis; this will slow the response of the circuit. Note that for a well-defined diode junction area, a reverse CMOS process must be used so that the diode can be fabricated from an n-well with a p^+ diffusion.

The SPICE 2G program is used to simulate this inverter with a pulse train input. Nominal values of C_1, C_2, and C_3 are used (C_{gd} = 0.1 pF, $C_2/A = 2.4 \times 10^{-4}$ F/cm^2, $C_{jo} = 1.554 \times 10^{-13}$ F). The result is shown in Figures 5.13 and 5.14. It is obvious that the compensated circuit does suffer in the frequency response at 25°C and 300°C. However, the response time is not significantly worse but logic levels are restored at 300°C.

5.6 Dynamic Logic: A Shift Register

To further investigate the influence of leakage current at high temperature, a simple two-phase shift register is simulated by the SPICE 2G program. The schematic diagram is shown in Fig. 5.15, where

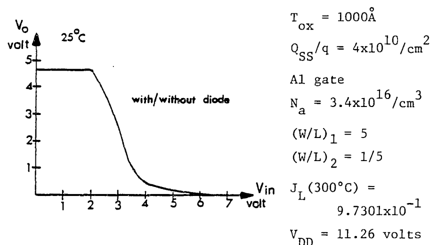


Fig. 5.13(a) Transfer Curve of a Saturated-Load Inverter at 25°C, With and Without Diode Compensation

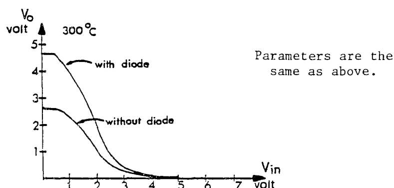
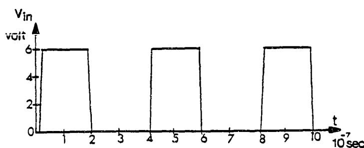
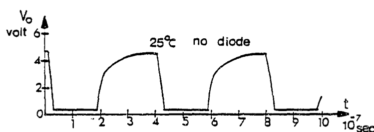


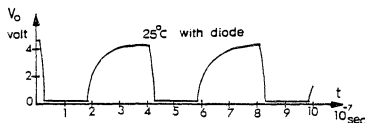
Fig. 5.13(b) Transfer Curve of a Saturated-Load Inverter at 300°C Showing the Effect of Diode Compensation



(a) Input Waveform



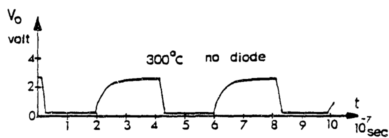
(b) Output at 25°C, no compensation



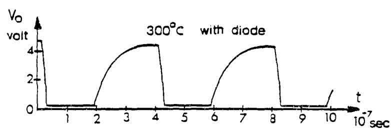
(c) Output at 25°C, with compensation

$C_{diode} = 24\ nF/cm^2$

Fig. 5.14 Transient Response of the Saturated-Load Inverter, Process Parameters as Before

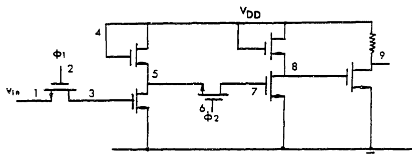


(d) Output at 300°C,
no compensation

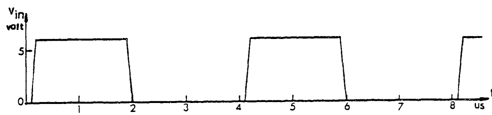


(e) Output at 300°C,
with compensation

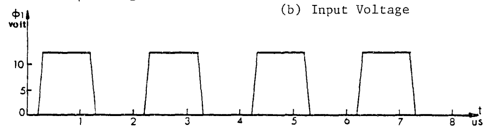
Fig. 5.14 (Continued)



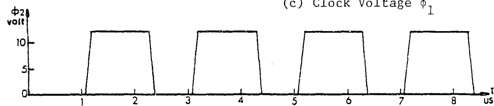
(a) Schematic Diagram



(b) Input Voltage



(c) Clock Voltage ϕ_1



(d) Clock Voltage ϕ_2

Fig. 5.15 A Saturated-Load Shift Register

the last stage transistor is used as a buffered output stage. The input voltage V_{in}, and clock voltages ϕ_1 and ϕ_2 are also shown in Fig. 5.15.

Figure 5.16 shows voltages at nodes 3, 5, 7 and 8 at 25°C. At this temperature V_{th} = 2.49 volts and leakage currents are too small to change the characteristics of this shift register. Comparing V_{in} and V_8, we see that a shifted input voltage is obtained at node 8.

At 300°C, leakage currents become important in determining the voltage level at each node. Accurate amounts of leakage currents at each node must be included. Again, voltage-controlled current sources are used for simulation of the diffusion components of junction leakage currents.

While the capacitances at node 3 and 7 will hold the charge when clock signals ϕ_1 and ϕ_2 are turned off for 25°C operation, it is not true at 300°C. Leakage currents of the source/substrate and drain/substrate junctions will cause the capacitance to discharge. In addition to this effect, voltages at node 5 and 8 decrease because of the junction leakage currents at these nodes. These two effects together will cause logic errors at the output.

When ϕ_1 is off, the voltage at node 3 is

$$V_3 = V_{in}\Big|_{\text{when } \phi_1 \text{ is turned off}} - \frac{I_{L3}}{C_1}\Delta T_1 \qquad (5.43)$$

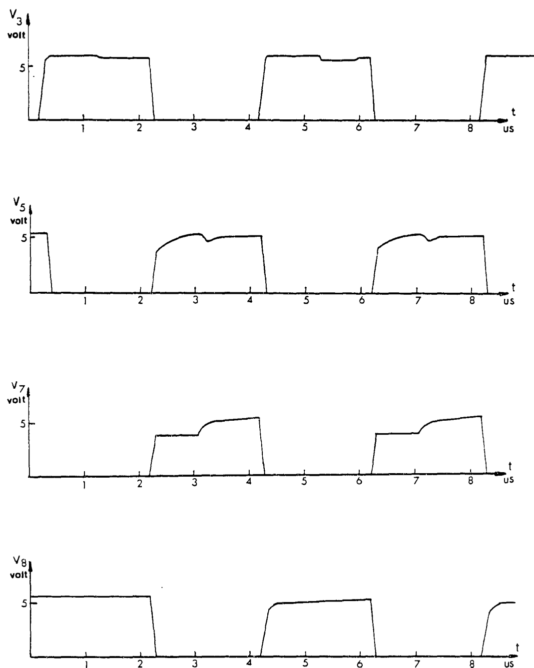


Fig. 5.16 Voltages at Node 3, 5, 7 and 8 of Above Shift Register at 25°C, Where Leakage Current is Negligible

where

ΔT_1 is the time span after ϕ_1 is turned off

C_1 is the total capacitance associated with node 3

I_{L3} is the total leakage current at node 3

The voltage at node 7 when ϕ_2 is off is

$$V_7 = V_5 \Big|_{\text{when } \phi_2 \text{ is turned off}} - \frac{I_{L7}}{C_2} \Delta T_2 \qquad (5.44)$$

where

ΔT_2 is the time span after ϕ_2 is turned off

C_2 is the total capacitance associated with node 7

I_{L7} is the total leakage current at node 7

At 300°C, with $J_L = 9.7301 \times 10^{-2}$ A/cm^2, the voltages at nodes 3, 5, 7, and 8 are as shown in Fig. 5.17a. Note that the voltages at node 3 and 7 drop when the clock signals are turned off. Too large a voltage drop at node 7 will cause a bump in V_8 as shown in Fig. 5.17a.

One way to increase the voltage level at V_7 is to increase the signal and clock frequencies. This is shown in Fig. 5.17b, where all frequencies have been doubled. The bump has been reduced in magnitude from 3.2 volts to 1.3 volts.

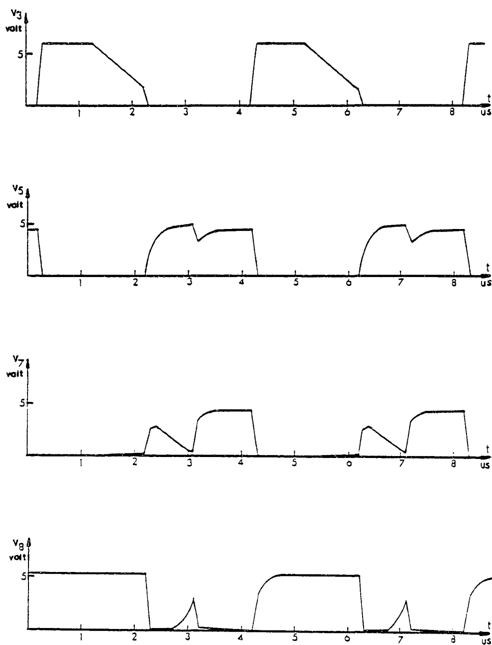


Fig. 5.17(a) Voltages at Node 3, 5, 7, 8 at 300°C.
$J_L = 9.7301 \times 10^{-2} A/cm^2$ is assumed.

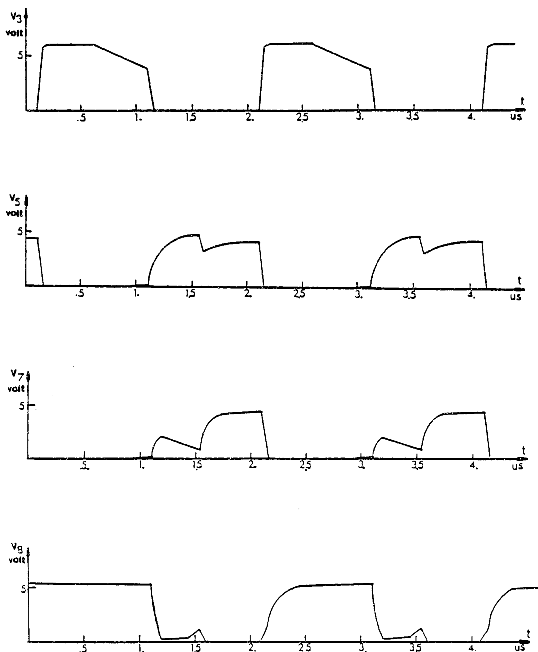


Fig. 5.17(b) Voltages at Node 3, 5, 7, 8 at 300°C.

$J_L = 9.7301 \times 10^{-2} A/cm^2$ is assumed.

All the frequencies are doubled from previous figure.

However, when leakage current density is too high, reducing ΔT_2 would produce little improvement in the response. This is shown in Figures 5.18a and 5.18b where $J_L = 9.7301 \times 10^{-1}$ A/cm^2 is assumed. By increasing frequencies by an order of magnitude, one will restore the voltage at node 3 in suitable shape to turn on the first stage inverter. But the voltage drop at first inverter output node at 5 makes V_7 too small to turn on the second inverter, thus resulting in a logic error. One then must rely on techniques to increase V_5 and V_7; diode compensation probably is the simplest. Note that in designing the diode, one must take into account the source-substrate junction leakage current of the clock transistor. For example, for the first inverter stage, it requires

$$J_d \quad A_d = J_L \quad [\sum_i A_i \Big|_{\text{at node 5}}] + J_L \quad A_s \Big|_{\text{transistor } \phi_2} \tag{5.45}$$

For the second stage inverter

$$J_d \quad A_d = J_L \quad [\sum_i A_i \Big|_{\text{at node 8}}] \tag{5.46}$$

Also, the frequency response will be degraded because of the introduction of the reverse-biased junctions.

In conclusion, general design guidelines for this type of shift register for high temperature operations are:

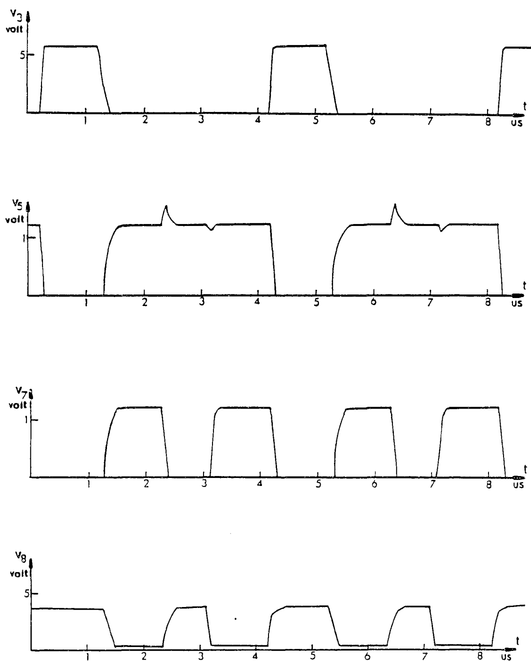


Fig. 5.18(a) Voltages at Node 3, 5, 7, 8 at 300°C.

$J_L = 9.7301 \times 10^{-1} A/cm^2$ is assumed.

Obvious logic error shown at V8.

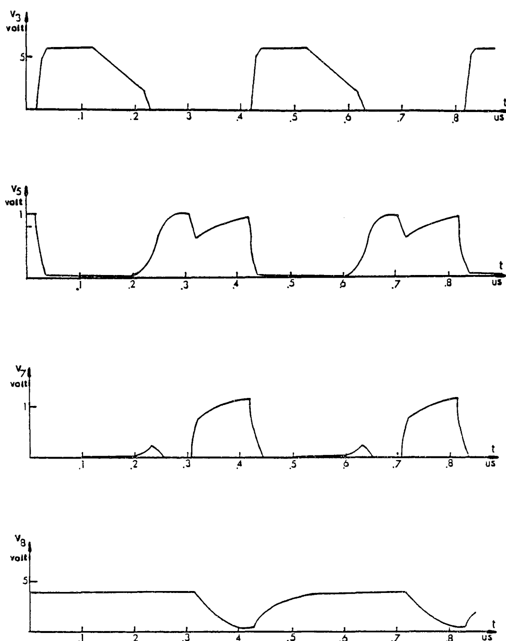


Fig. 5.18(b) Voltage at Node 3, 5, 7, 8 at 300°C. $J_L = 9.7301 \times 10^{-1}$ A/cm^2 and all frequencies are ten times higher than previous figure. $V_{o_{max}}$ drop at Node 7 is the main reason for logic error shown at V8.

(i) To increase the voltage output of the inverter stages, diode compensation is very powerful and easy to implement.

(ii) Increasing the frequency of operation (within the frequency response limit of the circuit) improves the waveforms.

Although increasing the capacitances at nodes 3 and 7 will reduce the voltage drop due to leakage current effects, their values are determined by processes and thus they are not considered to be variable parameters here.

5.7 Diode Protection in Logic Circuits

Another concern in MOS high temperature logic design is the necessity for protection diodes. Diode protection is extensively used in commercial logic circuits to prevent possible damage to the MOS gate oxide from static charge (and voltage) accumulation. Conventional methods of diode protection use a resistor and avalanche diodes to prevent gate voltage from reaching breakdown. Figure 5.19 shows a resistor-diode type protection circuit for N channel, P channel and CMOS circuits. Figure 5.20 shows another type of protection circuit where a pair of thick-oxide MOS devices is diode connected.

At high temperature, reversed-biased diode leakage current becomes important. If the two reverse diode currents do not cancel exactly, a net current will flow through the resistor. This causes additional V_{in} shift in the input-output characteristics. Experimental data for a CMOS inverter with diode protection is shown in Fig. 5.21a [Coquat, 1980]

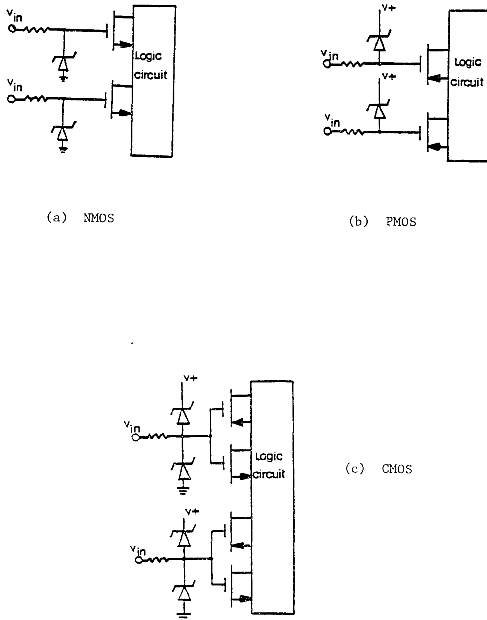


Fig. 5.19 Diode Protection Schemes in NMOS, PMOS and CMOS Logic Circuits

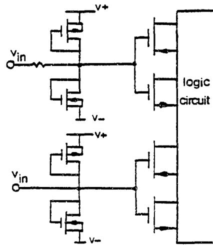
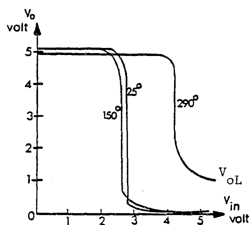
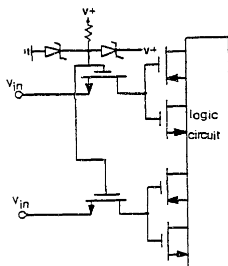


Fig. 5.20 Thick Oxide Diode Connected Transistor Protection Scheme in CMOS Circuits



(a) CMOS NAND Gate Transfer Characteristics; Note the Shift of V_{in} and Elevated V_{oL} at 290°C



(b) New Protection Scheme in CMOS Logic Circuits

Fig. 5.21 CMOS Characteristics and a New Protection Scheme

Note that at 290°C, a net current flows through the protection resistor (into the input gate) causing an equivalent V_{in} shift of 1.3 volts. The output of the inverter may be unable to turn on the next stage, and the circuit will malfunction. This happens when

$$V_{in} - (\Delta I_{diode} \cdot R) < V_{th}(\text{N channel}) \tag{5.47}$$

For the fabrication of these protection diodes, P well and n^+ diffusion regions are used for anode and cathode, respectively. This has the disadvantage that a larger P well region is required and the leakage currents are higher. Another diode protection scheme shown in Fig. 5.21b is therefore suggested. Here a protection transistor is added between every signal input and the gates of logic circuits. The gates of these protection transistors are protected by conventional methods. However, each logic input is connected to the source of the corresponding protection transistor. The loading on the input signal is mainly due to the source-to-substrate leakage current of the protection transistor and is less in magnitude than protection diode leakage currents of the conventional method. The frequency response of these protection transistors is determined by the source-to-substrate and drain-to-substrate capacitances and the turn-on resistance R_{on} of the protection transistor. If R_{on} is designed to be very small, this protection scheme will have response time comparable to the conventional method. Another advantage of this protection scheme is that

only one pair of diodes is used for all inputs; thus, the total area used for protection circuits is reduced.

Note that in Fig. 5.2.b, V^+ must be large enough to allow for sufficient swing of V_{in}. For this particular circuit V^+ should be at least equal to the summation of V_{in}, V_{th} (including source-body effect) and the voltage drop across the protection resistor.

5.8 Summary

In this chapter, we have investigated the existence of a zero-temperature-coefficient drain current. It has been shown that while such a current may not be obtainable, it is possible with proper bias conditions to design for equal drain currents at two different temperatures, provided that the highest temperature is below that at which leakage current becomes dominant.

Because of the wide use in digital circuits, an MOS inverter and a dynamic shift register at high temperature have been studied. Degradation of circuit performance is caused mainly by the leakage currents. Inverter characteristics can be improved by geometry layout change, charge pumping and diode compensation. Among them, diode compensation has the best performance. We have also shown that the depletion-load inverter has better high-temperature performance than saturated- or unsaturated-load inverters. Finally, a new diode protection scheme which avoids the large area P well - n^+ diffusion diodes in the signal path has been devised.

CHAPTER 6

MAXIMUM OPERATING TEMPERATURE

6.1 Introduction

In this chapter, we investigate the maximum operating temperature for MOS devices and integrated circuits. The physical limits are studied first; other limits on circuit performance and device characteristics are then examined. Finally, since leakage current is the most important concern for high-temperature MOS operation, its influence on the maximum temperature is discussed.

It is noteworthy that temperature limits imposed by circuit performance constraints are usually somewhat lower than those imposed by device physics constraints.

6.2 Substrate Concentration Considerations

6.2.1 Effect on Surface Channel Depletion Region

If the temperature is high enough such that $n_i = N_a$ where N_a is the substrate concentration, the $\phi_f = 0$, $x_{d_{max}} = 0$ and band bending is reduced to zero. No depletion region exists between the surface inversion channel and the bulk substrate. The device then appears to be a bulk resistor with resistivity ρ_s determined by n_i. This sets a temperature limit above which MOS function does not exist.

$$n_i(T_{max}) = N_a \tag{6.1}$$

Figure 6.1a shows the relation between T_{max} and N_a. Figure 6.1b shows T_{max} versus ρ_s for p-type and n-type substrates with process parameters of $Q_{SS}/q = 4 \times 10^{10}/cm^2$, $T_{ox} = 1000Å$.

This temperature limit can be increased by using non-zero substrate bias voltages. However, in this case, enhancement MOS transistors do not exist.

6.2.2 Effect on Threshold Voltages

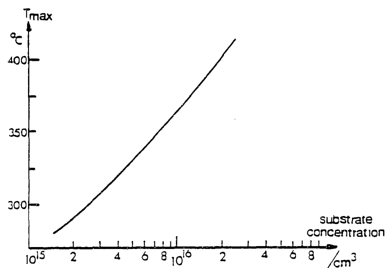
As temperature increases, threshold voltage decreases for N-channel devices and increases for P-channel devices. It will approach zero in both cases as temperature increases. To ensure $V_{th} \geq 0$ for N-channel ($V_{th} \leq 0$ for P-channel) enhancement type devices, there is a minimum substrate doping concentration that must be used. For devices without substrate bias voltage:

$$N_{a(min)} = \left[\left(\phi'_{MS(T_{max})} + 2\phi_{f(T_{max})} - \frac{Q_{SS(T_{max})}}{C_{ox}} \right) C_{ox} \right]^2 \frac{1}{4\phi_{f(T_{max})} q \, \varepsilon_s} \tag{6.2}$$

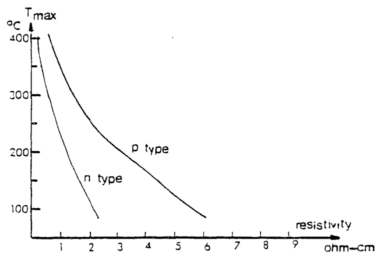
For devices with substrate bias voltage:

$$N_{a(min)} = \left[\left(\phi'_{MS(T_{max})} + 2\phi_{f(T_{max})} - \frac{Q_{SS(T_{max})}}{C_{ox}} \right) C_{ox} \right]^2$$

$$\frac{1}{2\left[2\phi_{f(T_{max})} - V_{sub}\right] q \, \varepsilon_s} \tag{6.1}$$



(a) T_{max} versus Substrate Concentration



(b) T_{max} versus Resistivity

Fig. 6.1 Maximum Operation Temperature versus Substrate Concentration and Resistivity to Ensure $N_a > n_i\ (T_{max})$

Therefore, for a given substrate doping concentration the maximum operating temperature T_{max} is found by using the relation

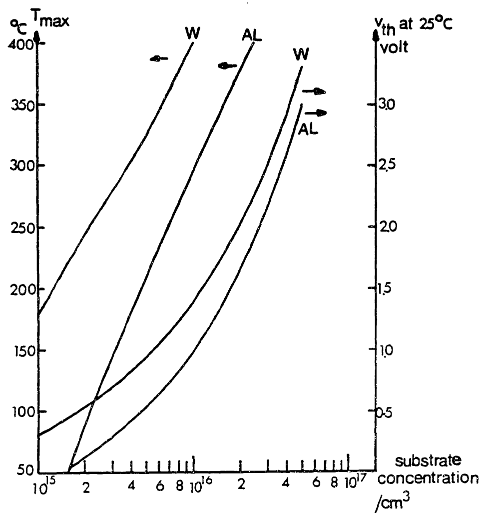
$$V_{th}\ (T_{max}, N_a, Q_{SS}, C_{ox}, \phi'_{MS}) = 0 \qquad (6.4)$$

Figure 6.2 shows T_{max} versus N_a for two different gate material N-channel devices without substrate bias. In this figure

$$\frac{Q_{SS}}{q} = 4.0 \times 10^{10}/cm^2\ ,$$

and T_{ox} = 1000Å are used. For example, to ensure $V_{th} \geq 0$ at 300°C for a grounded substrate N-channel device, $N_{a(min)}$ is found to be $4.0 \times 10^{15}/cm^3$ and $1.1 \times 10^{16}/cm^3$ for tungsten and aluminum gate MOS, respectively. The right-hand side of Fig. 6.1 shows V_{th} at room temperature. For these values of $N_{a(min)}$, the room-temperature threshold voltages are 1.5 volts for tungsten gate MOS and 1.0 volt for aluminum gate MOS devices. A similar plot for 500Å oxide thickness is shown in Fig. 6.3.

For P channel devices, the situation is different. Here ϕ_f is negative due to the upward band bending (Fig. 6.4a) of the type substrate. This will keep V_{th} negative for a reasonable range of substrate concentrations. Figure 6.4b shows V_{th} at 400°C and 25°C versus substrate concentration (with no substrate bias applied) with T_{ox} = 1000Å, $Q_{SS}/q = 7.5 \times 10^{10}/cm^2$. It can be seen that for substrate concentration lower than $10^{16}/cm^3$, V_{th} is very stable as far as N_a is concerned and remains negative even at 400°C. A similar plot for 500Å oxide thickness is shown in Fig. 6.5.



T_{ox} = 1000Å

$Q_{SS}/q = 4.0 \times 10^{10}/cm^2$

N-Channel

Fig. 6.2 T_{max} versus Substrate Concentration to Assure $V_{th}(T_{max}) > 0$. Also shown is the threshold voltage at 25°C.

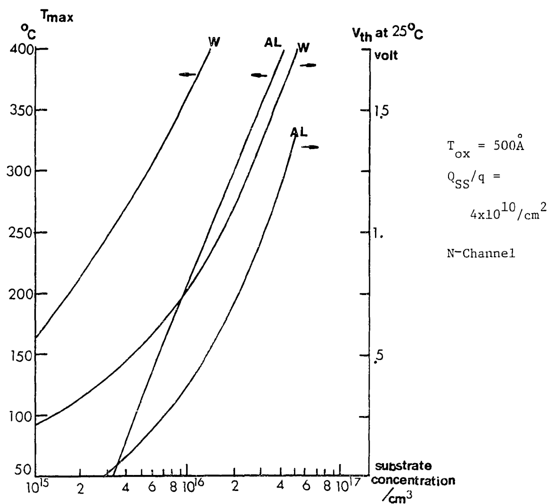


Fig. 6.3 T_{max} versus Substrate Concentration to Assure $V_{th}(T_{max}) > 0$. Also shown is the threshold voltage at 25°C.

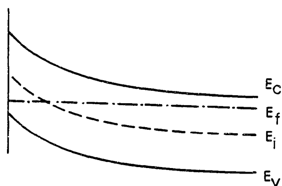


Fig. 6.4(a) Band Bending Diagram for a PMOS Device

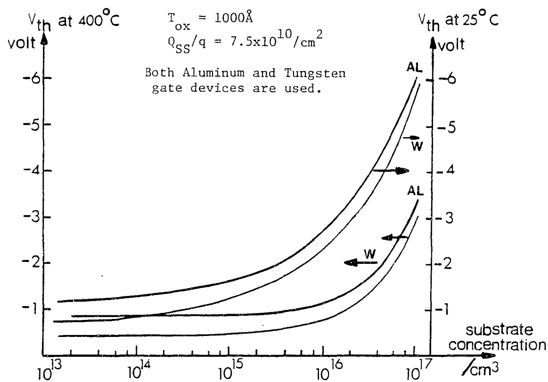
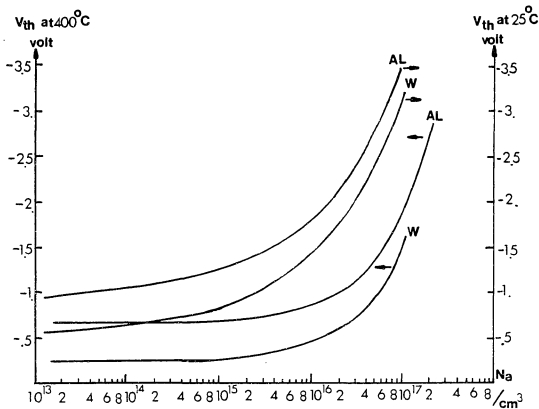


Fig. 6.4(b) V_{th} versus Substrate Concentration at 25°C and 400°C for a P-Channel MOS Device



$T_{ox} = 500\text{Å}$

$Q_{SS}/q = 7.5 \times 10^{10}/cm^2$

Fig. 6.5 V_{th} versus Substrate Concentration at 25°C and 400°C for a P-Channel MOS Device

6.3 Temperature Limit Due to Leakage Current Density

Leakage current density is a major concern in high-temperature MOS integrated circuits. Its effects are also widely different according to application. Here we consider several aspects of this leakage current density.

6.3.1 Channel to Leakage Current Ratio Consideration

This ratio R is the channel current divided by the summation of all leakage components at the drain, and is given by

$$R = \frac{I_{chan}}{I_{L_{d-sub}} + 0.5\ I_{L_{chan-sub}} + I_{sub}} \qquad (6.5)$$

where $I_{L_{d-sub}}$ is the drain to substrate leakage current (g-r plus diffusion leakage current).

In analog circuit applications this ratio is an important parameter. It determines the shape of the load line characteristics of MOS devices. Low channel-to-leakage current ratio devices have only small voltage swing and limited signal input range as shown in Fig. 6.6. Figure 6.7 shows R value versus temperature for a particular process with three different (W/L) dimensions. Leakage current densities are taken from measured data (under 300°C) and extrapolated data (above 300°C). It is clear that unless leakage current density could be reduced dramatically (for example, by using a gettering process) a device operated at 375°C has a ratio R of 1 for W/L = 160 μm/5 μm.

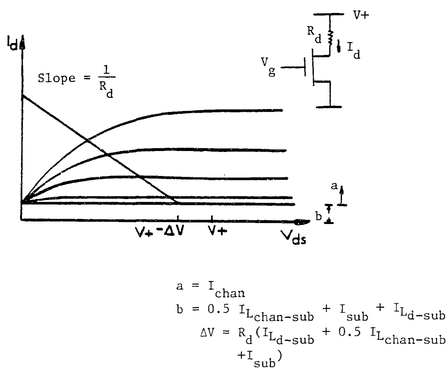
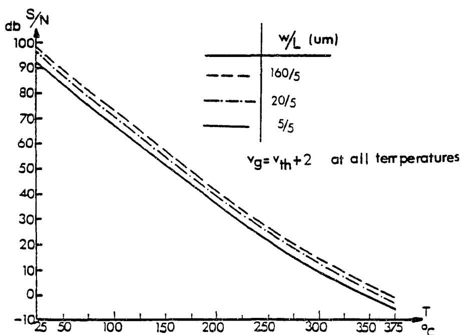


Fig. 6.6 Load Line Characteristics of a Resistor-Loaded MOS Transistor at High Temperature



$N_a = 3.4\mathrm{x}10^{16}/\mathrm{cm}^3$, $\frac{Q_{SS}}{q} = 4\mathrm{x}10^{10}/\mathrm{cm}^2$,

$\mu_o = 423\ \mathrm{cm}^2/\mathrm{V{\cdot}sec}$, $T_{ox} = 1000\text{Å}$,

Al gate, $N_T = 3.124\mathrm{x}10^{20}/\mathrm{m}^3$,

$\sigma_T = 1\mathrm{x}10^{-20}\mathrm{m}^2$, $\tau_n = \tau_p = 1\mathrm{x}10^{-5}$

$5\mathrm{x}10^{-6}\ (300/T)^{\frac{1}{2}}$ sec,

5μ design rules are used for the calculation of the junction area.

Fig. 6.7 Channel-to-Leakage Current Ratio for a Typical NMOS Device

6.3.2 Leakage Current Effect on the Source Junction

As shown in Fig. 6.8a, when the source node is connected to the substrate, there is a built-in junction potential barrier V_{bi} across source and substrate. At high temperature, leakage current flowing through the substrate introduces voltage drop across the bulk silicon resistor. If the voltage drop is larger than the potential barrier, the source-substrate junction is then forward-biased and parasitic bipolar transistor operation takes place with source as emitter, substrate as base and drain as collector, as shown in Fig. 6.8b. This situation is similar to the thermal "latch-up" process to be discussed later. It occurs when

$$\left(\rho_{epi} \frac{L_{sub}}{WL} \right) \left(J_{L(T_{max})} \; W \; L \right) = V_{bi(T_{max})} \tag{6.6}$$

If small-dimension devices are used, a high electric field exists at the drain-substrate junction. Latch-up will then occur at lower temperature because the thermally generated current is multiplied by avalanche multiplication.

$$\left(\rho_{epi} \frac{L_{sub}}{WL} \right) \left(J_{L(T_{max})} \; W \; L \right) M_{(T_{max})} = V_{bi(T_{max})} \tag{6.7}$$

where $M(T_{max})$ is the multiplication factor at the maximum temperature.

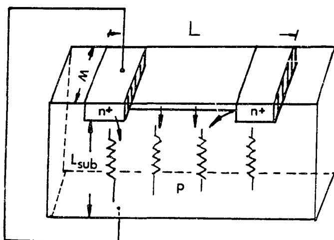
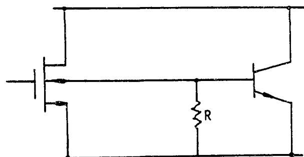


Fig. 6.8(a) MOS Device at High Temperature Showing the Leakage Current Components



R is the equivalent resistance of the substrate.

Fig. 6.8(b) MOS Device with Parasitic Bipolar Transistor Action

As an example, for a P-type substrate of $3 \times 10^{16}/cm^3$ doping concentration ($\rho_{epi} = 0.6\ \Omega$-cm) with source doping of $1.0 \times 10^{19}/cm^3$, $L_{sub} = 400\mu$, $J_{L(T_{max})} = 0.5$ amp/cm^2 and $M(T_{max}) = 9.6$, then we obtain

$$0.6 \times \frac{400 \times 10^{-4}}{W\ L}\ 0.5 \times 9.6 \times W\ L = 0.5 \text{ volt} \cong V_{bi}\ (328°C) \tag{6.8}$$

The maximum temperature is found to be approximately 328°C.

To increase this temperature limit, a small substrate bias voltage of the order of V_{bi} has to be applied to increase the source junction barrier height. In a logic circuit the load device already has body bias (except for CMOS). A small separate source body bias to the driver device could avoid this problem at the price of using another supply voltage. It also changes the threshold voltage of that device and may be impractical in the circuit design. Also, a thinner substrate increases this temperature limit.

6.3.3 Leakage Current Effect on Inverter Characteristics

As we learned in Section 5.4, in the single channel inverter circuit leakage current will force the load transistor to conduct, causing a maximum output voltage $V_{o(max)}$ drop. But, $V_{o(max)}$ has to be larger than the threshold voltage of the driver transistor of the next stage to be able to turn on the next stage transistor.

$$V_{o(max)} \geq V_{th(T_{max})} \tag{6.9}$$

This is shown graphically in Fig. 6.9 for an enhancement saturated-load inverter. Two different leakage current density parameters are used:

(a) $\sigma_T = 1 \times 10^{-20}/m^2$; $N_T = 3.124 \times 10^{20}/m^3$ (typical value),

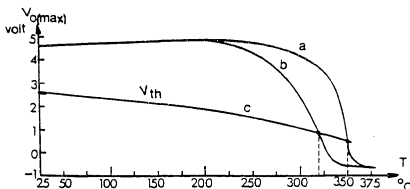
and

(b) $\sigma_T = 1 \times 10^{-20}/m^2$, $N_T = 3.124 \times 10^{22}/m^3$ (worst case value).

$V_{o(max)}$ versus temperature is shown in Curves (a) and (b), respectively. Curve (c) shows threshold voltage variation with temperature for this particular MOS process. T_{max} is found to be 320°C and 350°C for Cases (a) and (b), respectively.

Note that $V_{o(max)}$ increases slightly before T = 200°C due to the threshold voltage drop at high temperature. This $V_{o(max)}$ drop due to the leakage currents varies with different types of logic circuit. As shown in Section 5.6, depletion-load inverter has the lowest $V_{o(max)}$ drop for the same amount of leakage current at the source of the load transistor. This saturation would be improved if one increases the drain voltage of the load transistor for saturated-load and depletion-load inverters and the gate voltage of unsaturated-load inverter.

At the output node of a CMOS inverter, there is an internal cancellation of the substrate-to-drain plus substrate-to-channel leakage currents of P-channel and the drain-to-substrate plus channel-to-substrate leakage current of N-channel device. The $V_{o(max)}$ drop is not important.



(a) Typical value $N_T = 3.124\text{x}10^{20}/\text{m}^3$, $\sigma_T = 1\text{x}10^{-20}\text{m}^2$,

J_L (300°C) = 9.7301×10^{-1} A/cm^2

(b) Worst case value $N_T = 3.124 \times 10^{22}/\text{m}^3$, $\sigma_T = 1 \times 10^{-20}\text{m}^2$,

J_L (300°C) = 9.7301×10^{0} A/cm^2

(c) Threshold voltage versus temperature

$N_a = 3.4\text{x}10^{16}/\text{cm}^3$, $Q_{SS}/q = 4\text{x}10^{10}/\text{cm}^2$, $T_{ox} = 1000\text{Å}$, Al gate.

Inverter parameters are the same as those in Section 5.3.

Fig. 6.9 Transfer Characteristics of a Saturated-Load Inverter

6.3.4 Maximum Fan-In Number in a Logic Circuit

Another circuit consideration influenced by the leakage current effect is the maximum number of logic inputs for a NAND circuit. This is especially important in CMOS circuits. As shown in Fig. 6.10, when the output is low (all the input voltages are high), leakage currents of P-channel devices all flow into output node V_o. At the drain of Q_{1_n}, drain-to-substrate and channel-to-substrate leakage currents are flowing into the substrate. The difference current ΔI_1 between total P-channel leakage current ($n\ I_p$) and leakage current at the drain of Q_{1_n} flows through Q_{1_n} and induce voltage drop V_{oL_1}. Similarly, the difference current between ΔI_1 and the leakage current at the source of Q_{1_n} and drain of Q_{2_n} flows through Q_{2_n} to induce another voltage drop V_{oL_2}, etc. Transistor Q_{1_n} has the largest amount of current and Q_{n_n} has the least amount of current flowing. Also due to the source body effect, Q_{1_n} has the highest and Q_{n_n} has the lowest value of threshold voltage. As a result, among N-channel transistors V_{oL_1} is the highest and V_{oL_n} is the lowest voltage drop.

$$\Delta I_1 = nI_p - I_{1n}\Big|_d \tag{6.10}$$

$$\Delta I_1 - I_{sub_{1_n}} = \beta_n[(V_{in} - V_{th_{1_n}}) - \frac{1}{2} V_{oL_1}]V_{oL_1} \tag{6.11}$$

$$\Delta I_2 = \Delta I_1 - I_{sub_{1_n}} - I_{1n}\Big|_s - I_{2n}\Big|_d \tag{6.12}$$

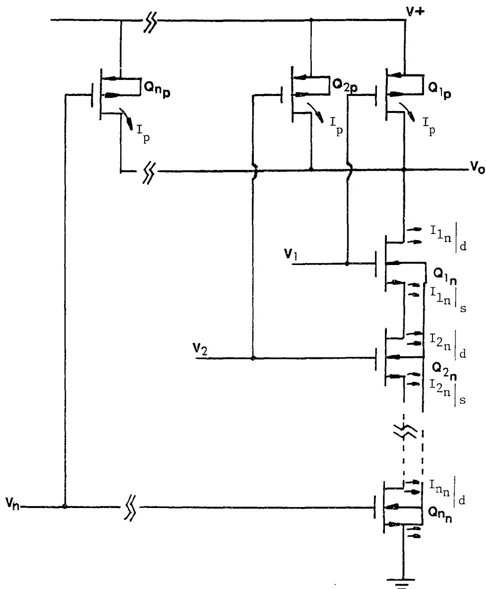


Fig. 6.10 A CMOS NAND Gate Circuit Showing All Important Leakage Components When All the Inputs Are High.

$$\Delta I_2 - I_{sub2_n} = \beta_n[(V_{in} - V_{th2_n}) - \frac{1}{2} V_{oL2}] V_{oL2} \quad (6.13)$$

$$\vdots$$

$$\Delta I_n = \Delta I_{n-1} - I_{sub_{n-1_n}} - I_{n-1_n}\Big|_s - I_{n_n}\Big|_d \quad (6.14)$$

$$\Delta I_n - I_{sub_{n_n}} = \beta_n[(V_{in} - V_{th_{n_n}} - \frac{1}{2} V_{oLn}] V_{oLn} \quad (6.15)$$

and a necessary operating condition is

$$V_{oL1} + V_{oL2} + V_{oL3} + \ldots + V_{oLn} < V_{th(T_{max})} \quad (6.16)$$

This will limit the maximum number of logic inputs to a NAND gate. As an example, for a particular CMOS process with the inverter characteristics as shown in Fig. 5.21a, n is found to be less than 3.

The situation is different in a CMOS NOR gate circuit. As shown in Fig. 6.11, the substrate-to-drain leakage current of Q_{2_p} cancels the substrate-to-source leakage current of Q_{1_p}, (it is a substrate-P^+-P^+-substrate structure). For the same reason, substrate-to-drain leakage current of Q_{3_p} cancels the substrate-to-source leakage current of Q_{2_p}, etc. The only important leakage current from P-channel devices is the

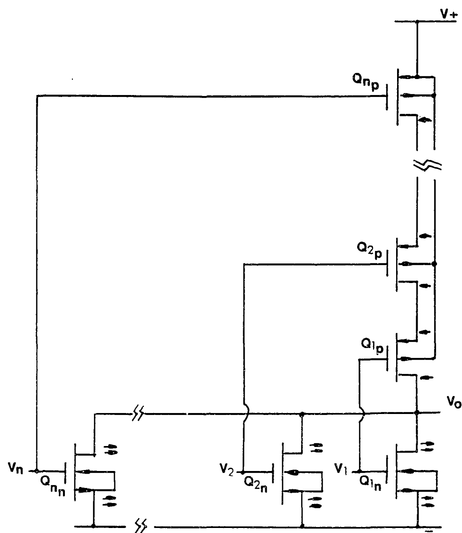


Fig. 6.11 A CMOS NOR Gate Circuit Showing All Important Leakage Components When at Least One of the Inputs Are High

substrate-to-drain leakage current of Q_{1p}. Furthermore, the N-channel devices have zero source body effect and if more than one input is high, leakage current at the drain of Q_{1p} is divided among the high voltage input devices. This induces only a small V_{oL} value at the output node.

It is then clear that in CMOS logic circuits, the NOR gate is a better choice than the NAND gate for high-temperature circuit operation. The maximum fan-in number will be determined by current driving and sinking capabilities instead of leakage current density.

6.3.5 Thermal/Radiation Latch-Up in CMOS Circuit

CMOS latch-up is a commonly encountered problem, especially in a radiation or high-temperature environment. Latch-up is a high-current, low-voltage state of positive feedback PNPN device action due to forward-biased P-N junctions. It is caused by voltage surges, photo current, radiation damage or leakage currents.

(1) Voltage Surge

As shown in Fig. 6.12, a voltage surge at the output of this CMOS inverter V_o such that $V_o \geq V_{DD} + V_{bi}$ will forward bias the p^+-N substrate junction where V_{bi} is the built-in voltage of that junction. For good expitaxial material the diffusion length is very large and injected minority carriers (holes in this case) diffuse through the N-substrate layer and are collected by the P-well diffusion at the p^+ node. Since V_{SS} is less than V_{DD}, a bipolar transistor exists with p^+ diffusion, N-substrate and P-well diffusion as emitter, base and collector, respectively. When the amount of injected carriers is large,

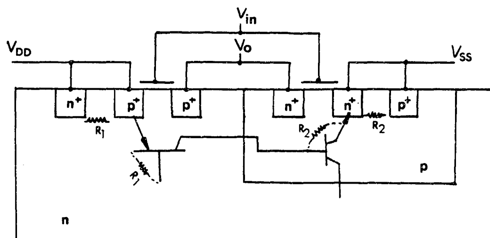


Fig. 6.12 Latch-Up in CMOS Due to Voltage Surge

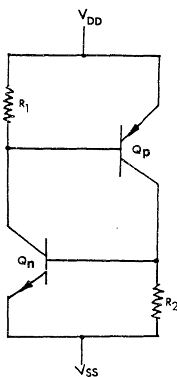


Fig. 6.13 Equivalent Circuit of the Latch

as they are collected at p^+ node a substantial drop is developed across R_2. Since the N-channel source and the P-well substrate (p^+ diffusion) are connected, the P-N junction at the N-channel source side tends to become forward-biased, the NPN transistor is turned on with N-substrate as collector. Electrons are then injected from n^+ diffusion region, diffused through P-well and are collected by N-substrate at n^+ diffusion region. This electron flow again develops a potential drop across R_1, causing the emitter of the PNP transistor to become more forward biased. This completes a positive feedback loop and a latch or positive PNPN device exists between V_{out} and V_{SS}. A circuit diagram showing these parasitic transistors, R_1 and R_2, is given in Fig. 6.13.

Similarly, a voltage surge at V_o such that $V_o \leq V_{SS} - V_{bi}$ causing a latch to exist between V_{DD} and V_o. Because V_o is an external pad, an incorrect power connection will lead to these types of latch.

Once one of the junctions is turned on, a regeneration process takes place. There is a minimum gain requirement for the transistors to sustain the positive feedback condition. This is shown in Fig. 6.14. Assume an initial current I_{in} is flowing through the base of the PNP parasitic transistor. By using the following equations

$$g_{mp} r_{be'_p} = \beta_p \tag{6.17}$$

$$g_{mn} r_{be'_n} = \beta_n \tag{6.18}$$

we find

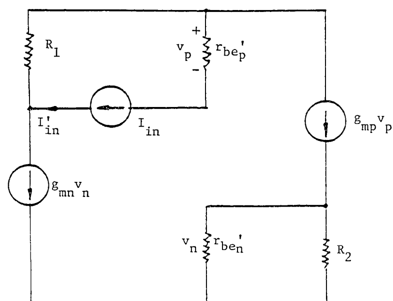


Fig. 6.14 Small Signal Circuit Diagram of Fig. 6.13 Showing the Feedback Loop

$$I'_{in} = \beta_p \beta_n \, I_{in} \frac{R_2}{R_2 + r_{be'_n}} - \frac{r_{be'_p} \, I_{in}}{R_1} \tag{6.19}$$

The current gain along the loop is

$$\frac{I'_{in}}{I_{in}} = \left(\beta_p \beta_n \frac{R_2}{R_2 + r_{be'_n}} - \frac{r_{be'_p}}{R_1} \right) \tag{6.20}$$

for a stable latch condition, $I'_{in}/I_{in} \geq 1$ or

$$\beta_p \beta_n \geq \left(\frac{R_2 + r_{be'_n}}{R_2} \right) \left[1 + \frac{r_{be'_p}}{R_1} \right] \tag{6.21}$$

This gives the minimum beta product of these parasitic transistors necessary to maintain a stable latch.

At high temperature it is easier for the latch to occur due to the following effects:

(i) The Junction Current Increases

The junction current is

$$I_E = I_S \exp\left(\frac{qV_{BE}}{kT}\right) = \frac{A_{EB} D_B \, n_i^2(T) q}{W_B \, N_B} \exp\left(\frac{qV_{BE}}{kT}\right) \tag{6.22}$$

where A_{EB} is the emitter-base area, D_B is the minority carrier diffusion coefficient in the base, V_{BE} is the forward base emitter bias voltage, W_B is the base width and N_B is the base concentration. Because of the

n_i^2 dependence, I_E increases with temperature. More minority carriers are injected into the base for a given voltage drop across the junction.

(ii) Resistances R_1 and R_2 increase with temperature.

(iii) V_{bi} decreases with temperature. For example, assume $N_{n+} = 1.0 \times 10^{19}/cm^3$, $N_{nepi} = 3 \times 10^{16}/cm^3$; then V_{bi} (25°C) = 0.84V and V_{bi} (328°C) = 0.5 volt. This makes the transistors easier to be turned on.

(iv) β_n and β_p increase with temperature due to the base-collector leakage current; making the loop gain of the regeneration process higher.

(2) Thermal Latch Up

As shown in Fig. 6.15, thermally generated carriers in a depletion region or in bulk silicon are collected by the P-well (at the p^+ diffusion) and by the source/drain of P-channel device. Furthermore, if avalanche multiplication occurs, multiplied holes in an N-channel device will flow through R_2, and multiplied electrons in a P-channel device will flow through R_1. They both help to develop enough potential to forward bias these junctions through R_1 and R_2. The analysis and latch condition is similar to that in Case (1).

It is also easier for this type of latch to occur at high temperature because of the following effects:

(i) R_1, R_2 increase

(ii) V_{bi} decreases

(iii) beta increases

(iv) thermally generated currents (g-r and diffusion current) are larger

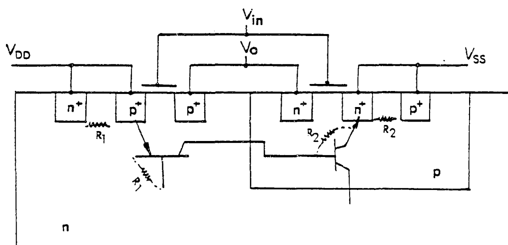


Fig. 6.15 Thermal Latch-Up in CMOS Circuit

(v) although the multiplication factor is smaller, multiplied current could be higher at high temperature due to the large value of seed current.

Efforts have been made to reduce the β values of these parasitic bipolar transistors by neutron radiation or gold doping; this is not suitable for high temperature devices because the reduction of lifetime will reduce the diffusion length and increase leakage current density. Another way to decrease β is to increase the Gummel number, $\int_0^{W_B} Q_B(x)\,dx$, of the parasitic transistor by using N^+ substrate, P^+ buried layer and an N-epi layer as shown in Fig. 6.16. This reduces β of these parasitic transistors as well as the diffusion leakage current, making it very desirable for high-temperature operation. But two more steps, buried layer and epi growth, are needed.

To eliminate the problem, one has to make a design change in the process. For example, one can use a guard ring structure with p^+ diffusion around the P-channel MOS device and tie it to the most negative supply voltage to suppress hole collection by the P-well, and n^+ diffusion around the N-channel MOS device and tie it to the most positive supply voltage to suppress electron collection by the N-substrate. Another good design is to merge n^+ and p^+ diffusion together to short out the transistors as shown in Fig. 6.17. Also, CMOS latch-up problems could be much reduced if only NMOS devices are used at the output of the circuits.

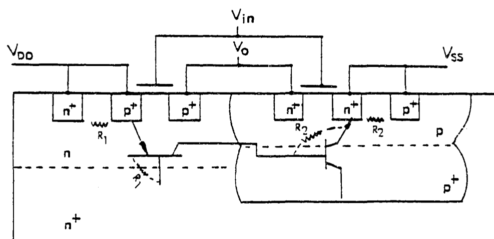


Fig. 6.16 Epi CMOS to Reduce Betas of the Parasitic Transistors and Also the Diffusion Leakage Current Density

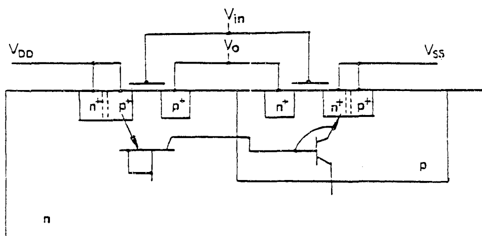


Fig. 6.17 Process Modification in CMOS to Eliminate R_1 and R_2

6.4 Ion Implanted Depletion-Load Devices

One reason to use a depletion device as an active load transistor in an inverter circuit is to utilize the high transconductance of the load transistor to increase the switching speed. For the inverter to operate properly, the load transistor has to be on at $V_{gs} = 0$ for all operating temperatures. For lightly doped P-substrate or shadow implant depletion type MOS devices, threshold voltage decreases with temperature, or the device is more "depletion type" at high temperature. For ion implanted depletion devices with implant depth larger than the substrate-to-implant-layer depletion width without any bias, the turn on voltage is the important concern. Turn on voltage is defined as the gate voltage required such that conduction charge at the source is greater than zero. From Appendix C, its value is found from

$$-Q_i + K_m \sqrt{V_S + \phi_B - V_{sub}} + K_n \left\{ \left[(V_S - V_g + V_{fb}) + \frac{K_n^2}{4C_{ox}^2} \right]^{\frac{1}{2}} - \frac{K_n}{2C_{ox}} \right\} = 0 \tag{6.23}$$

or

$$V_g \Big|_{\text{turn on}} = V_{fb} + V_S - \left(\frac{K_m}{K_n} \right)^2 [V_S + \phi_B - V_{sub}] + \frac{K_m \sqrt{V_S + \phi_B - V_{sub}}}{\left[C_{ox} / \left(1 + \frac{C_{ox}}{C_i} \right) \right]} - \frac{Q_i}{\left[C_{ox} / \left(1 + \frac{C_{ox}}{2C_i} \right) \right]} \tag{6.24}$$

where

$V_S \triangleq$ source voltage relative to ground

$$K_m \triangleq \sqrt{2\varepsilon_s q \frac{N_a N}{N_a + N}}$$

$N \triangleq N_d - N_a$ (net concentration at implant layer)

$$K_n \triangleq \sqrt{2\varepsilon_s qN}$$

$$\phi_B \triangleq 2 \frac{kT}{q} \ln \frac{N}{n_i} + \frac{kT}{q} \ln \frac{N_a}{n_i}$$

$V_{sub} \triangleq$ substrate voltage relative to ground

$C_{ox} \triangleq$ gate oxide capacitance

$$C_i \triangleq \frac{\varepsilon_s}{x_I}$$

$Q_i \triangleq qN\, x_I$ (net implanted charge)

At high temperature, the built-in voltage across the implant layer-to-substrate junction decreases, resulting in a decrease of depletion width from bottom of implant layer. Also, flat band voltage of the surface layer (N-type) decreases with temperature, resulting in a decrease of surface depletion width for the same gate bias. These factors make the amount of conduction charge in the implant layer larger at high temperature; i.e., the device remains in conductive state at high temperature if it is conductive at low temperature. Also, note

that the rate of decrease of turn on voltage is less compared to lightly doped or shadow implant depletion load devices.

Another concern about ion implanted depletion type devices is that we can pinch off the channel by gate voltage only. This is possible if channel pinchoff occurs before surface inversion as shown in Fig. 6.18a, assuming piecewise uniform impurity profile. The depletion width of substrate-to-implant layer junction in the implant layer x_n is

$$x_n = \sqrt{\frac{2\varepsilon_s}{q} \frac{N_a}{N(N_a+N)} (-V_{sub} + V_{bi})} \tag{6.25}$$

V_{bi} is the built-in voltage across the junction

$$V_{bi} = \frac{kT}{q} \ln \frac{N_a}{n_i} + \frac{kT}{q} \ln \frac{N}{n_i} \tag{6.26}$$

The maximum surface depletion width is

$$x_{d_{max}} = \sqrt{\frac{2\varepsilon_s}{q} \frac{1}{N} (-V_{sub} + 2|\phi_f|)} \tag{6.27}$$

To be able to pinch off the device by gate voltage, it is required that x_I is less than the summation of the above two distance

$$\begin{aligned} x_{I_{max}} &< x_n + x_{d_{max}} \\ &\cong \sqrt{\frac{2\varepsilon_s}{qN} (-V_{sub} + V_{bi})} \left[1 + \sqrt{\frac{N_a}{N_a + N}} \right] \end{aligned} \tag{6.28}$$

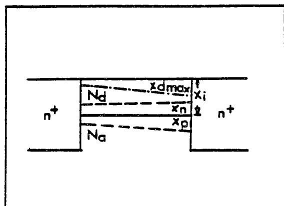


Fig. 6.18(a) Ion-Implanted Depletion Type NMOS with Box Profile Approximation

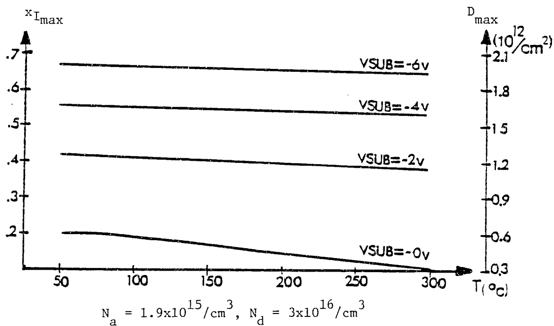


Fig. 6.18(b) Maximum Implant Depth versus Temperature to Ensure that the Device Could Be Pinched Off by Gate Bias; Also Shown Is the Corresponding Maximum Implant Dose

and the maximum implanted concentration per unit area is then

$$D_{max} = N_d \, x_{I_{max}} \qquad (6.29)$$

If the implant is too deep or too heavy such that channel inversion occurs earlier than channel pinch-off by gate voltage, we have to apply high drain voltage to pinch off the device at the drain, resulting in high drain saturation voltage. This has the disadvantages of reduction in signal swing, and reduction of output resistance in an inverter circuit.

At high temperature, V_{bi} and $|\phi_B|$ decrease, resulting in lower $x_{I_{max}}$ and D_{max}. Figure 6.18b shows the relation between $x_{I_{max}}$ and D_{max} versus temperature for an ion implanted N-channel depletion load device with $N_a = 1.9 \times 10^{15}/cm^3$ and $N_d = 3 \times 10^{16}/cm^3$ for different substrate voltages. It is shown that in order to pinch off the device with gate bias at high-temperature operation, shallower and light implant should be used.

6.5 Summary

In this chapter, we have studied various temperature limits of MOS devices and circuit operations. The important parameters for determining the maximum operating temperature include substrate concentration, oxide thickness, gate material, geometry dimensions, bias conditions, and leakage current densities.

The substrate concentration will determine the absolute maximum operating temperature. All other factors, however, need to be taken

into account in proper design. Among these, leakage current densities is the most important consideration. High leakage current densities cause a maximum output voltage drop for PMOS and NMOS inverters, limit the maximum fan-in number and induce thermal latch-up in CMOS logic circuit. General guidelines and design rules have been given to increase the maximum temperature of operation.

As has been shown in this chapter, use of a non-zero substrate bias voltage increases the maximum temperature limit in general. However, it increases the electric field near source/drain junctions and the multiplication factor. It also requires lower substrate doping concentration for reasonable V_{th} at room temperature. Process modifications such as guard ring structure and epitaxial MOS permits use of relatively low substrate concentration in the active MOS region and keeps the leakage currents at acceptably low level.

For a typical MOS process, series degradation of device behavior in analog circuits occurs when the temperature is higher than 150°C (Section 5.2). For digital circuits that typical value is 250°C (Section 6.3.3). However, dielectric isolated (DI) CMOS process allows digital circuit operation at 350°C [Beason, Moore, Mohammed and Draper, 1981] and epitaxial MOS process is expected to permit circuit operation at this temperature too.

The maximum temperature limit can be increased by process or circuit modifications. For example:

(i) Increase the substrate concentration or oxide thickness to raise the temperature limit imposed by the threshold voltage

consideration (Section 6.2.1). But a relatively high room-temperature threshold voltage must be tolerated in this case.

(ii) Increase the substrate bias voltage to raise the temperature limit imposed by the existence of surface channel-to-substrate potential barrier.

(iii) Use an epitaxial process and thinner substrate material to reduce the problem of forward-biasing the source junctions.

(iv) Use an epitaxial process, or a guard-ring structure to reduce the effectiveness of leakage currents at high temperature.

The concern for leakage currents will be less important as the technology advances with reduction of leakage current density and junction area. Another maximum temperature limit independent of leakage current density exists when the thermal carrier concentration in the surface inversion layer is comparable to the gate-voltage-induced inversion layer carrier concentration. This is more serious for transistors with smaller current levels. Taking the ratio to be 1:10, the maximum temperature limit is found from

$$n_i\ (T_{max})\ x_{d_{max}} = Q_n(y) \cdot 0.1 \qquad (6.30)$$

A typical value of T_{max} is about 420°C for this limit.

CHAPTER 7

SCALING AT HIGH TEMPERATURES

7.1 Introduction

One of the features of MOS LSI and VLSI design is the scaling property. By adjustment of processing parameters and biasing voltages, one can predict MOS terminal behavior accurately as the device dimensions are reduced. In this chapter we examine scaling theory in detail, with emphasis on those parameters which do not scale linearly. These include junction leakage current, subthreshold leakage current, etc. Unfortunately, these parameters are of major concern at high temperatures. Therefore, high-temperature oriented scaling rules are devised and discussed.

7.2 Conventional Scaling Rules

Scaling of MOS devices is done by the following changes [Dennard, Gaensslen, Yu, Rideout, Bassous, and Leblanc, 1974]:

a) Linear dimensions are scaled down by a factor of K, ($L' = \frac{1}{K} L$)

b) Supply voltages are scaled down by a factor of K, ($V' = \frac{1}{K} V$)

c) Substrate doping level is scaled up by a factor of K, ($N_a' = K \quad N_a$)

An example is shown in Fig. 7.1.

If the above rules are used, then we expect the following changes:

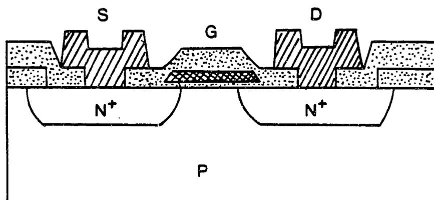


Fig. 7.1(a) A Polysilicon Gate NMOS Device

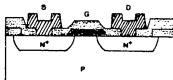


Fig. 7.1(b) Same Device after Scaling with K = 3.81

1. Maximum Depletion Layer Width

The maximum depletion layer width is

$$x_{d_{max}} = \sqrt{\frac{2\varepsilon_s(\phi_B - V_{sub})}{qN_a}} \quad (7.1)$$

which becomes after scaling

$$x'_{d_{max}} = \sqrt{\frac{2\varepsilon_s(\phi'_B - V'_{sub})}{q\ K\ N_a}} \quad (7.2)$$

where

$$\phi_B = 2\phi_f = \frac{2kT}{q}\ \ell n\ \frac{N_a}{n_i} \quad (7.3)$$

$$\phi'_B = \frac{2kT}{q}\ \ell n\ \frac{K\ N_a}{n_i} = \frac{2kT}{q}\ \ell n\ K + \phi_B \quad (7.4)$$

If $V'_{sub} = V_{sub}/K$, then

$$\frac{x'_{d_{max}}}{x_{d_{max}}} = \frac{1}{K}\left[1 + \frac{K\phi'_B - \phi_B}{\phi_B - V_{sub}}\right]^{\frac{1}{2}} \cong \frac{1}{K} \quad (7.5)$$

If we assume

$$\phi_B - V_{sub} >> (K-1)\phi_B$$

or if we choose V'_{sub} such that

$$\frac{\phi_B - V_{sub}}{K} = \phi'_B - V'_{sub} \tag{7.6}$$

then

$$\frac{x'_{d_{max}}}{x_{d_{max}}} = \frac{1}{K} \tag{7.7}$$

2. Threshold Voltage

The threshold voltage before scaling is

$$V_{th} = \frac{T_{ox}}{\varepsilon_{ox}} \left[-Q_{SS} - \sqrt{2\varepsilon_s q N_a (\phi_B - V_{sub})} \right] + \phi'_{MS} + \phi_B \tag{7.8}$$

where

$$\phi'_{MS} = \phi_{Mox} - \phi_{Sox} - \frac{E_g}{2q} - \frac{1}{2}\phi_B \tag{7.9}$$

After scaling, we obtain

$$V'_{th} = \frac{(T_{ox}/K)}{\varepsilon_{ox}} \left[-Q_{SS} - \sqrt{2\varepsilon_s q (K\,N_a)(\phi'_B - V'_{sub})} \right] + \phi''_{MS} + \phi'_B \tag{7.10}$$

If we choose V'_{sub} such that

$$\phi'_B - V'_{sub} = \frac{\phi_B - V_{sub}}{K}$$

then

$$\frac{V'_{th}}{V_{th}} = \frac{1}{K}\left[1 + \frac{(\phi''_{MS} - \phi'_{MS}/K) + (\phi'_B - \phi_B/K)}{V_{th}}\right] \qquad (7.11)$$

For aluminum gate N-channel devices or for a polysilicon gate with doped impurity opposite to that of the substrate,

$$\phi'_{MS} + \phi_B = \phi_{Mox} - \phi_{Sox} - \frac{E_g}{2q} + \phi_f$$

and

$$\phi''_{MS} + \phi'_B = \phi_{Mox} - \phi_{Sox} - \frac{E_g}{2q} + \phi'_f \qquad (7.12)$$

approach zero. (Table 7.1 shows this value for different concentration and temperature.) Therefore

$$\frac{V'_{th}}{V_{th}} = \frac{1}{K} \qquad (7.13)$$

3. Channel Current

It is shown in Section 5.2 that the square-law and linear approximations are good expressions for the drain characteristics if a pre-scaling factor is included.

(a) Linear Region

$$I_{chan} = \frac{\varepsilon_{ox}}{T_{ox}} \mu \left(\frac{W}{L}\right) \left[(V_g - V_{th}) - \frac{1}{2}(V_D - V_S)\right] (V_D - V_S) \qquad (7.14)$$

After scaling, this becomes

Table 7.1 Value of $\phi'_{MS} + \phi_B$ for Aluminum Gate and N^+ Polysilicon Gate MOS Devices

	Aluminum Gate			N^+ Polysilicon Gate		
N_a 1/cm^3 / T °C	10^{15}	10^{16}	10^{17}	10^{15}	10^{16}	10^{17}
25	-0.29	-0.23	-0.17	-0.23	-0.17	-0.11
135	-0.40	-0.32	-0.24	-0.34	-0.26	-0.18
212	-0.48	-0.39	-0.29	-0.42	-0.33	-0.23
300	-0.52	-0.46	-0.35	-0.46	-0.40	-0.29

$$I'_{chan} = \frac{\varepsilon_{ox}}{(T_{ox}/K)} \mu \left(\frac{W/K}{W/K}\right)\left[\frac{V_g}{K} - V'_{th} - \frac{1}{2}\left(\frac{V_D - V_S}{K}\right)\right]\left[\frac{V_D - V_S}{K}\right] \quad (7.15)$$

Since $V'_{th} \cong \frac{V_{th}}{K}$, we obtain

$$I'_{chan} = \frac{I_{chan}}{K} \quad (7.16)$$

(b) Saturation Region

$$I_{chan} = \frac{1}{2} \frac{\varepsilon_{ox}}{T_{ox}} \mu \left(\frac{W}{L}\right) (V_g - V_{th})^2 \quad (7.17)$$

After scaling, this becomes

$$I'_D = \frac{1}{2} \frac{\varepsilon'_{ox}}{(T_{ox}/K)} \mu \left(\frac{W/K}{L/K}\right)\left(\frac{V_g}{K} - V'_{th}\right)^2 \quad (7.18)$$

Since $V'_{th} \cong \frac{V_{th}}{K}$, we obtain

$$I'_{chan} = \frac{I_{chan}}{K} \quad (7.19)$$

4. Oxide Capacitance

The oxide capacitance is

$$C_{ox} = \frac{\varepsilon_{ox}}{T_{ox}} W L \quad (7.20)$$

After scaling, this becomes

$$C'_{ox} = \frac{\varepsilon_{ox}}{(T_{ox}/K)} \left(\frac{W}{K}\right)\left(\frac{L}{K}\right) \tag{7.21}$$

from which we obtain

$$\frac{C'_{ox}}{C_{ox}} = \frac{1}{K} \tag{7.22}$$

5. Transconductance

In the saturation region, the transconductance is

$$g_m = \left(\frac{\varepsilon_{ox}}{T_{ox}}\right)\mu \left(\frac{W}{L}\right) (V_g - V_{th}) \tag{7.23}$$

After scaling, this becomes

$$g'_m = \frac{\varepsilon_{ox}}{(T_{ox}/K)} \mu \left(\frac{W/K}{L/K}\right)\left(\frac{V_g}{K} - V'_{th}\right) \tag{7.24}$$

If $V'_{th} = V_{th}/K$, then we obtain

$$g'_m = g_m \tag{7.25}$$

6. Short-Channel-Effect Factor f

According to the trapezoidal depletion approximation in Section 2.4

$$f = 1 - \frac{r_j}{L}\left(\sqrt{1 + \frac{2x_{dmax}}{r_j}} - 1\right) \tag{7.26}$$

After scaling, this becomes

$$f' = 1 - \frac{(r_j/K)}{(L/K)}\left(\sqrt{1 + \frac{2x'_{d_{max}}}{(r_j/K)}} - 1\right) \tag{7.27}$$

because

$$x'_{d_{max}} = \frac{1}{K}\, x_{d_{max}} \tag{7.28}$$

Then we obtain

$$f = f'$$

This means that the short channel effect remains unchanged, a fact of fundamental importance in scaling theory.

7. Sheet Resistance, Line Delay and Power Consumption

Before scaling we have

$$R = \rho_s \frac{L}{W\ T}$$

$$\tau = CR$$

$$P = IV \tag{7.29}$$

After scaling we obtain

$$R' = \rho_s \frac{(L/K)}{(W/K)(T/K)} = KR$$

$$\tau' = C'R' = \tau$$

$$P' = I'V' = \frac{P}{K^2} \tag{7.30}$$

8. Breakdown Voltage V_{BV}

V_{BV} is found from the depletion approximation to be

$$\frac{1}{2} E_{crit} \quad W = V_{BV} + \phi_B \tag{7.31}$$

$$W = \sqrt{\frac{2\varepsilon_s(\phi_B + V_{BV})}{qN_a}} \tag{7.32}$$

where E_{crit} is the critical field strength at the onset of avalanche breakdown. The breakdown voltage is thus given by

$$V_{BV} = \frac{\varepsilon_s E_{crit}^2}{2qN_a} - \phi_B \tag{7.33}$$

After scaling we obtain

$$V'_{BV} = \frac{\varepsilon_s E_{crit}^2}{2q(K \quad N_a)} - \frac{2kT}{q} \ln K - \phi_B \tag{7.34}$$

It will normally be true that

$$V'_{BV} >> 2 \frac{kT}{q} \ln K$$

and

$$V'_{BV} >> \phi_B$$

so

$$\frac{V'_{BV}}{V_{BV}} = \frac{1}{K}$$

9. Punch-Through Voltage

This is an important parameter since for small dimensions the punch-through condition will be encountered prior to avalanche breakdown. Punch-through will occur when the depletion width from either source or drain end reaches L/2 where L is the channel length. Since

$$\frac{L}{2} = \sqrt{\frac{2\varepsilon_s (V_P + \phi_B)}{qN_a}} \tag{7.35}$$

We find that

$$V_P = \frac{qN_a L^2}{8\varepsilon_s} - \phi_B \tag{7.36}$$

After scaling, we obtain

$$V_P' = \frac{q(K \quad N_a)(L/K)^2}{8 \quad \varepsilon_s} - \phi_B' \tag{7.37}$$

$$\frac{V_P'}{V_P} = \frac{1}{K} + \frac{1}{V_P} \left[\left(\frac{1}{K} - 1\right)\phi_B - \frac{2kT}{q} \ell n \ K\right] \tag{7.38}$$

Again, because

$$V_P >> \left|\left(\frac{1}{K} - 1\right) \phi_B\right|$$

and

$$V_P >> 2 \frac{kT}{V_q} \ell n \ K$$

in most cases, it will be true that

$$\frac{V'_p}{V_p} = \frac{1}{K} \tag{7.39}$$

Next, we examine components that are important at high temperatures but do not scale linearly.

10. g-r Junction Leakage Current

The generation-recombination leakage current density within the depletion region is

$$J_{g-r} = \frac{qn_i}{2\tau_o} \sqrt{\frac{2\varepsilon_s [V + \phi_B]}{q\ N_a}} \tag{7.40}$$

If we assume τ_o is independent of substrate concentration, then after scaling we obtain

$$J'_{g-r} = \frac{qn_i}{2\tau_o} \sqrt{\frac{2\varepsilon_s [V/K + \phi'_B]}{q\ K\ N_a}} \tag{7.41}$$

$$\frac{J'_{g-r}}{J_{g-r}} = \frac{1}{K} \left[1 + \frac{K\phi'_B - \phi_B}{V + \phi_B} \right]^{\frac{1}{2}} \tag{7.42}$$

When $V >> (K-2)\phi_B$, we obtain

$$\frac{J'_{g-r}}{J_{g-r}} \cong \frac{1}{K} \tag{7.43}$$

and the g-r leakage current ratio is

$$\frac{I'_{g-r}}{I_{g-r}} = \frac{J'_{g-r}}{J_{g-r}} \frac{A/K^2}{A} \cong \frac{1}{K^3} \tag{7.44}$$

11. Diffusion Leakage Component

The diffusion leakage current density is

$$J_{diff} = \frac{qn_i^2}{N_a}\left(\frac{L_n}{\tau_n}\right) \tag{7.45}$$

where L_n and τ_n are average diffusion length and lifetime, respectively. We assume that these are independent of substrate concentration; then after scaling we find

$$J'_{diff} = \frac{qn_i^2}{K\ N_a}\left(\frac{L_n}{\tau_n}\right) \tag{7.46}$$

and

$$\frac{J'_{diff}}{J_{diff}} = \frac{1}{K} \tag{7.47}$$

and the diffusion leakage current ratio is

$$\frac{I'_{diff}}{I_{diff}} = \left(\frac{J'_{diff}}{J_{diff}}\right) \frac{(A/K^2)}{A} = \frac{1}{K^3} \tag{7.48}$$

12. Subthreshold Current

This is the channel diffusion current between weak inversion (b=1) and strong inversion where band bending is (b=2). This current is

$$I_{sub} = \left(\frac{W}{L}\right)\left(\frac{\varepsilon_s kT}{q^2 N_a}\right)^{\frac{1}{2}} \left(\frac{kT}{q}\ \mu\right) q\ \frac{n_i\ \exp[(b-1)u_f][1-\exp(-u_{ds})]}{[2(u_{sx} + |\ u_f\ | - 1)]^{\frac{1}{2}}} \quad (7.49)$$

where

$$b \triangleq \text{band bending factor}$$

$$u_f \triangleq \frac{q}{kT}\ \phi_f$$

$$u_{ds} \triangleq \frac{q}{kT}\ (V_D - V_S)$$

$$u_{sx} \triangleq \frac{q}{kT}\ (V_S - V_{sub}) \quad (7.50)$$

The value of b is obtained by solving the following equation:

$$V_g = \phi'_{MS} - \frac{Q_{SS}}{\varepsilon_{ox}}\ (T_{ox}) + b\phi_f + \frac{T_{ox}}{\varepsilon_{ox}}\ [2\varepsilon_s\ qN_a(b\phi_f - V_{sub})]^{\frac{1}{2}} \quad (7.51)$$

Maximum subthreshold leakage current occurs at the onset of strong inversion when b = 2, or $V_g = V_{th}$.

$$V_{th} = \phi'_{MS} - \frac{Q_{SS}(T_{ox})}{\varepsilon_{ox}} + 2\phi_f + \frac{T_{ox}}{\varepsilon_{ox}}\ [2\varepsilon_s\ qN_a\ (2\phi_f - V_{sub})]^{\frac{1}{2}} \quad (7.52)$$

After scaling, at strong inversion we find

$$V'_{th} = \phi''_{MS} - \frac{Q_{SS}(T_{ox}/K)}{\varepsilon_{ox}} + b'\phi'_f + \frac{(T_{ox}/K)}{\varepsilon_{ox}} [2\varepsilon_s q(KN_a)(b'\phi'_f - V'_{sub})]^{1/2} \tag{7.53}$$

If we choose V'_{sub} such that

$$\phi'_B - V'_{sub} = \frac{\phi_B - V_{sub}}{K}$$

and assume

$$\phi'_{MS} + 2\phi_f = \phi''_{MS} + 2\phi'_f = 0$$

as in Eq. (7.12), then

$$V'_{th} = \frac{1}{K} V_{th}$$

$$b' = b = 2 \tag{7.54}$$

Note that after scaling we have the same band bending factor at the threshold point. Now we neglect $\exp(-u_{ds})$ and approximate $[2(u_{sx} + |u_f| - 1)]$ by $[2(u_{sx} + |u_f|)]$. Then we obtain (for b=2)

$$I_{sub} = \left(\frac{W}{L}\right)\left(\frac{\varepsilon_s kT}{q^2 N_a}\right)^{1/2} \left(\frac{kT}{q} \mu\right) q \frac{n_i \exp|u_f|}{[2(u_{sx} + |u_f|)]^{1/2}} \tag{7.55}$$

After scaling we find

$$I'_{sub} = \left(\frac{W/K}{L/K}\right)\left(\frac{\varepsilon_s \; kT}{q^2 K N_a}\right)^{1/2}\left(\frac{kT}{q}\,\mu\right) q \; \frac{n_i \; \exp|u'_f|}{\left[2(u'_{sx} + |u'_f|)\right]^{1/2}} \tag{7.56}$$

$$\frac{I'_{sub}}{I_{sub}} = \frac{1}{\sqrt{K}} \; \frac{\exp \frac{q}{kT}\,\phi'_f}{\exp \frac{q}{kT}\,\phi_f} \left[\frac{\frac{q}{kT}\,(-V_{sub} + \phi_f)}{\frac{q}{kT}\,(-V'_{sub} + \phi'_f)}\right]^{1/2} \tag{7.57}$$

But note that

$$\begin{aligned} \frac{\exp \frac{q}{kT}\,\phi'_f}{\exp \frac{q}{kT}\,\phi_f} &= \exp\left[\frac{q}{kT}\,(\phi'_f - \phi_f)\right] \\ &= \exp\left[\frac{q}{kT}\;\frac{kT}{q}\,\ell n \; K\right] \\ &= \exp\;(\ell n \; K) = K \end{aligned} \tag{7.58}$$

Now if we choose V'_{sub} such that

$$\phi'_B - V'_{sub} = \frac{\phi_B - V_{sub}}{K} \quad ,$$

then we obtain

$$\left[\frac{\frac{q}{kT}\,(-V_{sub} + \phi_f)}{\frac{q}{kT}\,(-V'_{sub} + \phi'_f)}\right]^{1/2} = \left[K + \frac{K\phi'_f - \phi_f}{(-V'_{sub} + \phi'_f)}\right]^{1/2} \cong \sqrt{K} \tag{7.59}$$

Finally we have

$$\frac{I'_{sub}}{I_{sub}} \cong \frac{1}{\sqrt{K}} K \sqrt{K} = K \tag{7.60}$$

It is important to note that the subthreshold current scales up K times; this makes small-dimension MOS devices very unattractive for high-temperature applications.

7.3 High-Temperature Oriented Scaling Rules

To avoid this, we note that the increase of subthreshold current is mainly due to the increase of the substrate doping concentration. If we keep the substrate doping concentration the same, scale down linear dimensions and supply voltages, but maintain the same substrate bias voltage such that

$$\phi'_B - V'_{sub} = \phi_B - V_{sub}, \ \phi'_B = \phi_B ,$$

then the above parameters change as follows:

1. Maximum depletion width

$$\frac{x'_{dmax}}{x_{dmax}} = 1 \tag{7.61}$$

2. Threshold voltage

$$\frac{V'_{th}}{V_{th}} = \frac{1}{K} \left[1 + \frac{(\phi''_{MS} - \phi'_{MS}/K) + (\phi'_B - \phi_B/K)}{V_{th}}\right] \cong \frac{1}{K} \tag{7.62}$$

3. Channel current

$$\frac{I'_{chan}}{I_{chan}} = \frac{1}{K} \quad \text{in linear region}$$

$$\frac{I'_{chan}}{I_{chan}} = \frac{1}{K} \quad \text{in saturation region} \tag{7.63}$$

4. Oxide capacitance

$$\frac{C'_{ox}}{C_{ox}} = \frac{1}{K} \tag{7.64}$$

5. Transconductance

$$\frac{g'_m}{g_m} = 1 \tag{7.65}$$

6. Short-channel-effect factor f

$$f' = 1 - \frac{(r_j/K)}{(L/K)}\left(\sqrt{1 + \frac{2x'_{dmax}}{(r_j/K)}} - 1\right)$$

$$f = 1 - \frac{r_j}{L}\left(\sqrt{1 + \frac{2x_{dmax}}{r_j}} - 1\right)$$

$$f' < f \tag{7.66}$$

This shows that the short channel effect is more important with the high temperature scaling rules.

7. Sheet resistance, line delay and power consumption

$$\frac{R'}{R} = K$$

$$\frac{\tau'}{\tau} = 1$$

$$\frac{P'}{P} = \frac{1}{K^2} \tag{7.67}$$

8. Breakdown voltage

$$\frac{V'_{BV}}{V_{BV}} = \left(\frac{\varepsilon_s E_{crit}^{\;2}}{2qN_a}\right) \Big/ \left(\frac{\varepsilon_s E_{crit}}{2qN_a}\right)^2 = 1 \tag{7.68}$$

9. Punch-through voltage

$$\frac{V'_P}{V_P} = \left[\frac{qN_a (L/K)^2}{8\varepsilon_s}\right] \Big/ \left[\frac{qN_a L^2}{8\varepsilon_s}\right] = \frac{1}{K^2} \tag{7.69}$$

10. g-r junction leakage current

$$\frac{J'_{g-r}}{J_{g-r}} = \sqrt{\frac{V' + \phi'_B}{V + \phi_B}} = \sqrt{\frac{V/K + \phi_B}{V + \phi_B}} = \frac{1}{K}\left[\frac{\phi_B - \frac{\phi_B}{K}}{V + \phi_B}\right] \approx \frac{1}{\sqrt{K}} \tag{7.70}$$

$$\frac{I'_{g-r}}{I_{g-r}} = \frac{1}{\sqrt{K}} \quad \frac{1}{K^2} = \frac{1}{K^{5/2}} \tag{7.71}$$

11. Diffusion leakage current

$$\frac{J'_{diff}}{J_{diff}} = \frac{(qn_i^2/N_a)}{(qn_i^2/N_a)} \frac{(L_n/\tau_n)}{(L_n/\tau_n)} = 1$$

$$\frac{I'_{diff}}{I_{diff}} = 1 \cdot \frac{1}{K^2} = \frac{1}{K^2} \tag{7.72}$$

12. Subthreshold current

$$\frac{I'_{sub}}{I_{sub}} = \frac{\exp\left(\frac{q}{kT}\phi_f\right)}{\exp\left(\frac{q}{kT}\phi'_f\right)} \left[\frac{\frac{q}{kT}(-V_{sub} + \phi_f)}{\frac{q}{kT}(-V'_{sub} + \phi_f)}\right]^{1/2} \cong 1 \tag{7.73}$$

These scaling rules make the high temperature component I_{sub} independent of scaling but reduce the punch-through voltage by $1/K^2$ times. Thus they are useful for relatively long channel devices operated at high temperatures.

7.4 Summary

In summary, conventional scaling rules designed for room temperature operation often ignore the importance of scaling of subthreshold current. This poses serious problems at high temperature. High-temperature oriented scaling rules keep the subthreshold current unchanged, but the punch-through voltage is reduced to 1/K times the value conventional scaling rules predict. Therefore, relatively long devices or small bias voltages must be used when high-temperature operation is required.

CHAPTER 8

EXPERIMENTAL DATA FOR LEAKAGE CURRENT REDUCTION

8.1 Introduction

In Section 3.6, epitaxial MOS and guard-ring MOS devices have been suggested as modified processes for high-temperature operation. Although the guard-ring structure is more effective in reducing the minority carrier collection at the source/drain regions, it requires a somewhat complicated process. Epitaxial MOS devices, on the other hand, require only one additional processing step and thus are easier to fabricate and are more cost-effective.

Epitaxial diodes have been fabricated which simulate the source/drain-to-substrate junctions of epitaxial MOS devices. Leakage current measurements are discussed in this chapter. Results show that reduction of the diffusion leakage current is indeed observed when a heavily doped p^+ substrate is used.

8.2 Background and Process Outline

The junction leakage currents in MOS devices include aource-to-substrate, drain-to-substrate and surface channel-to-substrate leakage currents. At high temperatures (> 135°C typically), diffusion leakage components are dominant. As was shown in Section 3.6, epitaxial MOS devices have lower diffusion leakage current, therefore less overall leakage current at high temperature.

However, completed epitaxial MOS devices are not necessary to verify the leakage current reduction effect. An epitaxial diode can simulate the source/drain-to-substrate junction with much simpler processing requirements.

Epitaxial MOS structures for high-temperature operation require a heavily-doped substrate ($N_a \cong 10^{18}/cm^3$) upon which is grown a 10 μm layer of lightly doped material ($N_a \cong 10^{16}/cm^3$). Unfortunately, such heavily-doped substrates are not used in the semiconductor industry, and it was not possible to obtain them for this experiment. Instead, ordinary substrates with $N_a \cong 2 \times 10^{16}/cm^3$ and $\langle 100 \rangle$ surface orientation were used, primarily for mechanical support. A thick epitaxial layer (40 μm) was then grown with $N_a \cong 10^{18}$ to simulate a heavily-doped substrate of thickness comparable to the minority carrier diffusion length. (At 300°C the diffusion length is approximately 60 μm for this material.) On a first group (a) of wafers, a 10 μm layer with $N_a \cong 10^{16}/cm^3$ was then grown for fabrication of the epitaxial transistors. On a second group (b) of wafers, the 10 μm layer was grown directly atop the ordinary substrate. As was noted in Chapter 3, the effectiveness of the epitaxial layer in reducing leakage current depends upon the material below the layer having a high impurity concentration. Use of the two groups of wafers permits comparision of the effectiveness of the heavily-doped material.

Following epitaxial growth, diodes were fabricated in both groups of wafers to simulate source/drain junctions, which are the major contributors to diffusion leakage current.

Major processing steps are listed in Appendix D.

8.3 Measured Leakage Currents for Epitaxial Diodes

Figure 8.1 shows the measured leakage currents versus temperatures for diodes in group a having both 40 μm and 10 μm layers, and in group b having only the 10 μm layer. The diode junction area is $1.08 \times 10^{-4} cm^2$ and the reverse bias is 2 volts, for all cases.

At room temperature, (a) has 60 times higher leakage current (12 nA versus 0.21 nA). This is due to the fact that the relatively poor quality of the 40 μm p^+ epi layer deteriorates the top layer, resulting in lower lifetime and higher g-r leakage current density. It is interesting to note that for $N_a'/N_a = 50$, $L_n = 60$ μm, and $d_{epi} =$ 8.5 μm, the predicted reduction ratio for diffusion leakage currents is 7.9 (Fig. 3.11). This is domonstrated by the low value of photocurrent for case (a) when an intensive light source is shown on the wafer

$$[I_{ph} = qG_L (L_p + L_n) A]$$

However, the rate of increase of leakage current decreases at high temperature for diode (a) because of the reduction of minority carrier concentation in the p^+ regions. At 300°C, the leakage current density for (a) is 1/5 of that in diode (b).

For high value of reduction ratio by this method, thinner tip layer epi, heavier p^+ doped substrate and good quality of both p^+ substrate and p layer are required. Leakage currents at 25°C for epitaxial diodes are expected to be comparable to those of ordinary

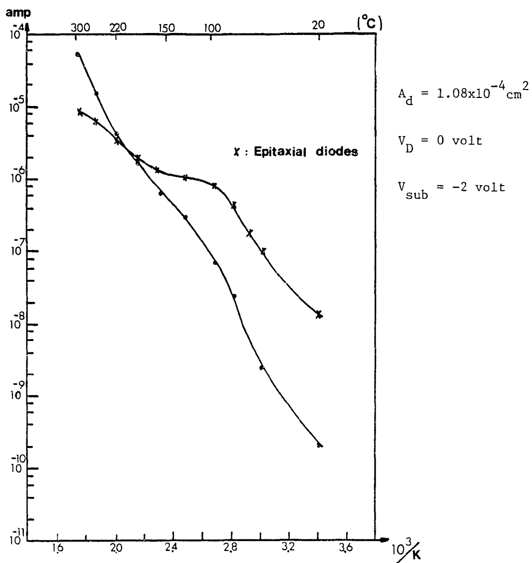


Fig. 8.1 Leakage Currents versus Temperature of Epitaxial (a) and Standard (b) Diodes

diodes, but at high temperature analysis shows that leakage currents for the former should be as much as 30 times lower than for the latter. For even higher leakage current reduction, one has to use guard-ring structure of dielectric isolation (DI) process.

CHAPTER 9

CONCLUSIONS AND RECOMMENDATIONS

9.1 Introduction

The main objectives of this dissertation were:

1. Identify all factors which influence high-temperature behavior of MOS devices;

2. Examine the maximum operating temperature T_{max} for MOS devices;

3. Investigate necessary process and circuit design modifications to optimize high-temperature behavior of MOS devices.

These objectives have been explored and a brief summary of the results is given in this chapter.

The junction leakage current density is 7-8 orders of magnitude larger at 300°C than at room temperature. This will have adverse effect on the operation of memory cells, logic circuits, gate protection diodes and many other circuits. Means to reduce the junction leakage current density have been explored. Experimental results for epitaxial MOS devices show 5 times lower leakage current density at 300°C than conventional structures; theoretical analysis shows that with good epitaxial material and good control of diffusion depth, the reduction can be as much as 20 times. Guard-ring structured MOS devices are expected to have as much as 60 times reduction of leakage current density at 300°C.

Maximum operation temperature has been investigated with regard to both device physics and circuit performance. Leakage current densities will doubtless decrease as technology advances and new process modifications are introduced, and thus subthreshold conduction current will become more important. Typical values of the maximum temperature limited by circuit performance is 350°C. Maximum operating temperature due to device physics limitations is determined by the substrate doping concentration near the surface and by the ratio of thermally generated electrons to gate-voltage-induced surface channel electrons. For typical MOS processes, both cases expect to have maximum operating temperature below 420°C.

In addition to the efforts to reduce the leakage current density itself by process modifications, circuit design and other techniques have been developed to reduce the effects of high leakage current densities at high temperature. In this respect, geometry layout change, charge-pumping and diode-compensation schemes have been investigated. For a typical N-channel saturated-load inverter considered in Chapter 5, simulated data show that the noise margin has been increased from 0.8 volts to 1.5 volts and 1.2 volts by using diode-compensation scheme and geometry layout modification, respectively.

One of the most important aspects of device design is the scaling of device size as the technology advances and achievable dimensions continue to become smaller. This leads to increased complexity of circuit functions that can be realized. It is of utmost importance to determine how the results of scaling influence device behavior at high

temperatures. If conventional scaling rules are used, the high-temperature subthreshold current increases by a factor of K, where K is the scaling factor. Therefore, scaling rules have been developed which keep the high-temperature subthreshold current independent of scaling. Other parameters which are important at high temperatures are not degraded by these scaling rules.

The variation of threshold voltage has been correctly modeled. A method to reduce the threshold voltage sensitivities with temperature by using temperature-dependent substrate bias voltages has been developed. This method permits the matching of threshold voltages between room temperature and a given higher temperature (for example, 300°C), with minimum variation of threshold voltage between these temperatures.

There is not sharp demarcation between the cutoff condition and the formation of an inversion channel in an enhancement-mode device. In this "subthreshold" region, the subthreshold drain current increase 7 order of magnitude between 25°C and 300°C. A model had been formulated to calculate the widening of the gate voltage range over which subthreshold current flows; it shows that this gate voltage range increases by a factor of 2.5 from 25°C to 300°C, for typical NMOS processes. This model then permits us to accurately simulate MOS inverter circuits especially near their threshold transition points.

9.2 Original Contributions of the Dissertation

The original contributions of this disseration can be summarized as follows:

1. Epitaxial MOS devices have been developed for high-temperature circuit applications. Reduction of leakage current density was demonstrated.

2. A guard-ring structure has been developed to further reduce the leakage current density at high temperature.

3. Various maximum temperature limits on device characteristics and circuit performance have been investigated and guidelines have been given to increase the temperature limit to its absolute maximum value.

4. Diode-compensation, geometry layout change and charge-pumping techniques have developed to improve high-temperature performance of saturated-load, unsaturated-load and depletion-load inverters.

5. High-temperature-oriented scaling rules have been developed.

6. A new temperature-dependent substrate voltage feedback bias scheme has been developed to stabilize the threshold voltage across a given temperature range.

7. A complete DC model with all leakage currents incorporated has been developed to simulate N-channel inverters.

8. A method to achieve zero-temperature-coefficient (ZTC) currents for analog circuit applications has been developed.

9. Demonstration that depletion-load inverters have the best high-temperature performance of all N-channel inverters. Demonstration that CMOS NOR gates have superior high-temperature performance to that of CMOS NAND gates.

10. A new diode-protection scheme has been suggested to reduce the problem of input voltage shift in CMOS inverter. Also, CMOS process

modifications have been adapted for latch-up-resistant CMOS inverters.

11. A model to calculate the threshold voltage of *in situ* ion-implanted enhancement type MOS devices has been developed.

9.3 Recommendations for Future Work

Future work is recommended in the following areas:

1. Experimental verification of leakage current reduction for the guard-ring structure should be carried out.

2. Various gettering techniques should be investigated to examine their effects on the reduction of leakage current density.

3. Other methods to gain better control of the junction areas of devices and the leakage current densitities should be investigated for the purpose of leakage current cancellation.

4. Other latch-up resistant methods for CMOS circuits should be investigated.

5. Effects of surface state ionization on threshold voltage variation should be investigated in more detail.

6. MOS devices with thin oxide ($\leq$ 500Å) and with other insulator material (i.e., Si_3N_4) should be investigated for their high-temperature applications.

7. Analog MOS circuits at high temperature should be further investigated.

APPENDIX A

THE USE OF DARE P PROGRAM

DARE P is a FORTRAN-based continous system simulation language developed at The University of Arizona. It can be used to solve rational differential equations.

The equation we are interested in is of the following type [Eq. (3.27)]:

$$\frac{dV}{dy} = \frac{AA \cdot y + GG}{BB \cdot \left[(CC-V) - DD\ [EE \cdot (FF + V)]^{\frac{1}{2}} \right]}$$

where

$AA = J_L(T)W$

$BB = \mu_n\ W\ C_{ox}$

$CC = V_g - V_{fb} - 2\phi_f$

$DD = \mu_n\ W$

$EE = 2\varepsilon_s\ qN_a$

$FF = 2\phi_f$

$GG = I_s$

with the initial condition

$V = 0$ at $y = 0$

In this program, potential distribution, electric field distribution, and surface charge distribution are all printed. Following are the card format:

```
JOB CARD
PW,---.
CALL DAREP,*EEDEPT.
7/8/9
$D1
     TIM=T
     AA=---
     BB=---
     CC=---
     DD=---
     EE=---
     FF=---
     GG=---
    BBB=---
   VDOT=(AA*TIM+GG)/(BB*(CC-V)-DD*(EE*(FF+V))**0.5)
     V.=VDOT
    QNY=-BBB*(CC-V)+((EE*(FF+V))**0.5)
END
*INITIAL CONDITIONS
     V=0.
     TMAX=---
     NPOINT=---
```

```
END
LIST V
LIST VDOT
LIST QNY
END
6/7/8/9
```

TMAX is the maximum number of y (i.e., the channel length). NPOINT is the number of points along y direction that the simulated data are printed.

The surface charge density is calculated as:

```
Qn(y) = -BBB*(CC-V)+((EE*(FF+V))**0.5)
```

where

$$BBB \triangleq C_{ox} = \frac{\varepsilon_{ox}}{T_{ox}}$$

APPENDIX B

INVERTER SIMULATION PROGRAM

For the saturated-load inverter discussed in Section 5.4, a separate program was written to simulate its performance at 25°C, 150°C, and 300°C, respectively. The advantage of this program is that important leakage current components such as g-r leakage current, diffusion leakage current, subthreshold conduction current and channel-to-substrate leakage current can be calculated separaterly and summed appropriately. This is not possible with the SPICE or SPICEG program. Figure B.1 shows the flow diagram of this program; the important parameters are:

V_{in} = input voltage of the inverter

ΔV_{in} = incremental input voltage of the inverter

V_o = out voltage of the inverter

ΔV_o = incremental output voltage of the inverter

V_{DD} = most positive power supply of the inverter

V_{SA1} = drain saturation voltage of the driver device (when the inverted charge at the drain of the driver is zero)

V_{SA2} = source saturation voltage of the load device (when the inverted charge at the source of the load transistor is zero)

$I_{L_{sat}}$ = saturation current of the load transistor

$I_{d_{sat}}$ = saturation current of the driver transistor

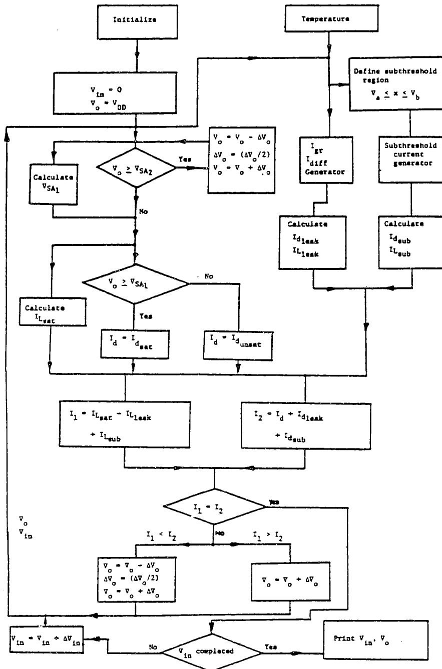


Fig. B.1 Flow Diagram of the Saturated-Load Inverter

$I_{d_{unsat}}$ = unsaturated current of the driver transistor

$I_{L_{sub}}$ = subthreshold current of load device

$I_{d_{sub}}$ = subthreshold current of the driver device

$I_{L_{leak}} = I_{L_{s-sub}} + 0.5\ I_{L_{chan-sub}}$

$I_{d_{leak}} = I_{d_{d-sub}} + 0.5\ I_{d_{chan-sub}}$

$I_{L_{s-sub}}$ = total g-r and diffusion leakage currents between source and substrate of the load transistor

$I_{d_{d-sub}}$ = total g-r and diffusion leakage currents between drain and substrate of the driver transistor

$I_{L_{chan-sub}}$ = channel-to-substrate leakage current of the load transistor

$I_{d_{chan-sub}}$ = channel-to-substrate leakage current of the driver transistor

```
      PROGRAM MOS(INPUT,OUTPUT,PLOT,TAPE6=OUTPUT,TAPE99=PLOT)

      DIMENSION TYPE(2)

C     COMMON BLOCK /IGFET/

      COMMON /IGFET/ TYP,T,DT,BFP,VFB,WFD,WFUN,QSSQ,ANA,AND,ANI,
     1CO,AL,Z,C,XJ,F,D,AJUNC,NT,SIGMA

      COMMON /MOSFET/ VTH,VDS
        COMMON /INVERT/ ZONL
      COMMON /CONST/ Q,ES,H,AKB,PI,CHI,MO

      COMMON /ATEMP/ TEMP(51)
      COMMON /ARRAY1/ PSI(51),FBV(51),FF(51),AK(51)
      COMMON /VDSAT/ VDSAT(51),VT(51)


      LOGICAL INV
      REAL MO,NT,JL
      INTEGER TYP

C     DEFINE FUNDAMENTAL CONSTANTS

      DATA Q,ES,H,AKB,PI,CHI,MO/1.602E-19,1.0625E-10,6.626E-34,1.38E-23,
     13.141592654,4.13,9.109E-31/

      DATA TYPE/1HN,1HP/

C     READ IN THE DEVICE PARAMETERS

      READ 1,TYP
      READ 3,SDL,DSDL
      READ 4,D,GL,Z,DSW,AJUNC
      READ 5,QSS
      READ 6,GATE,WFUN
      READ 8,XJ,BW
      READ 9,NT,SIGMA
      READ 4,GL1,GL2,Z1,Z2,VDD
      READ 19,INV

C     CALCULATE THE METALLURGICAL CHANNEL LENGTH
      AL=GL-(2.*XJ)

C     CALCULATE Z/L AND CO, AND CHANGE QSS,SDL,AND DSDL TO MKS UNITS
      ZONL=Z/AL
      CO=3.453E-11/(D*1.E-10)
      QSSQ=QSS*1.E+4

      ANA=1.E+6*SDL
      AND=1.E+6*DSDL

C     CALCULATE THE DRAIN JUNCTION AREA (IN CM2)
      IF(AJUNC.EQ.0.) AJUNC=(XJ*PI*(2.*XJ+DSW+Z)+DSW*Z)*1.E-8
C     INITIALIZE THE TEMPERATURE
      TC=25.
      DO 50 L=1,51
      TEMP(L)=TC
```

```
      T=TC+273.16

      CALL MOSFET
      PSI(L)=BFP
      FBV(L)=VFB
      VT(L)=VTH
      FF(L)=F
      VDSAT(L)=VDS

      UN=197.99*(T**(-1.5))
      UP=113.14*(T**(-1.5))
      IF(TYP.EQ.1) AK(L)=ZONL*UN*CO
      IF(TYP.EQ.2) AK(L)=ZONL*UP*CO

50    TC=TC+5.5

      PRINT 20,TYPE(TYP)
      PRINT 30,SDL,DSDL,XJ,D,CO,GATE,WFUN,QSS,AJUNC

1     FORMAT(I1)
3     FORMAT(2E10.2)
4     FORMAT(4F10.2,E10.2)
5     FORMAT(E10.2)
6     FORMAT(A10,F5.1)
8     FORMAT(2F5.2)
9     FORMAT(2E10.2)
10    FORMAT(1H1,1X,F6.4,5X,E14.7,5X,E14.7)
19    FORMAT(L1)
20    FORMAT(1H1,19X,37(1H-)/20X,37H*SUMMARY OF MOSFET DEVICE PARAMETERS
     +*/20X,37(1H-)//20X,*DEVICE TYPE = *,A1,*-CHANNEL*/)

30    FORMAT(20X,*SUBSTRATE IMPURITY CONCENTRATION   =*,E14.7,1X,*CM-3*/
     +20X,*DRAIN-SOURCE IMPURITY CONCENTRATION=*,E14.7,1X,*CM-3*/
     +20X,*DRAIN JUNCTION DEPTH              =*,F5.2,10X,*UM*//
     +20X,*OXIDE THICKNESS  =*,7X,F7.2,1X,*ANG.*/
     +20X,*OXIDE CAPACITANCE=*,E14.7,1X,*F/M2*//
     +20X,*GATE MATERIAL     =*,1X,A10/
     +20X,*GATE WORK FUNCTION=*,F6.3,1X,*VOLTS*,//
     +20X,*QSS=*,E14.7,1X,*CM-2*//
     +20X,*DRAIN JUNCTION AREA=*,1X,E14.7,1X,*CM2*//
     +20X,35(1H-)/)
      IF(INV) CALL INVER
4000  STOP
      END
      SUBROUTINE MOSFET
C-----------------------------------------------------------------------
C
C     SUBROUTINE MOSFET CALCULATES THE BULK FERMI POTENTIAL
C     FLATBAND VOLTAGE, AND THE THRESHOLD VOLTAGE
C     AT A SPECIFIED TEMPERATURE T.
C
C-----------------------------------------------------------------------

C     COMMON BLOCK /IGFET/

      COMMON /IGFET/ TYP,T,DT,BFP,VFB,WFD,WFUN,QSSQ,ANA,AND,ANI,
     1CO,AL,Z,C,XJ,F,D,AJUNC,NT,SIGMA

      COMMON /CONST/ Q,ES,H,AKB,PI,CHI,MO
```

```
      COMMON /MOSFET/ VTH,VDS
      COMMON /ATEMP/ TEMP(51)

      COMMON /ARRAY1/ PSI(51),FBV(51),FF(51)
      COMMON /VDSAT/ VDSAT(51),VT(51)

      INTEGER TYP

C     COMPUTE INTRINSIC CONCENTRATION
      EG=1.1557-((7.021E-4*T*T)/(T+1108.))
      ANI=(3.87298E+22)*(T**1.5)*EXP(-9.69210E-20/(T*AKB))

C     COMPUTE BFP AND VFB
      BFP=((AKB*T)/Q)*ALOG((ANA+SQRT(ANA**2.+4.*(ANI**2.)))/(2.*ANI))
      IF(TYP.EQ.2) BFP=-BFP
      WFD=WFUN-(CHI+BFP+(EG/(2.)))
      VFB=WFD-(Q*QSSQ)/CO

C     COMPUTE THE THRESHOLD VOLTAGE

      XDMAX=SQRT((2.*ES*(2.*ABS(BFP)+ABS(SSB)))/(Q*ANA))
      F=1.-(XJ/AL)*(SQRT(1.+(2.*XDMAX)/(XJ*1.E-6))-1.)
      QB=F*Q*ANA*XDMAX
      IF(TYP.EQ.1) VTH=VFB+2.*BFP+QB/CO+VS
      IF(TYP.EQ.2) VTH=VFB+2.*BFP-QB/CO+VS
      RETURN
      END
      SUBROUTINE MFET1(L)
C-----------------------------------------------------------------
C
C     SUBROUTINE MFET1 CALCULATES IDSAT AND VDSAT
C
C-----------------------------------------------------------------

C     COMMON BLOCK /IGFET/

      COMMON /IGFET/ TYP,T,DT,BFP,VFB,WFD,WFUN,QSSQ,ANA,AND,ANI,
     1CO,AL,Z,C,XJ,F,D,AJUNC,NT,SIGMA
      COMMON /CONST/ Q,ES,H,AKB,PI,CHI,MO

      COMMON /ATEMP/ TEMP(51)
      COMMON /ARRAY1/ PSI(51),FBV(51),FF(51)
      COMMON /VDSAT/ VDSAT(51),VT(51)

      INTEGER TYP


      C=((2./3.)*SQRT(2.*ES*Q*ANA))/CO
      A=((FF(L)**2.)*ES*Q*ANA)/(CO**2.)
      ALPHA=A*A+2.*A*ABS(VG-FBV(L)-VB)

C     CALCULATE VDSAT
      IF(TYP.EQ.1) VDS=VG-FBV(L)-2.*PSI(L)+A-SQRT(ALPHA)
      IF(TYP.EQ.2) VDS=VG-FBV(L)-2.*PSI(L)-A+SQRT(ALPHA)

      RETURN
      END
      SUBROUTINE INVER
```

```
C-----------------------------------------------------------------
C
C     SUBROUTINE INVER CALCULATE THE INVERTER CHARACTERISTICS
C
C-----------------------------------------------------------------

      DIMENSION VI(202),VOUT(202)
      DIMENSION AK1(51),AK2(51)

C     COMMON BLOCK /IGFET/
      COMMON /IGFET/ TYP,T,DT,BFP,VFB,WFD,WFUN,QSSQ,ANA,AND,ANI,
     1CO,AL,Z,C,XJ,F,D,AJUNC,NT,SIGMA
      COMMON /INVERT/ ZONL
      COMMON /CONST/ Q,ES,H,AKB,PI,CHI,MO

      COMMON /ATEMP/ TEMP(51)
C     COMMON BLOCK /ARRAY1/
      COMMON /ARRAY1/ PSI(51),FBV(51),FF(51),AK(51)
      COMMON /VDSAT/ VDSAT(51),VT(51)

      INTEGER TYP
      REAL JL

19    FORMAT(20X,E10.2,20X,E10.2)
21    FORMAT(4(10X,E10.3))
38    FORMAT(40X,4F20.9)
39    FORMAT(20X,3H*A*)

C     FIND VOMAX(N,L)

      AL1=GL1-2.*XJ
      ZONL1=Z1/AL1
      AL2=GL2-(2.*XJ)
      ZONL2=Z2/AL2
      ERR=1.E-10
      C=((2./3.)*SQRT(2.*ES*Q*ANA))/CO
      S=1.
      A=((S**2.)*ES*Q*ANA)/(CO**2.)

      DO 110 L=1,51,25
      T=TEMP(L)+273.16
      AK1(L)=AK(L)*ZONL1/ZONL
      AK2(L)=AK(L)*ZONL2/ZONL


      BFP=PSI(L)
      VFB=FBV(L)
      VTH=VT(L)

      VIN=0.
      DO 100 N=1,200
      VIN=VIN+0.03

      VO=0.
      DVO=1.

      ALPHA=A*A+2.*A*ABS(VDD-VFB)

      UP=VTH+0.12
      IF(L.EQ.1) DOWN=VTH-0.3
```

```
      IF(L.EQ.26) DOWN=VTH-0.4
      IF(L.EQ.51) DOWN=VTH-0.6

      VSA2=VDD-VFB-2.*BFP+A-SQRT(ALPHA)

      ALPHA1=A*A+2.*A*ABS(VIN-VFB)
      VSA1=VIN-VFB-2.*BFP+A-SQRT(ALPHA1)
      IF(VSA1.LE.0.) VSA1=0.

      IF(VIN.LE.DOWN) GO TO 5
      GO TO 10

C     HERE IS THE JUNCTION LEAKAGE IMBALANCE
5     IF(L.EQ.1) JL=1.E-08
      IF(L.EQ.26) JL=1.E-05
      IF(L.EQ.51) JL=5.E-03

10    IF(VIN.GE.DOWN .AND. VIN.LE.UP) GO TO 15
      GO TO 20

C     HERE ARE THE JUNCTION AND SUBTHRESHOLD LEAKAGE IMBALANCE
15    IF(L.EQ.1) JL=1.E-08+9.E-10*EXP((VIN-DOWN)/0.04)
      IF(L.EQ.26) JL=1.E-5+1.12E-10*EXP((VIN-DOWN)/0.036)
      IF(L.EQ.51) JL=5.E-03-03*EXP((VIN-DOWN)/0.72)


20    IF(VO.GE.VSA2) GO TO 40
      AID2=AK2(L)*((VSA2-VO)*(VDD-VFB-2.*BFP-0.5*VSA2-0.5*VO)
     1-C*S*((2.*ABS(BFP)+ABS(VSA2-.0))**1.5-(2.*ABS(BFP)+
     2ABS(VO-0.))**1.5))*1000.
      IF(VIN.LE.UP) AID1=JL
      IF(VIN.LE.UP) GO TO 25

      IF(VO.GE.VSA1) AID1=AK1(L)*((VSA1-0.)*(VIN-VFB
     1-2.*BFP-0.5*VSA1-0.)-C*S*((2.*ABS(BFP)+ABS(VSA1))
     2**1.5-(2.*ABS(BFP)+ABS(0.))**1.5))*1000.+JL
      IF(VO.GE.VSA1) GO TO 25

      AID1=AK1(L)*((VO-0.)*(VIN-VFB-2.*BFP-0.5*VO)
     1-C*S*((2.*ABS(BFP)+ABS(VO))**1.5-(2.*ABS(BFP)+
     2ABS(0.))**1.5))*1000.+JL

25    DIFF=AID2-AID1


      IF(ABS(DIFF).LE.ERR) GO TO 50
      IF(DIFF.GE.ERR) GO TO 30
      IF(DIFF.LE.ERR) GO TO 40

30    VO=VO+DVO
      GO TO 20

40    VO=VO-DVO
      DVO=DVO/2.
      VO=VO+DVO
      GO TO 20
50    CONTINUE

      VI(N)=VIN
```

```
      VOUT(N)=VO
      PRINT 38,DIFF,VOUT(N),VI(N)
100   CONTINUE

C     CALL THE PLOT ROUTINE
      IF(L.EQ.1) CALL NAMPLT

      CALL SCALE(VI,6.0,200,1,10.)
      CALL SCALE(VOUT,6.0,200,1,10.)
      IF(L.EQ.51) VOUT(201)=PP
      IF(L.EQ.51) VOUT(202)=QQ
      IF(L.NE.1) GO TO 115

      CALL AXIS(1.5,1.5,16HINPUT VOLTAGE(V),-16,6.0,0.,VI(201),
     1VI(202),10.)
      CALL AXIS(1.5,1.5,17HOUTPUT VOLTAGE(V),17,6.0,90.0,VOUT(201)
     1,VOUT(202),10.)

      IF(L.EQ.1) CALL PLOT(1.5,1.5,-3)
115   CALL LINE(VI,VOUT,200,1,0,0)
      IF(L.EQ.51) CALL ENDPLT
      IF(L.EQ.51) CALL EXIT

      IF(L.EQ.1) PP=VOUT(201)
      IF(L.EQ.1) QQ=VOUT(202)
110   CONTINUE

      RETURN
      END
```

APPENDIX C

ION IMPLANTED DEPLETION TYPE MOS DEVICES

This program, based on modified El-Mansy's model [El-Mansy, 1980] is used as a four-terminal device simulation tool. Important parameters are defined in the program. The flow diagram is shown in Fig. C.1.

To use this program, the following cared format should be used.

JOB

FTN

LGO

7/8/9

PROGRAM

7/8/9

INPUT DATA

6/7/8/9

The input data contain:

Card 1. Type of substrate (n or p)

Card 2. Substrate concentration and implant concentration

Card 3. Oxide thickness, gate length, gate width, drain diffusion window width DSW, and drain junction area

Card 4. Q_{SS}/q

Card 5. Gate material and ϕ_{MS}

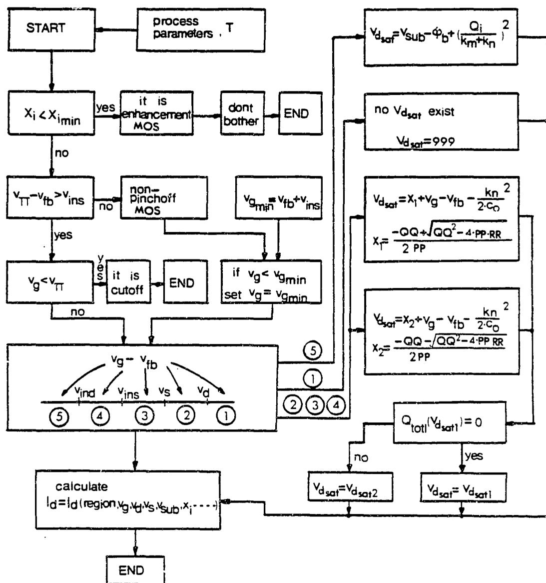


Fig. C.1 Flow Diagram of the Ion-Implanted Depletion MOS Device Simulation Program

Card 6. Junction depth and implant depth

Card 7. Trap density and average collision cross section area

Two more parameters are used, the implant concentration and the implant depth, compared to conventional enhancement type MOS devices.

```
      PROGRAM MOS(INPUT,OUTPUT,PLOT,TAPE6=OUTPUT,TAPE99=PLOT)
      DIMENSION TEMP(51),TYPE(2)

C     COMMON BLOCK /IGFET/

      COMMON /IGFET/ TYP,T,DT,BFP,VFB,WFD,WFUN,QSSQ,ANA,AND,ANI,
     1VDS,AIDS,CO,AL,Z,C,VTH,SSB,XJ,F,D,AJUNC,NT,SIGMA,AN

      COMMON /BIAS/ VD,VG,VS,VSUB
      COMMON /CONST/ Q,ES,H,AKB,PI,CHI,MO
      COMMON IREG
      COMMON /ATEMP/ BF,VINS,VIND
      COMMON /SATU/ VDS2,VDS5
      COMMON /AQSUR/ QSDV,KM,KN,VBB

      REAL MO,NT,JL,KM,KN
      INTEGER TYP

C     DEFINE FUNDAMENTAL CONSTANTS

      DATA Q,ES,H,AKB,PI,CHI,MO/1.602E-19,1.0625E-10,6.626E-34,1.38E-23,
     13.141592654,4.13,9.109E-31/

      DATA TYPE/1HN,1HP/

C     READ IN THE DEVICE PARAMETERS

      READ 1,TYP
      READ 3,SDL,DSDL
      READ 4,D,GL,Z,DSW,AJUNC
      READ 5,QSS
      READ 6,GATE,WFUN
      READ 8,XJ,XIMP
      READ 9,NT,SIGMA


C     CALCULATE THE METALLURGICAL CHANNEL LENGTH
      AL=GL-(2.*1.00*XJ)


C     CALCULATE Z/L AND CO, AND CHANGE QSS,SDL,AND DSDL TO MKS UNITS
      ZONL=Z/AL
      CO=3.453E-11/(D*1.E-10)
      QSSQ=QSS*1.E+4
      ANA=1.E+6*SDL
      AND=1.E+6*DSDL
      XI=XIMP*1.E-06


C     IMPLANTED LAYER CONCENTRATION IS COMPENSATED BY ANA
      AN=AND-ANA


C     CALCULATE KM AND KN C1 AND C2
      KM=SQRT(2.*ES*Q*(ANA*AN/(ANA+AN)))
      KN=SQRT(2.*ES*Q*AN)
      C1=CO/(1.+(CO*XI)/ES)
      C2=CO/(1.+(CO*XI)/(2.*ES))
```

```
C     CALCULATE THE DRAIN JUNCTION AREA (IN CM2)
      IF(AJUNC.EQ.0.) AJUNC=(XJ*PI*(2.*XJ+DSW+Z)+DSW*Z)*1.E-8

C     ADD THE BIAS
      VS=0.
      VS=6.5
      VSUB=0.
      VD=5.
      VD=12

      PRINT 20,TYPE(TYP)
      PRINT 30,SDL,DSDL,XIMP,XJ,D,CO,GATE,WFUN,QSS,AL,Z,AJUNC

20    FORMAT(1H1,19X,37(1H-)/20X,37H*SUMMARY OF MOSFET DEVICE PARAMETERS
     +*/20X,37(1H-)//20X,*DEVICE TYPE = *,A1,*-CHANNEL*/)
30    FORMAT(20X,*SUBSTRATE IMPURITY CONCENTRATION   =*,E14.7,1X,*CM-3*/
     +20X,*CHANNEL IMPLANT CONCENTRATION (N-TYPE)   =*,E14.7,1X,*CM-3*/
     +20X,*CHANNEL IMPLANT DEPTH                    =*,F5.2,10X,*UM*//
     +20X,*DRAIN JUNCTION DEPTH                 =*,F5.2,10X,*UM*//
     +20X,*OXIDE THICKNESS  =*,7X,F7.2,1X,*ANG.*/
     +20X,*OXIDE CAPACITANCE=*,E14.7,1X,*F/M2*//
     +20X,*GATE MATERIAL     =*,1X,A10/
     +20X,*GATE WORK FUNCTION=*,F6.3,1X,*VOLTS*,//
     +20X,*QSS=*,E14.7,1X,*CM-2*//
     +20X,*CHANNEL LENGTH=*,1X,F6.2,1X,*UM*/
     +20X,*CHANNEL WIDTH =*,1X,F6.2,1X,*UM*//
     +20X,*DRAIN JUNCTION AREA=*,1X,E14.7,1X,*CM2*//
     +20X,35(1H-)/)

C     INITIALIZE THE TEMPERATURE
      TEMP(1)=25.

      DO 50 L=1,3
      T=TEMP(L)+273.16
      VG=-7.5
      DO 48 M=1,101

C     THE MOBILITY (BULK AND SURFACE)
      UB=800.*1.E-04*(T/300)**(-1.5)
      US=450.*1.E-04*(T/300)**(-1.5)

C     CALL TEMPE FOR VALUES OF ANI,BF,VINS,VIND,VFB ETC
      CALL TEMPE

C     IF THE IMPLANT IS SHADOW, IT IS KIKE A E-TYPE MOS WITH VTH
C     DECREASE BY QI/CO

      XIMIN=SQRT((2.*ES*ANA*(BF-VSUB)/(Q*AN*(ANA+AN))))
      IF(XI.LT.XIMIN) PRINT 3000
      IF(XI.LT.XIMIN) GO TO 50
```

```
C     CALCULATE THE TURN ON VOLTAGE
      VTT=VS+VFB-(KM/KN)**2.*(VS+BF-VSUB)+(KM/C1)*SQRT(VS+BF-VSUB)-(1/C2
     1)*Q*AN*XI

      VTTFB=0.
      VTTFB=VTT-VFB

C     IF VTTFB.GE.VINS PINCHOFF IS POSSIBLE

      IF(VTTFB.GE.VINS .AND. VG.LT.VTT) PRINT 2500
      IF(VTTFB.GE.VINS .AND. VG.LT.VTT) GO TO 48

C     IF VTT.LT.VINS IT IS A NON-PINCHFOO DEVICE

      VGMIN=-999.
      IF(VTTFB.LT.VINS) VGMIN=VFB+VINS
C     IF(VTTFB.LT.VINS) PRINT 2000,VGMIN

      ICHAN=0
      IF(VTTFB.LT.VINS .AND. VG.LT.VGMIN) ICHAN=1
C     TEMPORARY SAVE THE VALUE OF VG
      IF(ICHAN.EQ.1) VGATE=VG
      IF(ICHAN.EQ.1) VG=VGMIN

C     ASSIGN REGIONS
      VTEST=VG-VFB

C     IN REGION1 BOTH SOURCE AND DRAIN ARE ACCUMULATED
      IF(VTEST.GE.VD) IREG=1

C     IN REGION2 SOURCE IS ACCUMULATE DRAIN IS DEPLETED
      IF(VTEST.GE.VS .AND. VTEST.LT.VD) IREG=2

C     IN REGION3 BOTH SOURCE AND DRAIN ARE DEPLETED
      IF(VTEST.GE.VINS .AND. VTEST.LT.VS) IREG=3

C     IN REGION4 SOURCE IS INVERTED DRAIN IS DEPLETED
      IF(VTEST.GE.VIND .AND. VTEST.LT.VINS) IREG=4

C     VBB IS USED IN SUBROUTINE QSURF REGION 4
      IF(IREG.EQ.4) VBB=((CO/KN)**2.)*((VG-VFB+BF)**2.)+VSUB-BF

C     IN REGION5 BOTH SOURCE AND DRAIN ARE INVERTED
      IF(VTEST.LT.VIND) IREG=5

C     CALCULATE SATURATION VOLTAGE IN DIFFERENT REGIONS
```

```
C     IF IN REGION 2,3,4
      IF(IREG.GT.4 .OR. IREG.LT.2) GO TO 35


      BB=-1.*BF+VSUB-VG+VFB+(KN/(2.*CO))**2.
      CC=KN/(2.*CO)
      DD=Q*AN*XI
      PP=1.-(KN/KM)**2.
      QQ=2.*(KN/KM)*(DD/KM+CC*(KN/KM))
      RR=-1.*BB-((DD/KM)+CC*(KN/KM))**2.


      AA1=((.5/PP)*(-1.*QQ+SQRT(QQ*QQ-4.*PP*RR)))**2.+VG-VFB-(KN/(2.*CO)
     1)**2.
      AA2=((.5/PP)*(-1.*QQ-SQRT(QQ*QQ-4.*PP*RR)))**2.+VG-VFB-(KN/(2.*CO)
     1)**2.

      CHEK=KN*(((AA1-VG+VFB)+(KN/(2.*CO))**2.)**.5-(KN/(2.*CO)))+KM*SQRT
     1(AA1+BF-VSUB)-Q*AN*XI

      CHEE=KN*(((AA2-VG+VFB)+(KN/(2.*CO))**2.)**.5-(KN/(2.*CO)))+KM*SQRT
     1(AA2+BF-VSUB)-Q*AN*XI
      PRINT 3,AA1,AA2,CHEK,CHEE
      VDS2=999.

      IF(CHEK.LT.1.E-14) VDS2=AA1
      IF(CHEE.LT.1.E-14) VDS2=AA2

      GO TO 45


C     IF IN REGION 5
35    IF(IREG.EQ.5) VDS5=VSUB-BF+(Q*AN*XI/(KM+KN))**2.

C     IF VDS5.GT.(VG-VFB) AN ERROR MESSAGE IS GIVEN


C     IF IN REGION 1
C     THERE IS NO SATURATION VOLTAGE IN REGION 1


45    CONTINUE

C     CALL QSURF FOR SURFACE CHARGE INTEGRAL
      CALL QSURF
C     SEND THE VALUE OF VG BACK
      IF(ICHAN.EQ.1) VG=VGATE


C     CALCULATE IMPLANT AND DEPLETION CHARGE INTEGRAL
      QIDV=Q*AN*XI*(VD-VS)
      QJDV=(2/3)*KM*((VD+BF-VSUB)**1.5-(VS+BF-VSUB)**1.5)


C     CALCULATE THE CURRENT
      AID=ZONL*UB*(QIDV-QJDV)-ZONL*US*QSDV

      PRINT 11,TEMP(L),IREG,VG,VDS2,AID,VGMIN
      PRINT 4,VS,VD,VINS,VIND
      PRINT 4,VM,VTEST,VBB
```

```
48     VG=VG+.2
50     TEMP(L+1)=TEMP(L)+125.


1      FORMAT(I1)
3      FORMAT(7E10.2)
4      FORMAT(4F10.2,E10.2)
5      FORMAT(E10.2)
6      FORMAT(A10,F5.1)
8      FORMAT(2F5.2)
9      FORMAT(2E10.2)
10     FORMAT(20X,I3)
11     FORMAT(20X,F10.2,I5,2F10.2,E10.4,F10.4)
2000   FORMAT(1X,*NON PINCHOFF DMOS, VGMIN   =*,F8.4,/)
3000   FORMAT(/20X,*DO NOT WASTE YOUR TIME,IT IS ESSENTIALLY AN E-TYPE*)
2500   FORMAT(1X,*THE DEVICE IS CUT-OFF*)
2800   FORMAT(20X,*BE CAUTION, TWO SOLUTIONS EXIST*)
4000   STOP
       END
       SUBROUTINE TEMPE

C-----------------------------------------------------------------
C
C      THIS SUBROUTINE CALCULATE THE TEMPERATURE-RELATED PARAMETERS
C
C-----------------------------------------------------------------

C      COMMON BLOCK /IGFET/

       COMMON /IGFET/ TYP,T,DT,BFP,VFB,WFD,WFUN,QSSQ,ANA,AND,ANI,
      1VDS,AIDS,CO,AL,Z,C,VTH,SSB,XJ,F,D,AJUNC,NT,SIGMA,AN
       COMMON /BIAS/ VD,VG,VS,VSUB
       COMMON /CONST/ Q,ES,H,AKB,PI,CHI,MO
       COMMON IREG
       COMMON /ATEMP/ BF,VINS,VIND
       COMMON /SATU/ VDS2,VDS5
       COMMON /AQSUR/ QSDV,KM,KN,VBB


C      COMPUTE INTRINSIC CONCENTRATION
       EG=1.1557-((7.021E-4*T*T)/(T+1108.))
       ANI=(3.87298E+22)*(T**1.5)*EXP(-9.69210E-20/(T*AKB))


C      COMPUTE BFP AND VFB
       BFP=((AKB*T)/Q)*ALOG((ANA+SQRT(ANA**2.+4.*(ANI**2.)))/(2.*ANI))
       BFN=((AKB*T)/Q)*ALOG((AN+SQRT(AN**2.+4.*(ANI**2.)))/(2.*ANI))
       BF=BFP+BFN


C      COMPUTE VINS VIND TO DETERMINE OPERATION REGION
       VINS=-1.*BF-(1./CO)*SQRT(2.*ES*Q*AN*(VS+BF-VSUB))
       VIND=-1.*BF-(1./CO)*SQRT(2.*ES*Q*AN*(VD+BF-VSUB))


C      COMPUTE VFB
       WFD=WFUN-(CHI-BFN+(EG/(2.)))
       VFB=WFD-(Q*QSSQ)/CO
```

```
      RETURN
      END
      SUBROUTINE QSURF
C-----------------------------------------------------------------
C
C     THIS SUBROUTINE CALCULATE THE SURFACE CHARGE INTEGRAL
C
C-----------------------------------------------------------------

C     COMMON BLOCK /IGFET/

      COMMON /IGFET/ TYP,T,DT,BFP,VFB,WFD,WFUN,QSSQ,ANA,AND,ANI,
     1VDS,AIDS,CO,AL,Z,C,VTH,SSB,XJ,F,D,AJUNC,NT,SIGMA,AN
      COMMON /BIAS/ VD,VG,VS,VSUB
      COMMON /CONST/ Q,ES,H,AKB,PI,CHI,MO
      COMMON IREG
      COMMON /ATEMP/ BF,VINS,VINO
      COMMON /SATU/ VDS2,VDS5
      COMMON /AQSUR/ QSDV,KM,KN,VBB
      REAL KM,KN


      IF(IREG.EQ.1) GO TO 10
      IF(IREG.EQ.2) GO TO 20
      IF(IREG.EQ.3) GO TO 30
      IF(IREG.EQ.4) GO TO 40
      IF(IREG.EQ.5) GO TO 50

10    QSDV=-1.*CO*(VG-VFB)*(VD-VS)+.5*CO*(VD**2.-VS**2.)
      GO TO 60


20    IF(VD.GT.VDS2) VM=VDS2
      IF(VD.GT.VDS2) GO TO 25
      QSDV=-1.*CO*(VG-VFB)*(VG-VFB-VS)+.5*CO*((VG-VFB)**2.-VS**2.)+(2./3
     +.)*KN*((VD-VG+VFB)+(KN*KN/(4.*CO*CO)))**1.5-(2./3.)*KN*(KN/(2.*CO)
     1)**3.-(KN*KN/(2.*CO))*(VD-VG+VFB)
      GO TO 60


25    QSDV=-1.*CO*(VG-VFB)*(VG-VFB-VS)+.5*CO*((VG-VFB)**2.-VS**2.)+(2./3
     +.)*KN*((VM-VG+VFB)+(KN*KN/(4.*CO*CO)))**1.5-(2./3.)*KN*(KN/(2.*CO)
     1)**3.-(KN*KN/(2.*CO))*(VM-VG+VFB)
      GO TO 60


30    IF(VD.GT.VDS2) VM=VDS2
      IF(VD.GT.VDS2) GO TO 35
      QSDV=(2./3.)*KN*(((VD-VG+VFB)+.25*KN*KN/(CO*CO))**1.5-((VS-VG+VFB)
     ++.25*KN*KN/(CO*CO))**1.5)-(.5*KN*KN/CO)*(VD-VS)
      GO TO 60


35    QSDV=(2./3.)*KN*(((VM-VG+VFB)+.25*KN*KN/(CO*CO))**1.5-((VS-VG+VFB)
     ++.25*KN*KN/(CO*CO))**1.5)-(.5*KN*KN/CO)*(VM-VS)
      GO TO 60
```

```
40     IF(VD.GT.VDS2) VM=VDS2
       IF(VD.GT.VDS2) GO TO 45
       QSDV=(2./3.)*KN*((VBB+BF-VSUB)**1.5-(BF-VSUB)**1.5)+(2.0/3.0)*KN*(
      +((VD-VG+VFB)+(KN/(2.*CO))**2.)**1.5-((VBB-VG+VFB)+(KN/(2.*CO))**2.
      1)**1.5)-(KN*KN/(2.*CO))*(VD**2.-VBB**2.)
       GO TO 60

45     QSDV=(2./3.)*KN*((VBB+BF-VSUB)**1.5-(BF-VSUB)**1.5)+(2.0/3.0)*KN*(
      +((VM-VG+VFB)+(KN/(2.*CO))**2.)**1.5-((VBB-VG+VFB)+(KN/(2.*CO))**2.
      1)**1.5)-(KN*KN/(2.*CO))*(VM**2.-VBB**2.)


       GO TO 60

50     IF(VD.GT.VDS5) VM=VDS5
       IF(VG.GT.VDS5) GO TO 55
       QSDV=(2/3)*KN*((VD+BF-VSUB)**1.5-(VS+BF-VSUB)**1.5)
       GO TO 60
55     QSDV=(2/3)*KN*((VM+BF-VSUB)**1.5-(VS+BF-VSUB)**1.5)
       GO TO 60


60     PRINT 3,VG
       RETURN
3      FORMAT(7E10.2)
       END
```

APPENDIX D

EPITAXIAL DIODE FABRICATION PROCEDURES

The epitaxial diodes are fabricated to simulate the source/drain junctions of epitaxial MOS devices. Starting material is p-type ⟨100⟩ substrate with resistivity of 0.6-1.4 Ω-cm, about 15 mils thick.

The fabrication processes were mainly divided into the following steps:

I. p^+ epitaxial growth ($\cong 10^{18}/cm^3$) to simulate the p^+ substrate

II. p epitaxial growth ($10^{16}/cm^3$)

III. n^+ diffusion for the diodes

IV. Open contact windows

V. Metallization and delineation

Each process is given in detail.

I. <u>p^+ epitaxial growth</u> (Group a)

1. Clean the wafer

   a. 1:1:1 Acetone:Methanol:Trichlorethylene at boiling temperature for 10 minutes

   b. Piranha clean - 3:1 $H_2SO_4:H_2O_2$ at 100°C ± 10°C for 5 minutes

   c. HF dip - 10:1 $H_2O:HF$ for 10 seconds

d. Nitric clean at 90°C ± 10°C for 10 minutes

e. HF dip 10:1 H_2O : HF for 10 seconds

f. Cascade rinse

2. HCl etch for 2 minutes at 1175°C in the epitaxial furnace

3. Epitaxial growth

Temperature: 1150°C

Gas: $H_2 + B_2H_6$ (0.2%) + SiH_2Cl_2

Growth rate: 1 μm/min

Time: 40 minutes

Resistivity: 0.06 Ω-cm

II. p epitaxial gorwth (Group a, b)

1. Clean wafer

2. HCl etch for 2 minutes at 1175°C in the epitaxial furnace

3. Epitaxial growth

Temperature: 1150°C

Gas: $H_2 + B_2H_6$ (24 ppm) + SiH_2Cl_2

Growth rate: 1 μm/min

Time: 10 minutes

Resistivity: 1.6 Ω-cm

III. n^+ diffusion

1. Grow SiO_2 of 6000Å on wafers (steam)

2. Open n^+ region windows (photoresist)

3. Predeposition

Temperature: 900°C

Gas: O_2 at 40, N_2 direct at 40, N_2 through $POCl_3$ at 50

Time: 1 minute

Gas: O_2 at 40, N_2 direct at 40 .

Time: 15 minutes

Resistivity: 32 Ω/□

4. Oxidation/drive-in

HF dip (10:1 - H_2O:Hf) for 15 seconds

Temperature: 1000°C

Time: 5 minutes dry O_2

Time: 30 minutes steam O_2

Resistivity: 14 Ω/□

x_j: 1.6 μm

IV. Open contact windows

V. Metallization and delineation

1. E-beam evaporation

2. Al etch 80:18:4 HPO_3 : H_2O : HNO_3 at 55°C ± 5°C

3. Sinter

Temperature: 450°C

Time: 1 minute

Gas: Forming gas (90% N_2 = 10% H_2)

LIST OF REFERENCES

Beasom, J. D., R. D. Moore, G. Mohammed, and B. L. Draper. "350°C CMOS Logic Process," International Electron Device Meeting Technical Digest, Paper 15.2, December 1981.

Black, J. R. "Electromigration--A Brief Survey and Some Recent Results," IEEE Trans. on Electron Devices, Vol. ED-16, April 1969.

Blacke, Douglas O. "Evaluation of CVD Tungsten Metallization for Integrated Circuit Application," M.S. Thesis, Department of Electrical Engineering, University of Arizona, Tucson, Arizona, 1980.

Blauschild, R. A., P. A. Tucci, R. S. Muller, and R. G. Meyer. "A New NMOS Temperature-Stable Voltage Reference," IEEE Journal of Solid-State Circuits, Vol. SC-13, No. 6, December 1978.

Conwell, E., and V. F. Weisskopf. "Theory of Impurity Scattering in Semiconductors," Physics Review, Vol. 77, No. 3, February 1950.

Coquat, J. A. "Semiconductors for High Temperature Active Devices Si, GaAs, and GaP," The Frontiers of High Temperature Electronics, Session 16, 1980 Electro-Professional Program, Boston, Mass., May 13-15, 1980.

Cosentino, Stephen J. "Analysis and Design of N-Channel MOS Transistors for Operation at 300°C," M.S. Thesis, Department of Electrical Engineering, University of Arizona, Tucson, Arizona, 1980.

Crowell, C. R., and S. M. Sze. "Temperature Dependence of Avalanche Multiplication in Semiconductors," Applied Physics Letters, Vol. 9, No. 6, September 1966.

Dennard, R. H. F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassoms, and A. R. Leblanc. "Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions," IEEE Journal of Solid-State Circuits, Vol. SC-9, No. 5, October 1974.

Eitan, B., and D. Frohman-Bentchkowsky. "Hot-Electron Injection into the Oxide in N-Channel MOS Devices," IEEE Trans. on Electron Devices, Vol. ED-28, No. 3, March 1981.

El-Mansy, Youssef. "Analysis and Characterization of the Depletion-Mode IGFET," IEEE Journal of Solid-State Circuits, Vol. SC-15, No. 3, June 1980.

Frohman-Bentchkowsky, D., and A. S. Grove. "Conductance of MOS Transistors in Saturation Regions," IEEE Trans. on Electron Devices, Vol. ED-16, No. 1, January 1969.

Gat, A. "Heat Pulse Annealing of Arsenic Implanted Silicon with a CW Arc Lamp," Electron Device Letters, Vol. EDL-2, No. 4., April 1981.

Grove, A. S. Physics and Technology of Semiconductor Devices. New York: John Wiley and Sons, Inc., 1967

Izumi, K., Doken, and H. Ariyoshi. "CMOS Devices Fabricated on Buried SiO_2 Layers Formed by Oxygen Implantation into Silicon," Electronics Letters, Vol. 14, No. 18, August 1978.

Johnson, N. M., D. J. Bartlink, and M. Schulz. "Transient Capacitance Measurement of Electronic States at the SiO_2-Si Interfaces," The Physics of SiO_2 and Its Interfaces. New York: Pergamon Press, 1978.

Jurgen, R. F. "High Temperature Electronics Applications in Space Explorations," Proceedings of the Conference on High Temperature Electronics, Session I, Tucson, Arizona, 1981.

Kingery, W. D., H. K. Bowen, and D. R. Uhlmann. Introduction to Ceramics. New York: John Wiley and Sons, Inc., 1976.

Lam, H. W., Z. B. Sobezak, R. F. Pinizzotto, and A. F. Tasch, Jr. "MOSFET Fabricated in {100} Single Crystal Silicon on Oxide Obtained by a Laser-Induced Lateral Seeding Technique," International Electron Device Meeting Technical Digest, Paper 15.2, December 1980.

Managan, W. W. "Needs for High Temperature Electronics in Fossil Energy Plants," Proceedings of the Conference on High Temperature Electronics, Session I, Tucson, Arizona, 1981.

Micheletti, F. B. "SOS Scales Best for VLSI," 158th Meeting of the Electrochem. Soc., Abstract No. 478, Hollywood, Florida, 1980.

Nieberding, W. C., and J. A. Powell. "High Temperature Electronic Requirements in Aeropropulsion Systems," Proceedings of the Conference on High Temperature Electronics, Session I, Tucson, Arizona, 1981.

Ning, T. H., C. M. Osburn, and H. N. Yu. "Emission Probablility of Hot Electron from Silicon into Silicon Dioxide," Journal of Applied Physics, Vol. 48, No. 1, January 1977.

Palkuti, J. P. "Integrated Circuit Characteristics at 260°C for Aircraft Engine Control Application," IEEE Trans. Components, Hybrids, and Mfg. Technology, Vol. CHMT-2, December 1979.

Raymond, L. S. Electrical Engineer, Microelectronics Laboratory, Department of Electrical Engineering, University of Arizona, Tucson, Arizona, Private Communication, 1980.

Richman, P. MOS Field-Effect Transistors and Integrated Circuits, New York: John Wiley and Sons, Inc., 1973.

Ryssel, H. "The Impact of Ion Implantation on VLSI," International Electronic Device Meeting Technical Digest, Paper 23.1, December 1978.

Ryssel, H., B. Schmiedt, and H. Kranz. "Minority Carrier Lifetime of Ion-Gettering Silicon," Semiconductor Silicon 1977, The Electrochemical Society Softbound Symposium Series, Princeton, New Jersey, 1977.

Sanders, N. H. "Present and Future Needs in High Temperature Electronics for the Well-Logging Industry," Proceedings of the Conference on High Temperature Electronics, Session I, Tucson, Arizona, 1981.

Taylor, G. W., H. M. Parley, R. C. Frye, and P. K. Chatterjee. "A Device Model for Ion-Implanted MOSFET," IEEE Trans. on Electron Devices, Vol. ED-26, No. 3, March 1979.

Tice, W. K., and H. J. Geipel. "Implantation Gettering in Semiconductors," 156th Meeting of Electrochemical Society, Abstract No. 496, Los Angeles, California, 1979.

Toyabe, T., K. Yamaguchi, S. Asai, and M. S. Mock. "A Numerical Model of Avalanche Breakdown in MOSFET," IEEE Trans. on Electron Devices, Vol. ED-25, No. 7, July 1978.

Troutman, R. R., and S. N. Chakravarti. "Subthreshold Characteristics of Insulated-Gate Field-Effect Transistors," IEEE Trans. Circuit Theory, Vol. CT-20, No. 6, November 1973.

Wang, K. L. "MOS Interface-State Density Measurements Using Transient Capacitance Spectroscopy," IEEE Trans. on Electron Devices, Vol. ED-27, No. 12, December 1980.

White, C. W., and P. S. Peercy. Laser and Electron Beam Processing of Materials. New York: Academic Press, 1979.