LEHIGH UNIVERSITY | Library & Technology Services

The Preserve: Lehigh Library Digital Collections

A Study Of Traps In The Metal-insulator - Semiconductor (mis) System With The Three Terminal Gated-diode Structure (devices, Mos, Microelectronics).

Citation

AGARWAL, ANANT KUMAR. *A Study Of Traps In The Metal-Insulator - Semiconductor (mis) System With The Three Terminal Gated-Diode Structure (devices, Mos, Microelectronics)*. 1984, https://preserve.lehigh.edu/lehigh-scholarship/graduate-publications-theses-dissertations/theses-dissertations/study-traps.

Find more at https://preserve.lehigh.edu/

This document is brought to you for free and open access by Lehigh Preserve. It has been accepted for inclusion by an authorized administrator of Lehigh Preserve. For more information, please contact preserve@lehigh.edu.

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify 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 complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. 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 illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

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

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

8418406

Agarwal, Anant Kumar

A STUDY OF TRAPS IN THE METAL-INSULATOR-SEMICONDUCTOR (MIS) SYSTEM WITH THE THREE TERMINAL GATED-DIODE STRUCTURE

Lehigh University PH.D. 1984

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

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __√__.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background __√__
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages __√__
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) __________ lacking when material received, and not available from school or author.
12. Page(s) __________ seem to be missing in numbering only as text follows.
13. Two pages numbered __________. Text follows.
14. Curling and wrinkled pages _____
15. Other ____

University
Microfilms
International

A STUDY OF TRAPS IN THE METAL-INSULATOR-SEMICONDUCTOR (MIS) SYSTEM WITH THE THREE TERMINAL GATED-DIODE STRUCTURE

by

Anant Kumar Agarwal

A Dissertation

Presented to the Graduate Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Electrical Engineering

Lehigh University

1984

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 8, 1984
(date)

Professor in Charge

Accepted May 8, 1984
(date)

Special committee directing
the doctoral work of
Anant Kumar Agarwal

Chairman

ACKNOWLEDGMENT

The valuable guidance, suggestions and encouragement provided by Prof. Marvin H. White, Major Professor, are acknowledged with deepest gratitude. A special word of thanks is due to Prof. Walter E. Dahlke, a committee member, for the tutorial guidance the author received from time to time. The assistance provided by Dr. R. H. Vogel, Messrs. Floyd Miller, Sanjay Jain, F. M. Rhodes and Chen-Chung Chao is acknowledged and appreciated.

In addition, the author is indebted to Dr. Lorenzo Faraone for introducing the author to the first principles of Microelectronics and to Professors Frank J. Feigl and Ralph J. Jaccodine for enlightening discussions. Finally, it is a pleasure to acknowledge the understanding and cooperation of my wife, Suman, throughout the course of this work.

The work reported herein was supported by the grants from The National Science Foundation, The Fairchild Foundation and the Hewlett-Packard Design Aids Program.

Table of Contents

1. INTRODUCTION 4

1.1 Motivation 4
1.2 Study of Interface Traps 10
1.2.1 Historical Review 10
1.2.2 Scope of the Present Work on Si-SiO_2 Interface Traps 12
1.3 Study of Bulk Traps in Insulators 13
1.3.1 Historical Review 13
1.3.2 Scope of the Present Work on Trapping in Insulators 16

2. THEORY OF NON-EQUILIBRIUM STATISTICS OF INTERFACE TRAPS 18

2.1 First Order Generation Recombination Statistics 18
2.2 Steady-State 22
2.2.1 Equilibrium steady-state 23
2.2.2 Nonequilibrium Steady-State 26
2.3 Small Signal Admittance 30
2.4 Normalization 34
2.5 Numerical Simulation 36
2.5.1 Steady-State Trap Occupancy 37
2.5.2 Pair Generation 40
2.5.3 Trap Admittance 41
2.6 Approximate Solution(Weak "Pinning" Regime) 45
2.7 Summary 47

3. EXPERIMENTAL EVALUATION OF Si-SiO_2 INTERFACE TRAPS 49

3.1 Fabrication 49
3.2 Experimental Setup 49
3.3 Experimental Results 52
3.4 Discussion 55
3.5 Summary 65

4. INTRODUCTION TO MULTI-DIELECTRIC STRUCTURES 70

4.1 Device Structure and the Principle of Operation 70
4.2 Historical Perspective 77
4.2.1 Theories 78
4.2.2 Experiments 86
4.3 Purpose of This Work 95

5. A NEW EXPERIMENTAL TECHNIQUE 96

5.1 Linear Voltage Ramp Technique 96
5.2 Charge Separation in the Semiconductor 97
5.3 Application to the Multi-Dielectric Structure 103

6. THEORY OF LINEAR VOLTAGE RAMP TECHNIQUE APPLIED TO THE MONOS STRUCTURE 115

6.1 Electrostatics of the MONOS Structure 115
6.2 Dynamic Analysis of the Ramp Measurements 119
6.3 Trapping in the Silicon Nitride Layer 121
6.4 Tunneling Model 135

6.5 The Overall Model 141
6.6 Results of the Numerical Simulation 145
6.6.1 Effect of Varying the Trap Capture Cross-section: 148
6.6.2 Effect of Varying the Trap Density: 150
7. EXPERIMENTAL RESULTS ON MNOS/MONOS STRUCTURES 152
7.1 Fabrication 152
7.2 Ramp Measurement 156
7.2.1 The Measurement Setup 161
7.2.2 Results on the P-Channel Device 163
7.2.3 Results on the N-Channel MNOS Device 177
7.2.4 On the Back-Tunneling 191
7.3 Summary 194
8. CONCLUSIONS 199
8.1 Studies on Si-SiO_2 Deep-Interface States 199
8.2 Recommendations for Future Work on Interface-Traps 201
8.3 Charge Injection and Trapping in MNOS/MONOS Devices: 201
8.4 Present and Future Investigation 203
References 204

LIST OF FIGURES

Figure	Page
1.1	6
1.2	8
2.1	20
2.2	25
2.3	28
2.4	35
2.5	38,39
2.6	43
2.7	44
3.1	50
3.2	53
3.3	57
3.4	58
3.5	59
3.6	60
3.7	61
3.8	62
3.9	63
3.10	64
3.11	66
3.12	67
3.13	68
4.1	71
4.2	73

Figure	Page
4.3	75
4.4	76
4.5	80
4.6	81
4.7	81
4.8	83
4.9	83
4.10	85
4.11	87
4.12	90
4.13	93
4.14	93
4.15	94
5.1	98
5.2	99
5.3	100
5.4	102
5.5	104
5.6	105
5.7	106
5.8	107
5.9	109
5.10	112
5.11	113
6.1	116
6.2	126

Figure	Page
6.3	129
6.4	134
6.5	138
6.6	139
6.7	140
6.8	143
6.9	144
6.10	147
6.11	149
6.12	151
7.1	157
7.2	158
7.3	159
7.4	160
7.5	162
7.6	165
7.7	166
7.8	167
7.9	169
7.10	170
7.11	172
7.12	174
7.13	175

Figure	Page
7.14	176
7.15	178
7.16	179
7.17	180
7.18	181
7.19	184
7.20	185
7.21	186
7.22	187
7.23	188
7.24	189
7.25	190
7.26	192
7.27	195
7.28	196
7.29	197

LIST OF SYMBOLS

Symbol	Description
A	Device area [cm^2]
$C_D = \sqrt{\frac{q^2 \varepsilon_s N_D}{2(\bar{U}_S - 1)kT}}$	Depletion capacitance [$F\ cm^{-2}$]
$C_{eff} = \frac{X_{oxB}}{\varepsilon_{oxB}} + \frac{X_N}{\varepsilon_N} + \frac{X_{oxt}}{\varepsilon_{oxt}}$	Effective capacitance of the MONOS sandwich [$F \cdot cm^{-2}$]
$C_{FN} = q^3/[16\pi^2 \hbar \phi_1]$	Pre-exponential factor in the modified Fowler Nordheim tunneling eq. (6.37) [$A \cdot V^{-2}$]
C_{it}	Interface trap capacitance for the statistically distributed continuum of traps [$F \cdot cm^{-2}$]
$C_{overlap}$	Source-gate overlap capacitance in the gated-diode structure in Fig. 5.6 [F]
C_{pad}	Gate bonding-pad capacitance in Fig. 5.6 [F]
D_{it}	Interface trap density [$cm^{-2} \cdot eV^{-1}$]
$e_n = \nu_n e^{(E_{TS} - E_{CS})/kT}$	Electron emission rate/active trap [s^{-1}]
$e_p = \nu_p e^{(E_{VS} - E_{TS})/kT}$	Hole emission rate/active trap [s^{-1}]
E_C, E_V	Conduction and valence band edges of Si in the bulk [eV]
E_{CS}, E_{VS}	Conduction and valence band edges of Si at the Si-SiO_2 interface [eV]
E_{Fn}, E_{Fp}	Free electron and hole quasi-Fermi levels in Si [eV]
E_{Fs}	Equilibrium Fermi-level (E_{Fn}=E_{Fp}) in Si [eV]
$E_{iB}=1/2[E_C+E_V-kT\ \ln(N_C/N_V)]$	Intrinsic level in the bulk [eV]
$E_{iS}=1/2[E_{CS}+E_{VS}-kT\ \ln(N_C/N_V)]$	Intrinsic level at the Si-SiO_2 interface [eV]
$E_{TP}=1/2[E_{CS}+E_{VS}+kT\ \ln(\nu_p/\nu_n)]$	Trap "pinning" level at the Si-SiO_2 interface, [eV]

E_{TS}	Trap level at the Si-SiO_2 interface, [eV]
$f_{FD}=[1+e^{(E_{TS}-E_{FS})/kT}]^{-1}$	Fermi-Dirac distribution
f_{ss}	Steady-state trap occupancy function
f_t	Trap occupancy function
f_{tD}	Occupancy function of the donor-like traps in the silicon nitride layer
G_{it}	Interface trap conductance for the statistically distributed, continuum of traps [$mho \cdot cm^{-2}$]
$i_{gen} = i_n = i_p$	Pair generated current per interface trap [A]
i_n, i_p	Pair generated electron and hole currents per interface trap [A]
$I_G(t)$	Displacement current at the gate [$A \cdot cm^{-2}$]
$J_c(x,t)$	Drift current in the nitride conduction band [$A \cdot cm^{-2}$]
$L_D = \sqrt{[2\varepsilon_s/qN_D\beta]}$	Extrinsic Debye length [cm]
m*	Effective mass of the electron in the oxide and nitride layer [Kg]
$\bar{n}=v_{th}\sigma_n n_s = \nu_n e^{(E_{Fn}-E_{CS})/kT}$	Electron capture probability [s^{-1}]
n_c	Electron density in the conduction band of the nitride [cm^{-3}]
n_i	Intrinsic carrier density in silicon [cm^{-3}]
$n_s = N_C e^{(E_{Fn}-E_{CS})/kT}$	Free electron concentration at the interface [cm^{-3}]
n_{so}	Steady-state free electron concentration at the interface [cm^{-3}]
N_C, N_V	Effective density of states in the conduction and valence bands of Si [cm^{-3}]

N_T	Discrete interface trap density [cm^{-2}]
N_{tA}	Density of Acceptor traps in the silicon-nitride layer [cm^{-3}]
N_{tD}	Density of Donor traps in the silicon-nitride layer [cm^{-3}]
$\bar{p} = v_{th}\sigma_p p_s = \nu_p e^{(E_{VS}-E_{Fp})/kT}$	Hole capture probability [s^{-1}]
$p_s = N_V e^{(E_{VS}-E_{Fp})/kT}$	Free hole concentration at the interface [cm^{-3}]
p_{so}	Steady-state free hole concentration at the interface [cm^{-3}]
$q = 1.6x10^{-19}C$	Electronic charge [C]
Q_G	Gate charge per unit device area [$C\cdot cm^{-2}$]
Q_{inj}	Injected charge into the nitride per unit device area [$C\cdot cm^{-2}$]
Q_n, Q_p	Free electron and hole charges in the conduction and valence bands, respectively
Q_N	Space charge in the nitride layer per unit device area [$C\cdot cm^{-2}$]
Q_S	Semiconductor charge per unit device area [$C\cdot cm^{-2}$]
$R = \sigma_n/\sigma_p$	Ratio of electron and hole capture cross-sections
$R_A = \bar{n}(1-f_t)$	Thermal electron capture rate [s^{-1}]
$R_B = e_n f_t$	Thermal electron emission rate [s^{-1}]
$R_C = \bar{p} f_t$	Thermal hole capture rate [s^{-1}]
$R_D = e_p(1-f_t)$	Thermal hole emission rate [s^{-1}]
$U_F = (E_{iB}-E_{Fn})/kT = -\ln(N_D/n_i)$	Normalized Fermi potential in the bulk. $U_F<0$ for n-type Si.

$U_G = E_G/kT$	Normalized band gap energy of Si
$U_S = (q/kT)\psi_S=(E_{iB}-E_{iS})/kT$	Normalized surface potential. $U_S>0$ in accumulation and $U_S<0$ in depletion for n-type Si.
$U_T = (E_{iS}-E_{TS})/kT$	Normalized interface trap energy. $U_T<0(U_T>0)$ for traps in the upper (lower) half of the bandgap.
$U_{TP} = (E_{iS}-E_{TP})/kT=1/2 \ln R$	Normalized interface trap "pinning" level
v_d	Drift velocity of electrons in the nitride conduction band [$cm\cdot s^{-1}$]
v_{th}	Thermal velocity of carriers [$cm\cdot s^{-1}$]
$V_B = -V_G$	Bias applied to the bulk Si relative to the gate [volt]
V_{FB}	Flatband voltage [volt]
$V_G = -V_B$	Bias applied to the Gate relative to the bulk Si [volt]
V_N	Voltage drop across the nitride layer [volt]
V_{oxB}	Voltage drop across the blocking oxide [volt]
V_{oxt}	Voltage drop across the tunneling oxide [volt]
V_{SB}	Reverse bias applied between source and bulk electrodes [volt]
$x_t = 1/(\sigma_{nD}N_{tD})$	Trapping distance in the nitride layer [cm]
X_N	Thickness of the nitride layer [cm]
$\bar{X}_N$	Centroid of the trapped charge in the nitride layer measured from the tunneling oxide-nitride interface [cm]
X_{oxB}	Thickness of the blocking oxide layer [cm]

Symbol	Definition
X_{oxt}	Thickness of the tunneling oxide layer [cm]
$Y_{it} = G_{it} + j\omega C_{it}$	Interface trap admittance for the statistically distributed continuum of traps [$mho\cdot cm^{-2}$]
Y_{s1}	Trap admittance of a discrete level [$mho\cdot cm^{-2}$]
α	Ramp rate [$V\cdot s^{-1}$]
$\beta = q/kT$	Inverse of the thermal voltage, $38.61V^{-1}$ at 300K [V^{-1}]
$\Delta Q_n(\Delta Q_b)$	Electron (hole) charge transferred between the semiconductor and the nitride film per unit device area [$C\cdot cm^{-2}$]
ΔV_{FB}	Shift in the flatband voltage due to a Write or Erase operation on the MNOS/MONOS device [V]
ε_N	Permittivity of the silicon nitride film $\approx 5.76x10^{-13}$ [$F\cdot cm^{-1}$]
$\varepsilon_{oxB}(\varepsilon_{oxt})$	Permittivity of the blocking (tunneling) oxide $\approx 3.46x10^{-13}$ [$F\cdot cm^{-1}$]
ε_s	Permittivity of silicon $\approx 1.04x10^{-12}$ [$F\cdot cm^{-1}$]
$\nu_n = v_{th}\sigma_n N_C$	Attempt to escape frequency of electrons [s^{-1}]
$\nu_p = v_{th}\sigma_p N_V$	Attempt to escape frequency of holes [s^{-1}]
ξ_N	Electric field in the nitride [$V\cdot cm^{-1}$]
$\xi_{oxB} = -\frac{Q_G}{\varepsilon_{oxB}}$	Electric field in the blocking oxide [$V\cdot cm^{-1}$]
$\xi_{oxt} = \frac{Q_S}{\varepsilon_{oxt}}$	Electric field in the tunneling oxide [$V\cdot cm^{-1}$]
$\rho_N(x,t)$	Density of the space-charge in the nitride film [$C\cdot cm^{-3}$]

Symbol	Definition
$\sigma_n(\sigma_p)$	Thermal capture cross-section of electron (hole) for the interface traps at the Si-SiO_2 interface [cm^2]
σ_{nD}	Thermal capture cross-section of electron for the donor-like trap in the nitride film [cm^2]
$\hat{\sigma}_{nD} = \sigma_{nD}(\frac{v_{th}}{v_d})$	Effective σ_{nD} [cm^2]
σ_s	Standard deviation corresponding to the normal probability distribution of surface potential fluctuations, equation (2.18)
$\tau_{no} = (\sigma_n v_{th} n_i)^{-1}$	Intrinsic majority carrier time constant [s]
$\psi_S = (E_{iB}-E_{iS})/q$	Surface potential, $\psi_S<0(>0)$ in depletion (accumulation) for n-type Si [volt]

ABSTRACT

The conventional MIS capacitor structure, frequently employed for the studies of trapping phenomena in MIS systems, does not provide any control over the minority carriers. The gated-diode structure overcomes this limitation since the third terminal can be used either to supply minority carriers or remove them from the inversion layer. In this thesis, we illustrate the use of the third terminal to study the deep interface traps at the $Si-SiO_2$ interface and bulk traps in the silicon nitride layers.

The Shockley-Read-Hall (SRH) statistics, when applied to the $Si-SiO_2$ interface traps in a deep-depleted gated-diode structure, predicts the "pinning" of the steady-state occupancy-function at a trap level which is (1/2)ln(R)kT eV below (above if R<0) the intrinsic level at the interface with $R = \sigma_n/\sigma_p$ being the ratio of thermal capture cross-sections. Rapid attenuation of the small signal ac conductance with frequency for the midgap interface traps is predicted and experimentally verified. A relatively high value of $R=10^4$ is measured corresponding to a low interface density of $7x10^{10} ev^{-1} cm^{-2}$ in the

midgap region indicating donor-like nature of the traps for n-type <100> silicon. Finally, it is emphasized that this study has verified one key aspect of the SRH statistics--that the capture and emission processes of minority carriers are present simultaneously and removal of minority carriers results in the elimination of the capture process, thereby making the minority carrier emission process dominant around the midgap region.

The charge-injection and trapping in the silicon nitride layers are studied with the Metal-Oxide-Nitride Oxide-Semiconductor (MONOS) structure. A new experimental technique based on the linear voltage ramp has been developed which minimizes back-tunneling of the injected charge and separates the charge carriers at the injecting boundary with a simultaneous measurement of the flatband voltage shift. The main conclusion from these studies is that the semiconductor injects electrons into the nitride layer for positive gate bias whereas the holes are injected from the semiconductor into the nitride layer under the negative gate bias. It is speculated that the recombination of electrons and holes takes place in the nitride layer via an "amphoteric" trap level. This result is in sharp contrast with the prveious interpretations based on a single carrier concept.

CHAPTER I

INTRODUCTION

1.1 Motivation

In the past two decades, progress in the development and application of Metal - Insulator - Semiconductor (MIS) devices has been very rapid. These developments were initiated by the successful experiments of Atalla et al [1] in 1959 in passivating silicon surfaces through the use of thermally grown SiO_2 layers. In effect, these experiments reduced the number of surface states available for trapping the charge induced by the gate electrode. The above work resulted in the realization of a practical MOS field-effect transistor, which was further developed by Hofstein and Heiman [2] in 1963. The advent of integrated circuit technology in the 1960's resulted in Metal - Oxide - Semiconductor (MOS) devices being utilized in such diverse applications as random-access and read-only memories, electronic calculators, data communications equipment and microprocessors. The study of the basic MIS structure has also led to the development of charge coupled devices [3], various non-volatile memory devices [4, 5] and various types of sensors [6, 7].

Traps are electrically active states within the energy

bandgap of a material and arise due to any imperfection in an otherwise regular atomic structure of the material. Foreign atoms, vacancies, dangling and stretched bonds, and various dislocations may result in electrically active traps in a material. As shown in Fig.1-1, traps can occur in the bulk of the semiconductor, bulk of the insulator and at the semiconductor - insulator and gate electrode - insulator interfaces.

The effect of the traps on device performance is usually detrimental. For instance, bulk traps in the semiconductor and at the semiconductor - insulator interface are responsible for leakage currents affecting refresh time of dynamic RAM cells, reducing the dynamic range of CCD arrays and increasing the power requirements in general. In addition, these traps are responsible for reduced signal/noise ratio and reduction in the charge transfer efficiency of CCD arrays, reduction of breakdown voltages etc. Likewise, bulk trapping in silicon oxide layers results in logic threshold instabilities and has been extensively studied [8]. However, traps have also been used to advantage. For instance, gold has been used in the past as a life-time killer in silicon for fast recovery of power rectifiers [9] and oxygen for intrinsic gettering of other undesirable impurities in silicon [10]. In addition, traps in dual-dielectric MIS structures have been used as charge storage sites resulting in a non-volatile memory action.

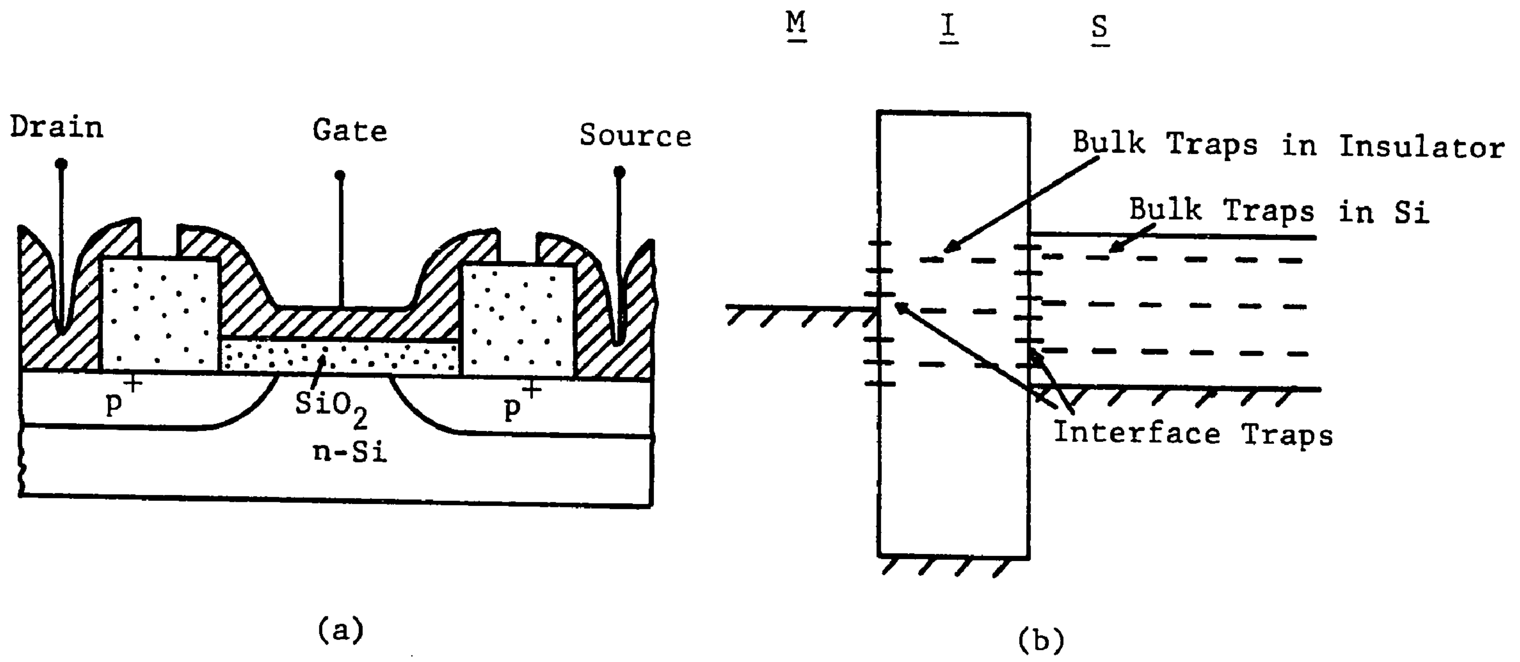
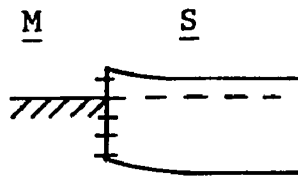


Fig. 1.1 (a) A typical MIS structure and (b) its energy band diagram illustrating the presence of the traps.

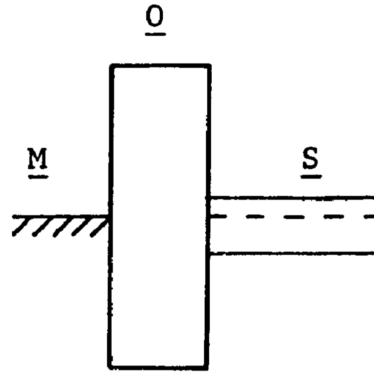
The variations in the MIS structure are illustrated in Fig.1-2. Although the traps affect the device performance in all cases, we have selected the MOS structure for studying the Si-SiO_2 interface traps and the MONOS structure for studying the trapping in the Si_3N_4 layers.

An electrically active trap can communicate with the free carriers in both the bands. The four basic band to trap and trap to band transitions, involving capture and emission of electrons and holes, formed the basis of the Schockley - Read - Hall (SRH) Statistics developed in 1952 [11, 12]. The SRH statistics was a first order phenomenological description of the various transitions and used phenomenological constants namely, electron(hole) capture cross-section, σ_n(σ_p) and the electron(hole) emission probability, $e_n(e_p)$. It was then possible to explain the low values of life-times of the indirect bandgap materials like silicon and germanium in terms of the deep-lying traps.

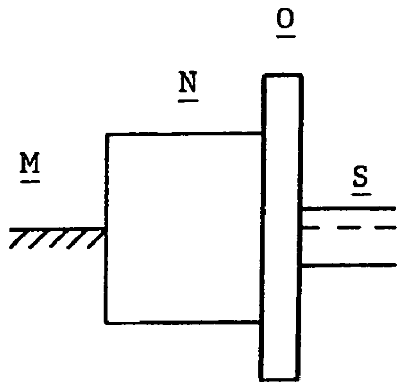
Based on the SRH statistics, numerous experiments [13, 14] have been designed to study the trap properties in the bulk silicon and at the Si-SiO_2 interface. When a voltage step is applied to a reverse-biased p-n junction or a MOS capacitor, capacitance and current transients are obtained. If the depletion region contains a deep-level trap, and if the temperature is low, then the transients are slow enough to be measured. Information can therefore be



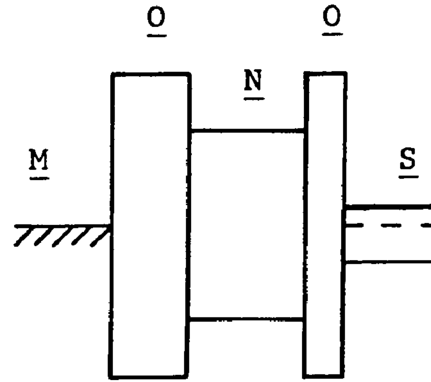
Schottky Barrier



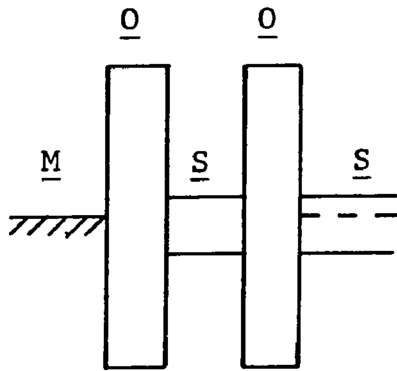
MOS Structure



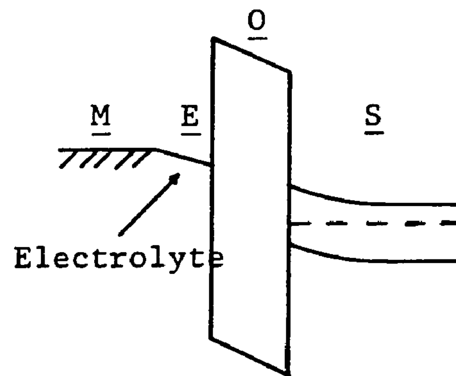
MNOS Structure



MONOS Structure



Floating Gate E^2PROM Structure



Ion-Sensing FET

Fig. 1.2 Various types of MIS structures.

obtained on the thermal emission rates of the trap. Experiments of this nature may have many variations since the voltage cycle may have several forms and the charge state of the impurities may be changed by photo-injection , current injection,or temperature of the specimen. Various measurements of this kind such as dark current, dark capacitance and photo-capacitance transient methods have been developed by Sah et al [13, 14]. However, all these techniques are limited to studying the capture and emission of only one type of carrier depending upon the energy of the trap being studied. It has not been in general, possible to study the emission and capture properties of the same trap for both electrons and holes.

The limitation to studying the transition between only one band and a trapping level arises due to the two terminal nature of the MOS capacitor usually employed for such measurements. The two terminal structure does not provide any control over the minority carriers and hence the interpretation of the results in terms of minority carrier transitions to the traps is at best ambiguous. To overcome this limitation, we have employed the three terminal gated-diode structure where the third terminal is a diffusion region (p^+ in n-type substrate or vice versa) which can be used to either supply the minority carriers or remove them from the inversion layer. In this thesis, we illustrate the use of the third terminal to study the deep interface traps at the $Si\text{-}SiO_2$ interface and bulk traps in the Si_3N_4 layers.

1.2 Study of Interface Traps

1.2.1 Historical Review

All of the previous mentioned techniques based on transients have been used extensively on the bulk traps in the semiconductor. Although such techniques are equally useful in studying the interface traps, it is in general difficult to separate the transients due to the bulk and interface traps. Therefore, the steady-state techniques based on the high frequency and quasi-static C-V curves have been used extensively [15]. These techniques measure the total charge stored in the interface traps and do not provide any information about the capture cross-sections of the traps. In addition, such techniques are limited in their sensitivity because the capacitance contribution due to the surface states is masked by the much larger insulator and depletion capacitances. This problem is overcome by the small signal admittance technique developed by Nicollian and Goetzberger [16].

The electrical behavior of the interface traps at the Si-SiO_2 interface has been studied extensively with the small signal admittance technique [16, 17, 18, 19]. As a result of these studies, we can conclude that (1) a continuum of states exists across the Si bandgap [20], (2) a dispersion of time constants has been observed and is attributed to random variations in fixed oxide charge and interface states [16], (3) the states lie mostly in a

thin "sheet" within the plane of the interface and do not spread appreciably into the insulator [17], and by careful processing, (4) interface trap densities, as low as 10^9 $cm^{-2}eV^{-1}$, can be obtained on Si samples with (100) orientation [18].

In the past, first-order rate kinetics [11] has been applied successfully to formulate the admittance of a continuum of states for majority carriers in depletion [20]. Subsequently, the concept of "surface potential fluctuations" was introduced [16] to explain the observed frequency dispersion in experimental plots of normalized trap conductance versus frequency measurements [$G/\omega C_{ox}$ vs f]. In these measurements, an unambiguous interpretation was possible only for majority carriers in depletion, where the minority carrier density at the Si-SiO_2 interface is small and minority carrier transitions to the interface traps may be neglected. When the surface is intrinsic or weakly inverted, minority carrier transitions and the associated inversion-layer capacitance may no longer be neglected. In this region, the frequency response of the trap conductance is influenced by both majority and minority carriers, and, for interface traps well into weak inversion, the statistical broadening due to the surface potential fluctuations is eliminated [16].

The interpretation of the trap conductance in the midgap and weak inversion regions requires the consideration of the capture

cross-sections for both the carrier types and the influence of the inversion layer capacitance [19]. Optically activated admittance measurements have been employed to study these states with assumptions regarding optical capture cross-sections and the degree of photo- ionization [21]. A recent review of the literature in this area describes the problems encountered in the midgap and weak-inversion regions [15].

1.2.2 Scope of the Present Work on $Si-SiO_2$ Interface Traps

In this work, we focus on the formulation of the nonequilibrium statistics of the interface traps and the manner in which it differs from the equilibrium case. The three-terminal, gated-diode structure provides a so-called "back-gate" bias to remove the minority carriers from the $Si-SiO_2$ interface. The removal of minority carriers allows an unambiguous interpretation of the trap conductance in the midgap and weak inversion regions. Our goal in these studies is two-fold: (1) to demonstrate, in the absence of minority carriers, the probability of the minority carrier emission from the interface traps in the midgap and weak inversion regions is indeed a dominant factor--as predicted by the first order SRH statistics. This effect results in a considerable deviation of the trap occupancy function from the Fermi-Dirac distribution and should be noticeable in the rapid attenuation or "pinning" of the measured trap conductance; (2) to study the interface-traps in the midgap and weak inversion regions with a unique method to extract the capture

cross-sectons, σ_n, σ_p, for both electrons and holes, respectively. The study of these deep-level traps is of major importance in the understanding of generation - recombination processes in solid-state devices. The differences in σ_n and σ_p are usually ignored in the calculation of the generation lifetime and the surface recombination velocity [22]. It will be shown that, in the deep-depleted silicon, relative values of capture cross-sections are extremely important in determining the position and the generation rate of the most efficient trapping site.

1.3 Study of Bulk Traps in Insulators

1.3.1 Historical Review

The traps in insulators are difficult to study because the free carriers are not available in the bands to interact with these traps. Thus, the carriers have to be injected into the bands and then, somehow, their subsequent capture has to be detected. Conversely, traps may be filled and then the carriers may be excited out of the traps by raising the temperature or shining light of the appropriate energy.

Historically, steady-state conduction through the thin insulating films was studied extensively. Mott and Gurney, in 1940, presented the first analysis of the Space Charge Limited Conduction (SCLC) in ideal (trap-free) insulators [23]. The work of Rose [24],

Smith [25] and Lampert [26] concentrated on the effect of trapping on SCLC. Rose [24] pointed out that a large fraction of injected charge will condense in the traps and, thus, the free carrier density will be much lower than that in a perfect insulator. Furthermore, since the occupancy of the traps is temperature dependent, SCLC will be temperature dependent. Lampert [26] noted that at low voltages, when volume generated free carrier density dominates over the injected free carrier density, ohmic conduction will dominate. Lampert also pointed out that if sufficient charge is injected into the insulator, all traps will be filled. Beyond this point, I-V curves would follow V^2 law according to the SCLC in trap-free insulators of Mott and Gurney [23]. Smith and Rose [25] presented experimental evidence of SCLC in CdS films of 20μm thickness.

Electrodes play an important role in the conduction through the thin insulating films as the charge has to be injected before it can be transported in the insulating layer. Frank and Simmons [27] presented the first analytical study of the transitions from SCLC to the Electrode Limited Conduction and vice-versa. At higher fields, the Poole-Frenkel Effect [28] in the insulators becomes dominant and $J \sim V^2$ relationship, characteristic of the SCLC, may be completely masked. In many cases it is not possible to identify J-ξ_{av} characteristic with either Schottky Effect at the electrodes or Poole-Frenkel Effect in the insulator. Pulfrey et al [29], in 1970,

presented an analytical model which clearly showed that depending upon the trap parameters, a spectrum of behavior ranging between Schottky Effect to the Poole-Frenkel Effect is possible.

Some of the very few experimental and theoretical studies of *transient* SCLC in insulators have been performed by Many et al [30, 31]. Many and co-workers have derived time-dependent current and space-charge distribution functions following the onset of injection via an ohmic contact under an applied voltage pulse. A single discrete trapping level has been assumed and detrapping via Poole-Frenkel Effect has been ignored. Later, Arnett [32] extended the analysis of Many et al to take into account the field dependent detrapping effect. This analysis is important in the context of Metal - Nitride - Oxide - Semiconductor (MNOS) device operation.

As can be seen from the discussion in preceding paragraphs, charge transport in the insulator is dominated by the traps, the trap-related space-charge & Poole-Frenkel conduction, whereas, injection of the charge is dominated by the nature of electrodes, work function difference between the electrode material and the insulator under study and the related Schottky effect. Electrical conductivity of the silicon nitride layer has been studied extensively [33, 34, 35]. It has been found [34] that the current is bulk controlled. At fields below 4 to 5 $MVcm^{-1}$, the conduction is ohmic whereas at higher fields, the conduction is dominated by

the Poole-Frenkel conduction.

Recently, Ginovker et al [36], Weinberg et al [37, 38, 39] and Schroder and White [40] have used a charge separation technique to investigate the steady charge transport through the MNOS structures. The main conclusion from these studies is that the current in the nitride layer is carried almost exclusively by holes in the steady-state [41]. However, these measurements are performed at fields in excess of 5 $MVcm^{-1}$ which is quite close to the breakdown field (7 $MVcm^{-1}$) of the silicon nitride.

1.3.2 Scope of the Present Work on Trapping in Insulators

In this work, we are interested in studying the trapping effects in the nitride layer. The steady-state conduction does not provide any insight into the trapping phenomenon as no net trapping occurs in the steady-state. Therefore, we are interested in measuring the transient when a bias is applied at the gate electrode relative to the bulk of the semiconductor in MNOS devices. In addition, we would like to limit the injection to only one electrode in order to simplify the interpretation of the results. In this study, the injecting electrode is chosen to be the semiconductor because it offers the possibility of separating the charge carriers by means of a p/n junction [36]. Once the carriers are injected in the nitride layer, they drift under the applied electric field and are captured by the deep-lying traps. Once trapped, the carriers

contribute to the space-charge and therefore modify the distribution of the electrostatic field which in turn reduces the injection of the carriers. If the bias at the gate is changed slowly, then the injection from the semiconductor and hence the trapping in nitride can be controlled. As long as all the injected charge becomes trapped, the displacement current at the gate electrode would be a sensitive indicator of the trapping and detrapping processes within the nitride layer. A detailed description of this study is given in chapters 4 to 7.

CHAPTER II

THEORY OF NON-EQUILIBRIUM STATISTICS OF INTERFACE TRAPS

2.1 First Order Generation Recombination Statistics

In this section, we discuss the first order rate equation for the trap occupancy in the dark and its steady-state solution for the equilibrium and non-equilibrium cases. In accordance with the Shockley-Read-Hall statistics [11], the four basic band-to-trap transitions that determine the occupancy of a discrete trapping level are indicated in Figure 2-1b. We can write the rate equation of the trap:

$$df_t/dt = R_A - R_B - R_C + R_D \ , \tag{2.1}$$

where f_t represents the probability that the trap is occupied with an electron and is known as the trap occupancy function. R_A, R_B, R_C and R_D represent various band-to-trap transition rates and are given by

$R_A \doteq$ thermal electron capture rate $[s^{-1}] = \bar{n}(1-f_t)$,

$R_B \doteq$ thermal electron emission rate $[s^{-1}] = e_n f_t$,

$R_C \doteq$ thermal hole capture rate $[s^{-1}] = \bar{p} f_t$,

$R_D \doteq$ thermal hole emission rate $[s^{-1}] = e_p(1-f_t)$,

where e_n and e_p denote the electron and hole emission probabilities.

$\bar{n}$ and $\bar{p}$ are, respectively, electron and hole capture probabilities $[s^{-1}]$ which can be written as

$$\bar{n} = v_{th}\sigma_n n_s = \nu_n \exp[(E_{Fn}-E_{CS})/kT] ,$$

and

$$\bar{p} = v_{th}\sigma_p p_s = \nu_p \exp[(E_{VS}-E_{Fp})/kT] .$$

Here, the thermal velocities of electrons and holes are considered to be the same (v_{th}) and any difference is implicitly absorbed in the definition of capture cross-sections, σ_n and σ_p. E_{Fn} and E_{Fp} denote the free electron and hole Quasi-Fermi levels. $\nu_n = \sigma_n v_{th} N_C$ and $\nu_p = \sigma_p v_{th} N_V$ are known as <u>attempt to escape frequencies</u> for electrons and holes and have dimensions of $[s^{-1}]$. n_s and p_s denote the free electron and hole concentration in cm^{-3} and are given by

$$n_s = N_C \exp[(E_{Fn}-E_{CS})/kT] ,$$
$$p_s = N_V \exp[(E_{VS}-E_{Fp})/kT] .$$

The rate at which the free electron concentration in the conduction band, $-Q_n/q$, increases due to the net emission from the discrete trap, is given by

$$-(1/q)dQ_n/dt = R_B - R_A + i_n/q \qquad (2.2)$$

with i_n the current density/trap, leaving the conduction band as

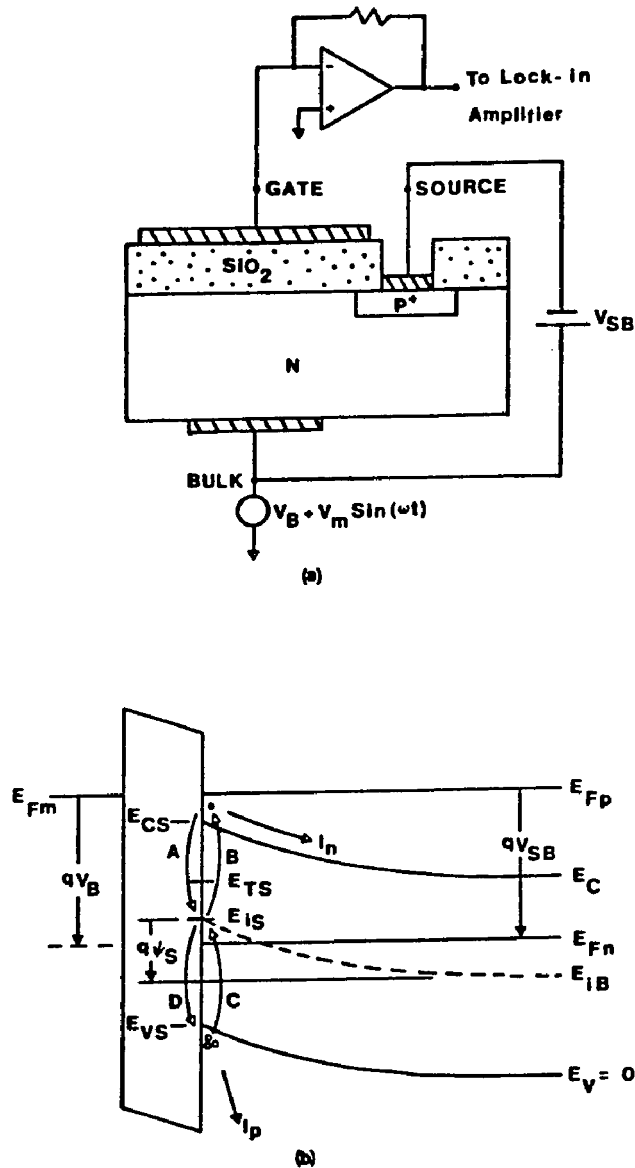


Fig. 2.1 (a) Schematic of the gated-diode structure under bias; (b) the energy band diagram. All energies are referenced to the top of the valence band in the bulk. Also shown are the four basic band-to-trap transitions as described in the text. See Fig. 3.2 for detailed block-diagram of the measurement station.

shown in Figure 2-1b. Similarly, the rate at which the free hole concentration in the valence band, Q_p/q, increases is

$$(1/q)dQ_p/dt = R_D - R_C - i_p/q \tag{2.3}$$

with i_p the current density/trap, leaving the valence band.

The emission probabilities, e_n and e_p, are related to the capture cross-sections, σ_n and σ_p, respectively, by the detailed balance in thermal equilibrium;

$$R_A = R_B, \ R_C = R_D \tag{2.4}$$

with the trap occupancy, f_t, given by the Fermi-Dirac distribution,

$$f_t = f_{FD} = 1/(1 + \exp[(E_{TS}-E_{Fs})/kT]) \quad , \tag{2.5}$$

where E_{Fs} is the equilibrium Fermi level and the degeneracy factor has been absorbed in E_{TS}. Equations (2.4) and (2.5) yield

$$e_n = \sigma_n v_{th} N_C \exp[(E_{TS}-E_{CS})/kT] = \nu_n \exp[(E_{TS}-E_{CS})/kT], \tag{2.6}$$

$$e_p = \sigma_p v_{th} N_V \exp[(E_{VS}-E_{TS})/kT] = \nu_p \exp[(E_{VS}-E_{TS})/kT]. \tag{2.7}$$

2.2 Steady-State

In steady-state, the thermal occupancy of any trap is a constant and, thus, the four processes which fill and empty the traps are in balance. Therefore, $df_t/dt=0$ and equation (2.1) yields

$$R_A - R_B = R_C - R_D \ , \tag{2.8}$$

which can be solved for the steady-state occupancy, f_{ss},

$$f_{ss} = (\bar{n} + e_p)/(e_n + e_p + \bar{n} + \bar{p}) . \tag{2.9}$$

In addition, $-dQ_n/dt = dQ_p/dt = 0$ in steady-state. Thus, equations (2.2), (2.3) and (2.8) yield the following expression for the pair generation current per trap, i_{gen}:

$$\begin{aligned} i_{gen} = i_p &= q(R_D - R_C) \\ &= -i_n = q(R_B - R_A) \\ &= q(e_p e_n - \overline{np})/(e_n + e_p + \bar{n} + \bar{p}) . \end{aligned} \tag{2.10}$$

It is noted that equation (2.9) is true for any trap and is independent of the energy distribution of the traps. As noted by Simmons and Taylor [42], f_{ss} can be multivalued at a particular value of the energy for different traps positioned at the same energy having different capture cross-sections. Thus, each trap has a unique f_{ss} associated with it. This observation has interesting

consequences in relation to the observability of a trap at a given energy with the admittance technique; however, a group of traps for which the ratio of the capture cross-sections, $R=\sigma_n/\sigma_p$, is a constant, is characterized by a unique distribution function.

2.2.1 Equilibrium steady-state

It is easy to see that in equilibrium ($n_s p_s = n_i^2$, $E_{FS}=E_{Fn}=E_{Fp}$, $R_A=R_B$, $R_C=R_D$), f_{ss} in equation (2.9) reduces to the Fermi-Dirac distribution of equation (2.5) and equation (2.10) reduces to $i_n=i_p=0$. Following Dahlke and Greve [43], various terms in equation (2.9) are plotted in Figure 2-2 on a logarithmic scale as functions of energy in the bandgap for different biasing conditions at T=300K. The emission times, $(e_n)^{-1}$ and $(e_p)^{-1}$ form a triangle with a peak at the trap "pinning" level, E_{TP}:

$$E_{TP} = (1/2)[E_{CS} + E_{VS} + kT\ln(\nu_p/\nu_n)], \qquad (2.11)$$

which is very close to the midgap. The sides of the triangle intercept the band-edges at $(\nu_n)^{-1}$ and $(\nu_p)^{-1}$ and have slopes of magnitudes $\pm\log_{10}(e)/kT$ [see equations (2.6) and (2.7)]. Assuming σ_n and σ_p are weak functions of trap energy over the bandgap, the capture times $(\bar{n})^{-1}$ and $(\bar{p})^{-1}$ are relatively constant across the bandgap and intersect $(e_n)^{-1}$ and $(e_p)^{-1}$, respectively, at the respective positions of E_{Fn} and E_{Fp}. It is noted that while $(e_n)^{-1}$ and $(e_p)^{-1}$ are functions of the trap energy and temperature and

independent of bias, $(\bar{n})^{-1}$ and $(\bar{p})^{-1}$ are relatively independent of trap energy and primarily functions of temperature and bias.

Although $R_A = R_B$, $R_C = R_D$ for all the traps and f_{ss} is given by the Fermi-Dirac distribution in equilibrium, it is instructive to examine further the plots in Fig.2-2.

FLAT BAND AND DEPLETION

At flat band or in depletion, $e_n \gg \bar{n} \gg \bar{p} \gg e_p$ for traps between E_{Fs} and E_{CS} and thus $R_A = R_B \gg R_C = R_D$. It follows that these traps are empty; i.e., $f_{ss} \ll 1$. For traps located at E_{Fs}, $e_n = \bar{n} \gg \bar{p} = e_p$ and $R_A = R_B \gg R_C = R_D$, and consequntly $f_{ss} = 1/2$. For traps between E_{TP} and E_{Fs}, $\bar{n} \gg e_n \gg e_p \gg \bar{p}$ and $R_A = R_B \gg R_C = R_D$, and thus these traps are occupied ($f_{ss} = 1$). Similar reasoning shows that, for traps between E_{TP} and E_1, $R_A = R_B \gg R_C = R_D$ with $f_{ss} = 1$ and, for traps between E_1 and E_{VS}, $R_C = R_D \gg R_A = R_B$ with $f_{ss} = 1$.

MIDGAP BIAS

The time constant and the associated band diagrams are shown in Figure 2-2b. E_{Fs} is placed at E_{TP} or close to it in this bias range. It is noted that, for traps between E_{TP} and E_{CS}, $e_n \gg \bar{n} = \bar{p} = e_p$, $R_A = R_B \gg R_C = R_D$ and therefore, $f_{ss} = 0$. These traps are thus empty. For traps located at E_{TP}, $e_n = e_p = \bar{n} = \bar{p}$ and thus $R_A = R_B = R_C = R_D$ with $f_{ss} = 1/2$. Similarly, the traps between E_{VS} and

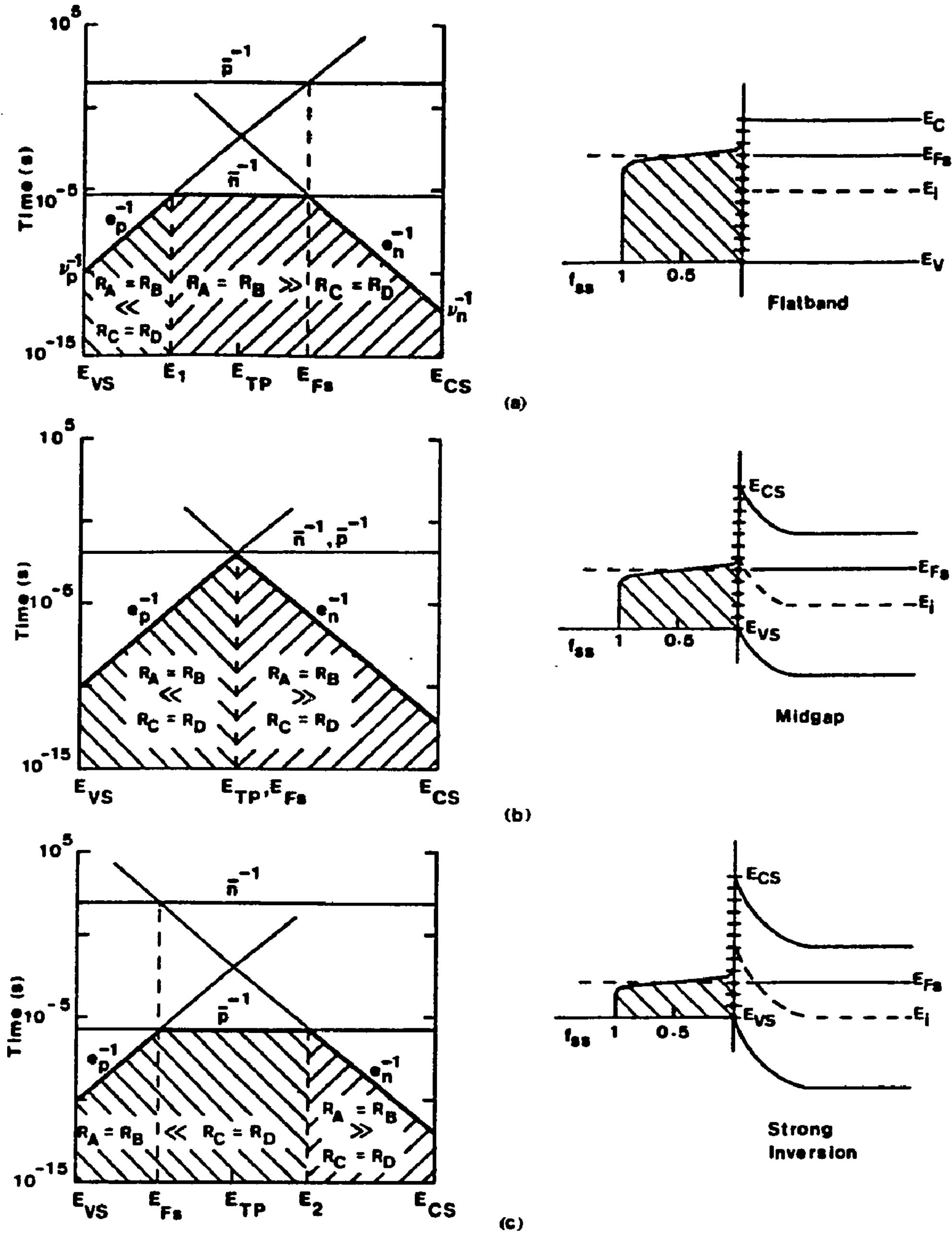


Fig. 2.2 The time-constant diagrams and the plots of the steady-state trap occupancy function f_{ss} vs trap energy in equilibrium (V_{SB}=0) for (a) flatband condition, (b) midgap bias range, and (c) the strong inversion region. The dominant time-constants are marked by the bold lines and the dominant rates are indicated by cross-hatching.

E_{TP} are characterized by $e_p >> \bar{n} = \bar{p} >> e_n$, $R_C = R_D >> R_A = R_B$ and $f_{ss} = 1$.

STRONG INVERSION

In strong inversion, E_{Fs} is placed between E_{VS} and E_{TP} as illustrated in Figure 2-2c. For traps between E_2 and E_{CS}, $e_n >> \bar{p} >> \bar{n}$, e_p and $R_A = R_B >> R_C = R_D$ with $f_{ss} << 1$. Similarly, between E_{Fs} and E_2, $\bar{p}$ is dominant with $R_C = R_D >> R_A = R_B$ and $f_{ss} <<1$. For traps located between E_{VS} and E_{Fs}, $e_p >> \bar{p} >> \bar{n} >> e_n$ and $R_C = R_D >> R_A = R_B$ with $f_{ss} =1$. Thus the traps in this region are occupied. As before, $f_{ss} = 1/2$ at E_{Fs}.

In summary, electron and hole emission probabilities e_n and e_p are independent of bias, whereas capture probabilities $\bar{n}$ and $\bar{p}$ are strongly dependent on bias. In depletion, $\bar{n}$ is dominant around E_{TP} and in weak-to-srong inversion $\bar{p}$ is dominant around E_{TP}.

2.2.2 Nonequilibrium Steady-State

With the application of V_{SB}, minority carriers (hole) are removed and consequently electron and hole quasi-Fermi levels are separated in the space-charge region by the applied reverse bias, qV_{SB}. This is clearly a nonequilibrium situation with $n_s p_s < n_i^2$. As a result of the removal of minority carriers, the hole capture rate, R_C, reduces to zero; however, the hole emission rate R_D, remains finite. The <u>detailed balance</u> must now involve both the

bands as opposed to the individual bands in the equilibrium case [see equation (2.4)]. In steady-state, equation (2.8) yields

$$R_C = 0; \quad R_B = R_A + R_D \quad , \tag{2.12}$$

for every interface trap across the bandgap. Thus, a steady-state dc current $i_p = -i_n = q(R_B - R_A)$ per trap, circulates between source and bulk electrodes. The steady-state occupancy function, f_{ss}, follows from equation (2.9) with $\bar{p} = 0$,

$$f_{ss} = (\bar{n} + e_p)/(e_n + e_p + \bar{n}) \quad . \tag{2.13}$$

In addition, equation (2.10) becomes

$$i_{gen} = i_p = -i_n = qe_pe_n/(e_n+e_p+\bar{n}). \tag{2.14}$$

The time constant and the related band diagrams are shown in Figure 2-3 for different biasing conditions. The quasi-Fermi levels, E_{Fn} and E_{Fp}, are shown to be separated as discussed before. We now discuss the implications of Figure 2-3 for each bias range.

FLATBAND AND DEPLETION

The time constant diagram in this range is identical to that in equilibrium (compare Figure 2-2a and 2-3a) with regard to the relative values of various time constants. It follows that f_{ss} is identical to the equilibrium distribution. In fact, although f_{ss} is

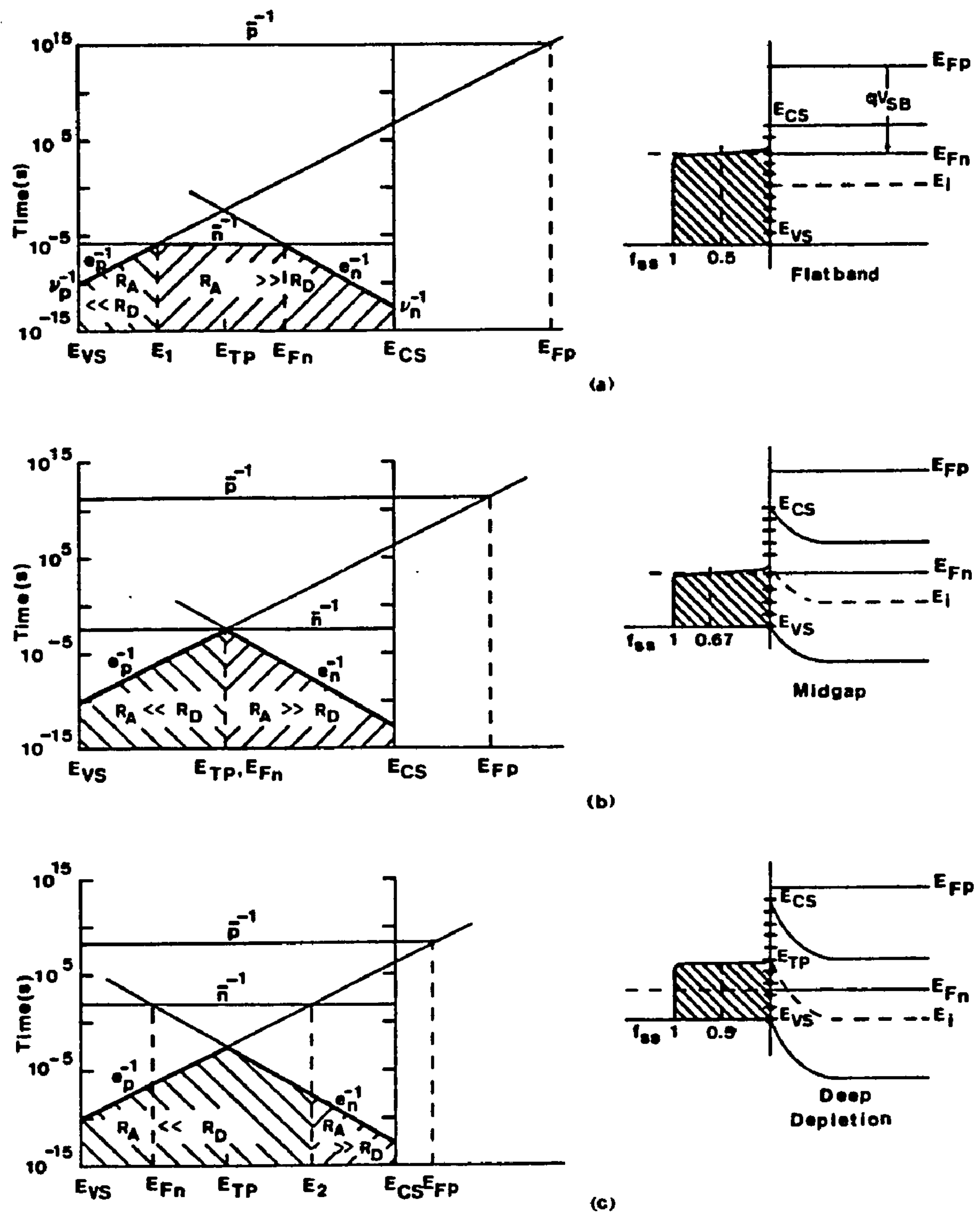


Fig. 2.3 The time-constant diagrams and the plots of f_{ss} vs trap energy in non-equilibrium ($V_{SB}<0$) for (a) the flatband condition, (b) midgap bias range, and (c) deep-depletion.

given by equation (2.13), it is quite similar to the Fermi-Dirac distribution. Electron emission rate, R_B, at all times and for all traps is equal to the sum of electron capture rate, R_A, and hole emission rate, R_D. For all traps between E_1 and E_{CS}, $R_A >> R_D$ with $R_B = R_A$ and for traps between E_1 and E_{VS}, $R_A << R_D$ with $R_B = R_D$.

MIDGAP BIAS

The time constant diagram is shown in Figure 2-3b. Comparing it with the corresponding diagram in equilibrium, we see that the hole capture time, $(\bar{p})^{-1}$, is very long in comparision to the electron capture time, $(\bar{n})^{-1}$, in contrast to their comparable values in equilibrium. This results in $f_{ss} = 2/3$ at $E_{Fn} = E_{TP}$ as opposed to $f_{ss} = 1/2$ in equilibrium. Thus, f_{ss} begins to deviate from the Fermi-Dirac distribution. For traps located between E_{Fn} and E_{CS}, $R_A >> R_D$ with $R_B = R_A$ and for traps located between E_{Fn} and E_{VS}, $R_A << R_D$ with $R_B = R_D$. The traps in the lower half bandgap engage in pair generation via hole and electron emission processes.

DEEP-DEPLETION

In this bias range, the electron Fermi level, E_{Fn}, is placed between E_{VS} and E_{TP}. The hole quasi-Fermi level, E_{Fp}, remains separated from E_{Fn} by the amount qV_{SB} and maintains the hole concentration at a negligible level. Thus, inversion charge is not present in this bias range and the device is said to be in a deep-depleted state. As a result, $\bar{p}$ remains negligible over the entire

bandgap, whereas in equilibrium, inversion charge is present and $\bar{P}$ is dominant around E_{TP} in this bias range. As in equilibrium, $\bar{n}$ is also negligible everywhere in the bandgap. This leaves $(e_p)^{-1}$ as the dominant time constant between E_{VS} and E_{TP} and $(e_n)^{-1}$ as the dominant time constant between E_{TP} and E_{CS}. It follows from equation (2.13) that $f_{ss} << 1$ (traps empty) between E_{TP} and E_{CS} and $f_{ss} = 1$ (traps full) between E_{VS} and E_{TP} while $f_{ss} = 1/2$ at E_{TP} because $e_n = e_p$. Thus, the traps between E_{Fn} and E_{TP} remain occupied regardless of the position of E_{Fn} as long as E_{Fp} remains close to E_{CS}. We call this phenomenon "pinning" of the trap occupancy at E_{TP}. This is quite different from the equilibrium case where traps above the equilibrium Fermi level are always empty and traps below are full. We also note that $R_A << R_D$ with $R_B = R_D$ between E_{VS} and E_2. These traps engage in pair generation and contribute to i_{gen} given by equation (2.14).

2.3 Small Signal Admittance

In a manner similar to the 2-terminal MOS capacitor [16, 15], equation (2.1) may be solved for a small sinusoidal perturbation about the dc bulk bias, V_B. Assuming $f_t = f_{ss} + \delta f_t$ and $n_s = n_{so} + \delta n_s$, proceeding with equations (2.1), (2.12), (2.13), $\delta f_t = f_m \exp(j\omega t)$ and $\delta n_s / n_{so} = (q/kT)\,\delta\psi_s$, where $\delta\psi_s$ is the perturbation in surface potential $\psi_s = (E_{iB} - E_{iS})/q$, we may write

$$\delta f_t = \frac{f_{ss}(1 - f_{ss})(q/kT)\delta\psi_s}{1 + (e_p/\sigma_n v_{th} n_{so}) + (j\omega f_{ss}/\sigma_n v_{th} n_{so})} \quad . \tag{2.15}$$

Equation (2.15) differs from a similar expression derived for the equilibrium case [16], in two respects: (a) f_{ss} is no longer represented by the Fermi Dirac distribution for the traps in the midgap region, as discussed earlier, and (b) a new term, $e_p/\sigma_n v_{th} n_{so}$, is added corresponding to the hole emission from these traps. For experimental conditions such that the holes can be removed in a time less than half the period of the sinusoidal excitation, the net ac responce of the processes R_A, R_B and R_D is given by the expression $i_T = qN_T df_t/dt = qN_T(j\omega\delta f_t)$ with N_T as the density of discrete interface traps [cm^{-2}]. The interface state admittance, Y_{s1}, for a single trap energy level takes the following form,

$$Y_{s1} = i_T/\delta\psi_s = \frac{j\omega q N_T f_{ss}(1 - f_{ss})(q/kT)}{1 + (e_p/\sigma_n v_{th} n_{so}) + (j\omega f_{ss}/\sigma_n v_{th} n_{so})} \quad . \tag{2.16}$$

The interface states appear as a continuum over the bandgap of the silicon. The admittance given by eq.(2.16) for a single trap level must therefore be modified. For a continuum of states at a finite temperature, capture and emission of carriers may occur via states located within a few kT/q on either side of the electron

Fermi level. This results in a time constant dispersion. In order to consider a continuum of interface states, N_T is replaced by D_{it} [$cm^{-2}eV^{-1}$] and eq.(2.16) is integrated over the bandgap. As discussed by Nicollian and Goetzberger [16], surface potential fluctuations in the plane of the interface cause further broadening of the time constant dispersion. Small fluctuations in the surface potential cause large fluctuations in the free carrier electron density n_{so} and thus broaden the time constant dispersion. It is shown in reference [16] that a random (normal) distribution of the built-in oxide charges and the interface states result in a normal distribution of the surface potential. We devide the plane of the interface in small but equal area segments. Within each segment, the normalized surface potential $U_S = \psi_S(q/kT)$ is considered to be uniform and the eq.(2.16) is integrated over the bandgap of silicon to account for the continuum of interface states. Admittance contributions from various segments are then combined in parallel to yield

$$Y_{it} = G_{it} + j\omega C_{it}$$
$$= \int_{-\infty}^{\infty} dU_S P(U_S) \int_{E_{VS}}^{E_{CS}} \frac{j\omega q D_{it} f_{ss}(1 - f_{ss})\cdot(q/kT)dE_{TS}}{1 + (e_p/\sigma_n v_{th} n_{so}) + (j\omega f_{ss}/\sigma_n v_{th} n_{so})} \quad , (2.17)$$

where G_{it} and C_{it} represent the interface trap conductance and capacitance, respectively, and the normal probability distribution

is given by

$$P(U_S) = (2\pi\sigma_s^2)^{-1/2}\exp[-(U_S-\bar{U}_S)^2/2\sigma_s^2] \ , \tag{2.18}$$

where $\bar{U}_S$ is the mean normalized surface potential and σ_s^2 the normalized variance of the distribution [16, 15].

The overall small signal equivalent circuit between gate and bulk terminals is shown in Figure 2-4. The source is an ac-short with the bulk and is not shown explicitly in the equivalent circuit.

C_D is the depletion capacitance per unit area and is given by

$$C_D = \sqrt{\frac{q^2\varepsilon_s N_D}{2(\bar{U}_S-1)kT}} \ , \tag{2.19}$$

where the depletion approximation has been used. Trap conductance and capacitance can be written in terms of the measured values of the parallel combination of the conductance G_m and capacitance C_m [see Fig. 2-4] in the following form:

$$G_{it}/\omega C_{ox} = \frac{G_m/\omega C_{ox}}{(G_m/\omega C_{ox})^2 + (1 - C_m/C_{ox})^2} \ , \tag{2.20}$$

$$(C_{it} + C_D)/C_{ox} = \frac{1 - C_m/C_{ox}}{(G_m/\omega C_{ox})^2 + (1 - C_m/C_{ox})^2} - 1 \ . \qquad (2.21)$$

2.4 Normalization

The previous equations take a simple form when all the energies are normalized to the thermal energy kT and referenced to the intrinsic level in the bulk, E_{iB}, with the exception of the interface trap energy, E_{TS}, which is referenced to the intrinsic level at the surface, E_{iS}. In addition, the top of the valence band edge in the bulk, E_V, is assigned zero energy [see Fig. (2.1)b]. E_{iB} and E_{iS} are given by

$$E_{iB} = (1/2)[E_C + E_V - kT \ln(N_C/N_V)],$$
$$E_{iS} = (1/2)[E_{CS} + E_{VS} - kT \ln(N_C/N_V)] \quad .$$

With reference to Fig. 2-1(b), various normalizations of interest are

$$U_S = (q/kT)\,\psi_S = (E_{iB} - E_{iS})/kT \quad ,$$
$$U_T = (E_{iS} - E_{TS})/kT \quad ,$$
$$U_{TP} = (E_{iS} - E_{TP})/kT = (1/2) \ln R \quad ,$$
$$U_F = (E_{iB} - E_{Fn})/kT = -\ln(N_d/n_i) \quad ,$$
$$U_G = E_G/kT \quad ,$$

where $R = \sigma_n/\sigma_p$ and E_{TP}, the trap "pinning" level, is given by Eq. (2.11).

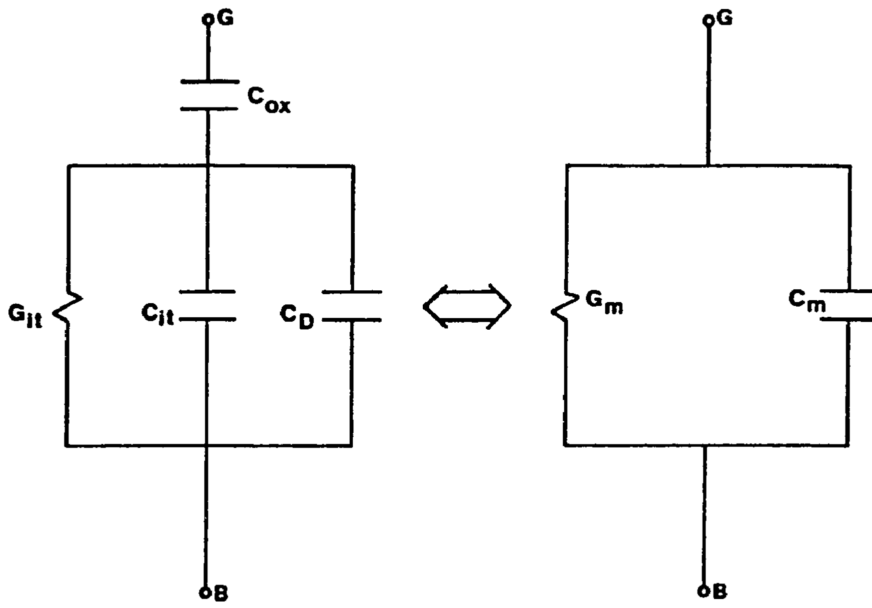


Fig. 2.4 The overall equivalent circuit diagram between the gate and bulk electrodes of the device.

In accordance with the definitions listed above, for n-type Si, $U_F<0$; $U_S<0$ in depletion, $U_S=0$ at flatband and $U_S>0$ in accumulation. $U_T>0$ for the interface states in the lower half of the bandgap and vice versa. Equations (2.13), (2.14) and (2.17) now take the following form:

$$f_{ss} = \frac{1 + R\exp(U_S - U_F - U_T)}{1 + R\exp(U_S - U_F - U_T) + R\exp(-2U_T)}, \tag{2.22}$$

$$i_{gen}\tau_{no}/q = [R\exp(-U_T) + \exp(U_T) + R\exp(U_S - U_F)]^{-1} \tag{2.23}$$

$$Y_{it} = G_{it} + j\omega C_{it}$$

$$= \int_{-\infty}^{\infty} dU_S P(U_S) \int_{-U_G/2}^{U_G/2} \frac{j\omega q D_{it} f_{ss}(1-f_{ss})dU_T}{1+R^{-1}\exp(U_F-U_S+U_T)+j\omega\tau_{no} f_{ss}\exp(U_F-U_S)}, \tag{2.24}$$

where $\tau_{no} = (\sigma_n v_{th} n_i)^{-1}$ is the intrinsic majority carrier time constant.

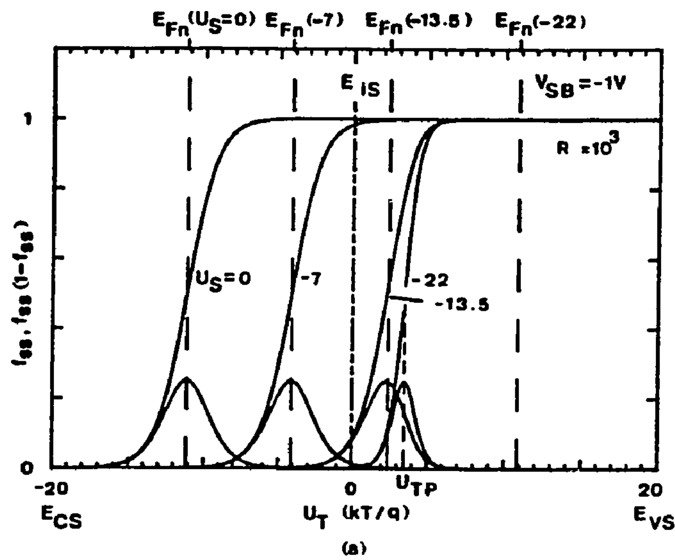
2.5 Numerical Simulation

In this section, we examine the nature of the steady-state occupancy, f_{ss}, in nonequilibrium [equation (2.22)], pair generation [equation (2.23)], and the frequency dispersion of the trap conductance and capacitance equation (2.24) and their sensitivity to various trap parameters.

2.5.1 Steady-State Trap Occupancy

Figure 2-5a illustrates the plots of f_{ss} and $f_{ss}(1-f_{ss})$ vs. trap energy, U_T, for different values of the surface potential, U_S. The position of the majority carrier Fermi level, E_{Fn}, is also indicated. For the sake of comparsion, the Fermi-Dirac distribution is plotted in Figure 2-5b for the same values of U_S. It is seen that the occupancy function of the traps is identical to the Fermi-Dirac distribution in depletion, deviates considerably from the latter in the midgap bias range and "pins" at E_{TP} (normalized trap level, $U_{TP} = (1/2)\ln R = 3.45$ for $R = 1000$) in deep-depletion. In the "pinning" region (deep-depletion), $f_{ss}(U_{TP})=1/2$ and the width of the $f_{ss}(1-f_{ss})$ function is approximately half of that in the "non-pinning" region.

The "pinning" trap level, $U_{TP} = (1/2)\ln R$, shifts with R. This effect is shown in Figure 2-5c. $U_S = 2U_F$ is chosen to create deep-depletion and thus "pinning" of f_{ss}. For R=1, $U_{TP} = 0$ (midgap), $f_{ss}(U_{TP}=0) = 1/2$ and $f_{ss}(1-f_{ss})$ peaks at midgap. R = 1000 moves the "pinning" level, $U_{TP} = (1/2)\ln(1000) = 3.45$, closer to the valence band and consequently, $f_{ss}(1-f_{ss})$ now peaks at $U_T = U_{TP} = 3.45$ with f_{ss} $(U_T = 3.45) = 1/2$. R = 1/1000 moves the "pinning" level, $U_{TP} = (1/2)\ln(1/1000) = -3.45$, closer to the conduction band edge. It will be shown in section 2.5.3 that the effect of R on U_{TP} puts a limitation on the extent to which the deep interface states can be profiled with the conductance technique and also provides a



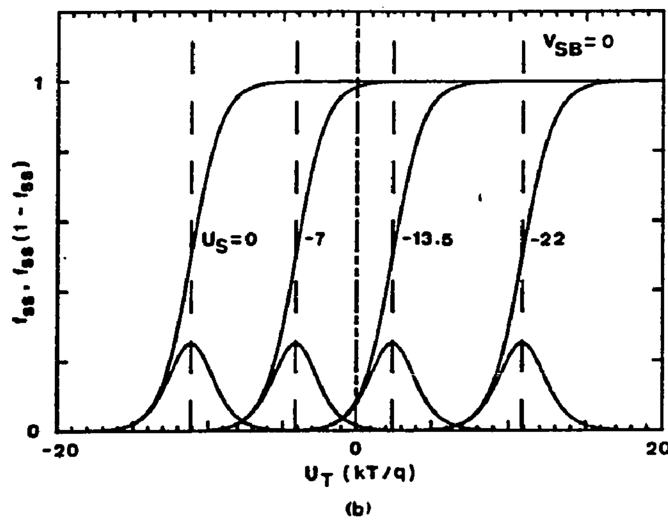


Fig. 2.5 Plots of the occupancy function f_{ss} and the scanning function $f_{ss}(1-f_{ss})$ vs the normalized trap energy U_T with the surface potential U_S as a parameter for (a) non-equilibrium ($V_{SB}<0$) and (b) equilibrium ($V_{SB}=0$) cases.

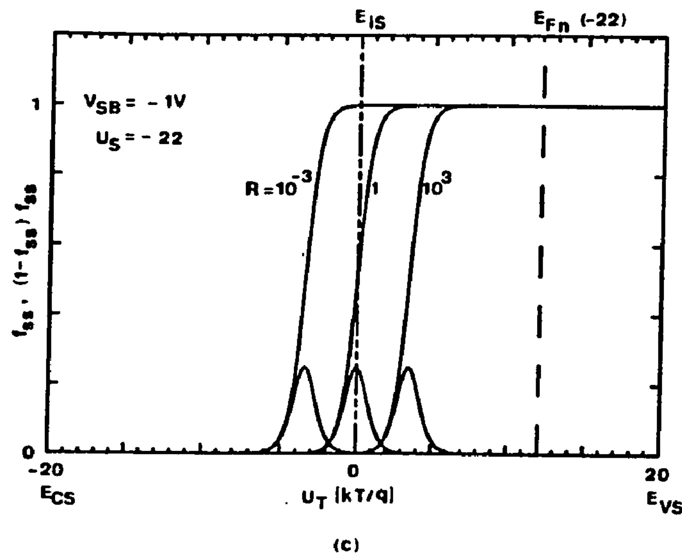


Fig. 2.5 (c) Plots of f_{ss} and $f_{ss}(1-f_{ss})$ vs U_T with $R=\sigma_n/\sigma_p$ as a parameter. The device parameters used are $U_F=-11.1$ ($N_D=10^{15}cm^{-3}$) and T=300K.

unique method for the determination of R.

2.5.2 Pair Generation

Equation (2.23) describes the normalized pair generation current per trap, $i_{gen}\tau_{no}/q$, as a function of R, U_T and U_S. At a fixed value of U_S, it is easy to see that $i_{gen}\tau_{no}/q$ is maximum at $U_T = U_{TP} = (1/2)\ln R$ and the maximum value is given by

$$(i_{gen})_{max}\tau_{no}/q = [2\sqrt{R} + R\exp(U_S-U_F)]^{-1} \qquad (2.25)$$

This implies that the interface traps located at the "pinning" level, U_{TP}, are the most efficient generating centers and, in addition, the maximum strength of pair generation depends on the ratio of capture cross-sections, R, for a given value of σ_n.

Figure 2-6a illustrates the variation of $i_{gen}\tau_{no}/q$ vs. U_T for two values of the surface potential, U_S. R is chosen to be unity so that maximum generation occurs at $U_{TP} = 0$. $U_S = U_F$ places E_{Fn} at midgap ($U_T = 0$) and, consequently, $(i_{gen})_{max}\ \tau_{no}/q = [2\sqrt{R} + R]^{-1} = 0.33$ from equation (2.25). $U_S = 2U_F$ places E_{Fn} closer to the valence band edge and since $U_F<0$ for n type semiconductor, $(i_{gen})_{max}\tau_{no}/q = 1/2\sqrt{R} = 0.5$. Figure 2-6b illustrates the effect of varying R at a fixed value of U_S and σ_n. $U_S = 2U_F$ is chosen so that the device is biased well into the "pinning" region. It is seen that the value of R has a pronounced effect on the peak value of the

generation current. R<1 places E_{TP} (or U_{TP}) in the upper half band-gap and the peak current increases [see equation (2.25)]. R>1 places U_{TP} in the lower half bandgap with a subsequent decrease in the peak current.

2.5.3 Trap Admittance

Figure 2-7 illustrates the variation of $G_{it}/\omega C_{ox}$ and C_{it}/C_{ox} vs. frequency, $\omega/2\pi$, from depletion to weak inversion. For simplicity, a constant value of D_{it} is assumed. The other parameters of interest are listed in the figure caption. In depletion ($\overline{U}_S$ = -6, -9), only majority carrier processes take place over each half cycle of the ac signal and, thus, the trap conductance and capacitance are identical to the results of equilibrium theory [16]. This is also the case for $\overline{U}_S$ = -12 (E_{Fn} is positioned approximately at midgap) because R = 1000 places U_{TP} = (1/2)ln(1000) = 3.45 units below midgap. Note, the constant peaks in $G_{it}/\omega C_{ox}$ vs. frequency curves for $\overline{U}_S$ ranging from -6 to -12 are due to the assumed constant value of D_{it}. In a true situation, however, D_{it} is varied to fit the experimental curves at various values of U_S, but is assumed constant at a fixed value of U_S while evaluating the integrals in equation (2.24). This is justified because the width of the scanning function, $f_{ss}(1-f_{ss})$, is only about 2(kT/q) in the depletion region and (kT/q) in the "pinning" region [see Fig. 2-5a]. As $\overline{U}_S$ is increased to -15 and -18 (E_{Fn} is moved into the lower half bandgap), two effects in $G_{it}/\omega C_{ox}$ vs. f

curves (solid line) are noticeable: (a) peak in the curves begins to decrease, (b) peak shifts less rapidly to lower frequencies for a given change in $\overline{U}_S$. Attenuation in the C_{it}/C_{ox} vs. f curves is also apparent. The attenuation in the peak of $G_{it}/\omega C_{ox}$ curves is due to the attenuation of $f_{ss}(1-f_{ss})$ term and the increase of the $(e_p/\sigma_n v_{th} n_{so})$ term in equation (2.17). The steady-state occupancy, f_{ss}, "pins" in the neighborhood of U_{TP} and $f_{ss}(1-f_{ss})$ term decreases rapidly as shown in Figure 2-5a. In the "pinning" region, $(e_p/\sigma_n v_{th} n_{so}) >> 1$ as illustrated in Figure 2-3c. This not only reduces the magnitude of $G_{it}/\omega C_{ox}$ but also slows down the shift in the position of its peak. In other words, as E_{Fn} moves into the deep-depletion region ($\overline{U}_S$ = -15, -18), the ac response of the hole emission process begins to compete with the majority carrier process. In one half cycle of the ac signal, electron emission to the conduction band occurs, and in the other half cycle, both electron capture from the conduction band and hole emission to the valence band occurs. The attenuation in the peak magnitude and the relatively slow displacement of its position to lower frequsncies are important experimental indicators of the "pinning" action. Figure 2-7 also illustrates the attenuation and retrograde action of $G_{it}/\omega C_{ox}$ vs. f curve as a function of R = 100, 10, 1 (dotted line) at fixed $\overline{U}_S$ = -15. This is again a manifestation of the "pinning" of $f_{ss}(1-f_{ss})$ and the increase in the hole emission term, $(e_p/\sigma_n v_{th} n_{so})$, in equation (2.17).

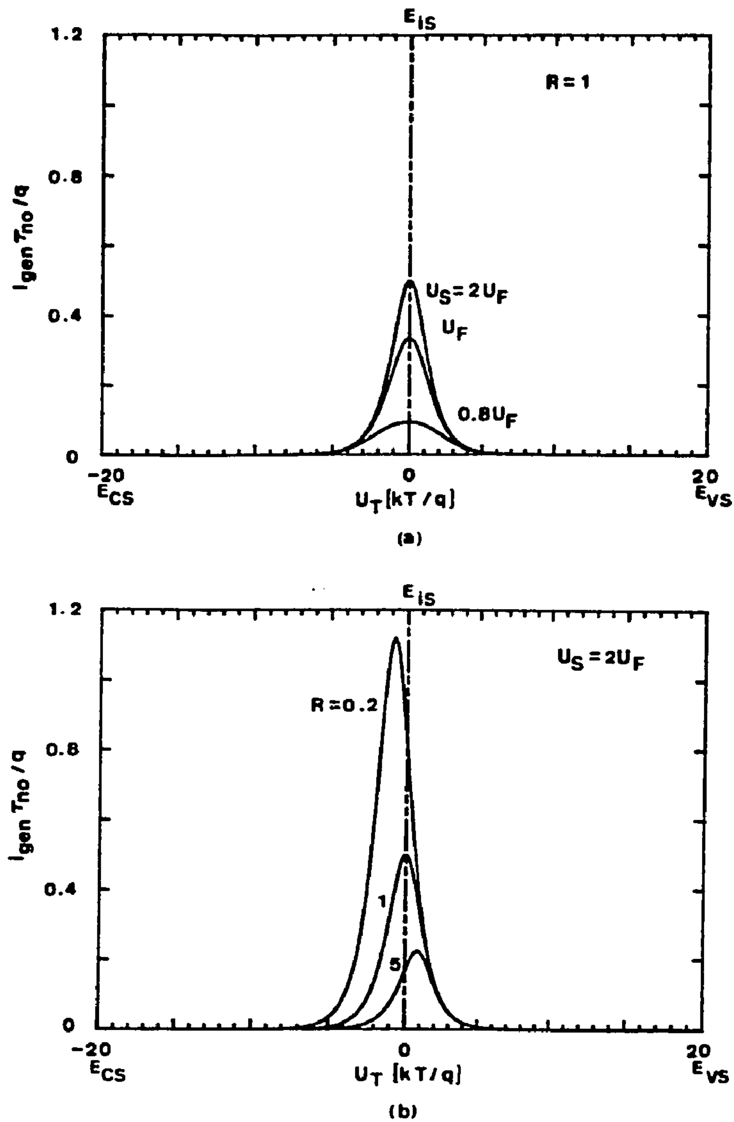


Fig. 2.6 Plots of the normalized generation current $i_{gen}\tau_{no}/q$ vs the normalized trap energy U_T for (a) various values of U_S with R=1, and (b) various values of R with $U_S=2U_F$ (deep-depletion). The parameters used are $U_F=-11.1$ ($N_D=10^{15}cm^{-3}$), T=300K and $V_{SB}=-1V$.

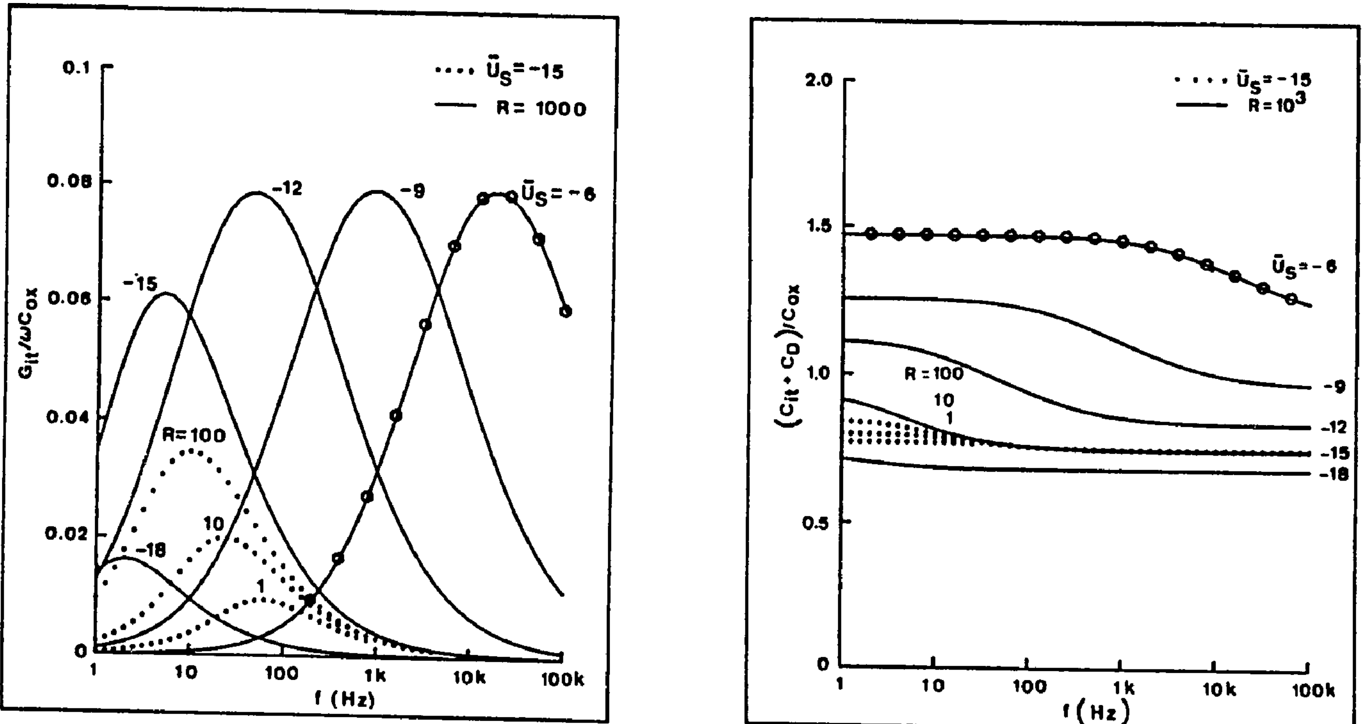


Fig.2.7 Plots of calculated (a) normalized trap conductance $G_{it}/\omega C_{ox}$, and (b) trap capacitance C_{it}/C_{ox} vs frequency f from Eq(2.24) for various values of $\bar{U}_S$ with R=1000 (solid line) and for a fixed value of $\bar{U}_S$ with R as aparameter (dotted line). The parameters used are U_F=-12.25 ($N_D=2\times10^{15}$ cm^{-3}), $\sigma_n=10^{-15}$ cm^2, σ_s=1.5, V_{SB}=-1V, $D_{it}=4\times10^{10}$ $cm^{-2}eV^{-1}$, $C_{ox}=2.79\times10^{-8}$ Fcm^{-2}, T=295K. The symbol O represents the points calculated in depletion ($\bar{U}_S$=-6) by the equilibrium theory in Ref.[16].

2.6 Approximate Solution(Weak "Pinning" Regime)

We may gain some physical insight into the "pinning" phenomenon, created by the nonequilibrium operation of the gated-diode structure, if we rewrite the admittance integral (2.24) in terms of the occupancy function, f_{ss}. We begin by defining

$$X \doteq R^{-1}\exp[U_T + U_F - U_S] = \frac{-1 + [1+\frac{4\alpha f_{ss}}{1-f_{ss}}]^{1/2}}{2} \tag{2.26}$$

where

$$\alpha \doteq R^{-1}\exp[2(U_F - U_S)] \tag{2.27}$$

After some algebraic manipulation, we may write the scanning function as

$$f_{ss}(1 - f_{ss}) = \frac{(1 + X)}{(1 + 2X)} \frac{df_{ss}}{dU_T} \quad . \tag{2.28}$$

Substitution of equations (2.26) and (2.28) into (2.24) yields

$$Y_{it} = \int_{-\infty}^{\infty} dU_S P(U_S) \int_0^1 \frac{j\omega q D_{it} df_{ss}}{[1 + \frac{4\alpha f_{ss}}{1-f_{ss}}]^{1/2} [1 + \frac{j2f_{ss}\omega\tau_n}{1 + [1+\frac{4\alpha f_{ss}}{1-f_{ss}}]^{1/2}}]} \tag{2.29}$$

where $\tau_n = \tau_{no}\exp[U_F - U_S]$. The second portion of the above integral represents the modified Lehovec solution [20], which may be obtained in the limit of $\alpha \to 0$ and constant D_{it}. For a given R,

we may achieve these conditions in depletion where $U_F-U_S<0$. As we near the trap energy $U_T = U_{TP} = U_F-U_S = (1/2)\ln R$ we have $\alpha \to 1$ and $f_{ss} \to 2/3$. As we continue further into the midgap and weak inversion regions, the value of $f_{ss} \to 1$ and the values of G_{it} and C_{it} are attenuated rapidly as seen from an inspection of equation (2.29). Thus, we may picture U_{TP} as an effective "pinning" trap energy which represents the limit of applicability for the Lehovec and Nicollian-Goetzberger equilibrium solutions. We can approximate equation (2.29) into the weak pinning region with the integral

$$\int_0^{f^*_{ss}} \frac{j\omega q D_{it} df_{ss}}{1 + j\omega\tau_n f_{ss}} \tag{2.30}$$

where f^*_{ss} represents the limit imposed by U_{TP} as can be seen by substitution in equation (2.22),

$$f^*_{ss} = f_{ss}(U_T = U_{TP}) = \frac{1}{1 + [1 + R^{1/2} e^{U_S - U_F}]^{-1}} \tag{2.31}$$

with

$$U_{TP} = \frac{1}{2} \ln R \tag{2.32}$$

We may solve equation (2.30) for a constant D_{it} to yield the non-statistically weighted trap conductance and capacitance,

$$\frac{g_{it}}{\omega} = \frac{qD_{it}}{2\omega\tau_n} \ln\, [1 + (\omega\tau_n f^*_{ss})^2] \tag{2.33}$$

$$c_{it} = \frac{qD_{it}}{\omega\tau_n} \text{ arc tan } (\omega\tau_n f^*_{ss}) \tag{2.34}$$

which reduces to the Lehovec solutions as $f^*_{ss} \to 1$. We can observe from equation (2.31) that as we proceed from depletion, where $U_F - U_S < 0$ and $f^*_{ss} = 1$, to midgap and weak inversion regions, the value of f^*_{ss} decreases as we enter the "pinning" regime, thereby decreasing the trap conductance and capacitance according to (2.33) and (2.34). The integral (2.30) is only valid in "weak" pinning and the attempt of this approximate solution is to illustrate the convergence of the nonequilibrium solution to the conventional equilibrium solution. Equation (2.29) is valid for equilibrium and nonequilibrium solutions. Note, f^*_{ss} represents the occupancy at a fixed trap "pinning" level, U_{TP}, with $U_S - U_F$ considered as the variable, whereas f_{ss} [the integration variable in equations (2.29) and (2.30)] scans the interface traps, U_T, across the bandgap, with U_S-U_F fixed during the integration.

2.7 Summary

In this chapter, SRH statistics is applied to the non-equilibrium condition at the Si-SiO_2 interface created by the removal of minority carriers with a gated-diode structure. The minority carrier capture process is eliminated in the absence of minority carriers whereas the emission of minority carriers remains active. The majority carrier capture and emission processes are

also present. The resulting steady-state occupancy function of the midgap interface traps deviates considerably from the Fermi-Dirac distribution. The occupancy function, in fact, "pins" at a trap level which is $(1/2)\ln(\sigma_n/\sigma_p)kT$ eV below (above if $\sigma_n/\sigma_p < 0$) the intrinsic level E_{is} at the interface. Interface traps below this "pinning" level always remain occupied due to the dominant hole emission process regardless of the surface potential in deep-depletion.

Based on this non-equilibrium occupancy function, expressions are derived for small signal trap conductance and capacitance as functions of the signal frequency, surface potential and various trap parameters. The theoretical simulations predict the rapid decay or "pinning" in the normalized conductance $(G_{it}/\omega C_{ox})$ vs the frequency curves at a critical value of surface potential which places the majority carrier Fermi level at the so called "pinning" trap level. This provides a unique method of determining the ratio σ_n/σ_p and the interface trap density for the midgap traps.

In addition, pair generation via the interface traps has been analyzed. It has been found that the traps located at the so called "pinning" trap level are the most efficient generating centers and the generation rate is proportional to $\sqrt{\sigma_n \sigma_p}$.

CHAPTER III

EXPERIMENTAL EVALUATION OF $Si-SiO_2$ INTERFACE TRAPS

3.1 Fabrication

The measurements were performed on a large area, gated-diode structure with n type, 4-8 Ω.cm, (100) Si substrate. The photograph of the gated-diode and its cross-section are shown in Fig.3-1. The processing sequence consisted of p^-, p^+ and n^+ diffusion steps followed by a Phosphorus gettering step. Gate oxidation was carried out at 900^0C in 5% HCl and dry O_2 followed by the 15 min N_2 anneal at 900^0C (d_{ox} = 960 A). Subsequently, 545 A of APCVD Si_3N_4 and 5000 A of n^+ doped poly-silicon films were deposited. After defining the gate region, Forming gas anneal at 500^0C for 15 min was done. Next, contact window were opened and Al was put down with electron beam, defined and sintered.

3.2 Experimental Setup

The experimental MOS gated-diode structure and the band diagram are shown in Figure 2-1 for an n-type substrate. A reverse-bias, V_{SB}, sufficient to remove minority carriers, is applied across the source to bulk junction. In response to a small-signal excitation (5Hz-30KHz) superimposed on a bias V_B, admittance measurements are performed between the gate and bulk electrodes. These measurements are recorded as a function of V_B and the measurement frequency

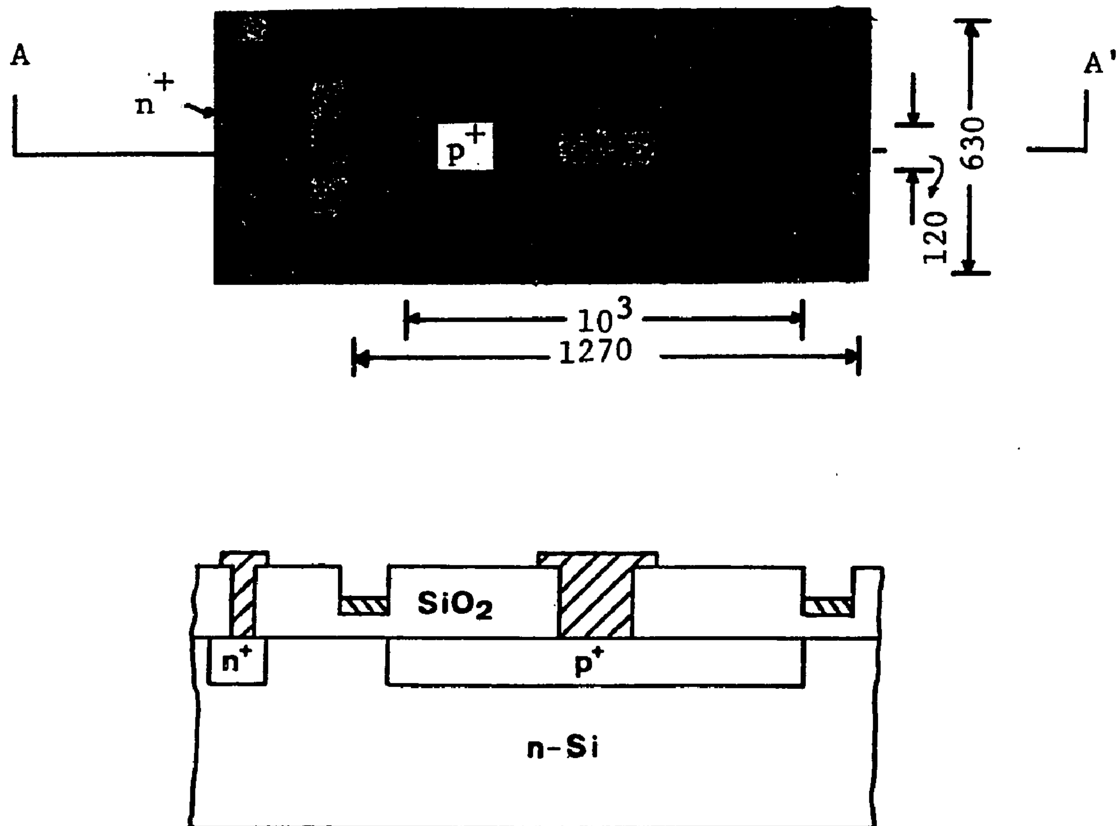


Fig. 3.1 Photomicrograph of the gated-diode structure and the schematic of its cross-section along the plane AA'. All dimensions are in μm.

by sensing the small signal displacement current in the gate electrode with a two-phase, lock-in amplifier. The reverese bias, V_{SB}, is not varied since its sole purpose is to effectively remove the minority carriers from beneath the gate electrode over the range of bias V_B. Variation of the bias, V_B, effectively scans the free electron Fermi level, E_{Fn}, across the entire bandgap and the $Si-SiO_2$ interface states.

The block diagram of the experimental setup is illustrated in Fig.3-2. The desktop HP-9836 computer controls a number of peripherals which include the plotter, printer, disk-drive, D/A and A/D converters and a function generator. A small sinusoidal signal (25 mV p/p) from the function generator is superimposed on a dc bias and applied to the bulk terminal of the device. A fixed reverse bias of 1.5 volt is maintained across the source and bulk electrodes with a battery. The small signal ac current at the gate is measured through a cascade of a current pre-amplifier and a two-phase lock-in amplifier. The analog outputs of the lock-in amplifier are sampled by the A/D converter and averaged by the computer. At low frequencies, the time-constant of the lock-in amplifier must be increased to reduce the effective noise bandwidth and this is followed by the time-averaging of the analog signals over a period comparable with the amplifier time-constant. The resultant signal to noise ratio is improved with this technique at low frequencies.

3.3 Experimental Results

The measurements are performed at different frequencies (5Hz-30KHz) as a function of bias. The conductance and capacitance at a fixed bias are extracted from the measured values by means of equations (2.20) and (2.21) and plotted as a functions of the frequency. In addition to the above measurements, high frequency (1 MHz) C-V, gated-diode I-V [44] and the quasi-static C-V measurements are performed. The doping density in the bulk, N_D, is calculated from the high frequency C-V curve. $\overline{U}_S$ vs V_G curve ($V_G = -V_B$) and an estimate of D_{it} is obtained from the quasi-static measurement. The gated-diode I-V curve provides an estimate of the surface leakage and a comparsion between interface and bulk trap densities [44].

Figure 3-3 illustrates the experimental points and the theoretical fits to the $G_{it}/\omega C_{ox}$ and $(C_{it}+C_D)/C_{ox}$ versus frequency data. The "pinning" of the trap conductance in the range $-16.5 > \overline{U}_S > -19.5$ is quite apparent. In addition, the peak of the $G_{it}/\omega C_{ox}$ curves does not shift much in frequency as discussed in section 5.3. The onset of "pinning" ($\overline{U}_S = -16.5$) corresponds to the trap energy $U_{TP} = (1/2)\ln R = U_F - \overline{U}_S = 4.5$. This value of U_{TP} gives $R = 10^4$, a relatively high value.

Figure 3-4 shows the quasi-static curves with V_{SB}=0 and -1.5 volt which are integrated to give $\overline{U}_S$ vs V_G curves shown in Fig. 3-5. The values of $\overline{U}_S$ used to fit the experimental data in Fig. 3-3 were

Fig. 3.2 Block diagram of the computer-controlled data-acquisition system used for measurements in this thesis.

obtained from the $\overline{U}_S$ vs V_G curve with V_{SB} = -1.5 volt. Approximate values of D_{it} are obtained from the quasi-static curve with V_{SB} = 0. N_D is calculated from the hf C-V curve to be $2.3x10^{15}cm^{-3}$ which corresponds to U_F=-12.0. These parameters are used in equation (2.24) to fit both $G_{it}/\omega C_{ox}$ and $(C_{it}+C_D)/C_{ox}$ curves. The parameters required for the fit are D_{it}, R, σ_n, σ_s and $\overline{U}_S$. Comparsion of the D_{it} obtained from the quasi-static and admittance measurements is illustrated in Fig.3-6. Figure 3-7 shows the energy distribution of σ_n. The reverse source current I_S, as a function of the bulk bias V_B with V_{SB} = -1.5 volt, is shown in Fig.3-8. It is noted that the contribution due to the generation via interface states is not present and I_S saturates at the onset of inversion to a value governed by the generation in the bulk. This observation is consistent with the high value of R as discussed in the next section.

Trap conductance and capacitance curves for another device with a higher value of D_{it} are shown in Fig.3-9. There is some structure in the conductance curves which may be attributed to either lateral non-uniformities or the rapid variation of the trap parameters (σ_n, σ_p, D_{it}) with energy. No detailed attempt is made to fit the theory to these curves. The line through the data points is only meant to distinguish between different curves. The values of $\overline{U}_S$ indicated in Fig.3-9 were obtained from the hf C-V curve and are consistent with the depletion capacitance obtained from the $(C_{it} + C_D)/C_{ox}$ curves

(Fig.3-9) at high frequencies. N_D as calculated from the hf C-V curve is $2.3x10^{15}$ cm^{-3}. σ_n and D_{it} are estimated from the $G_{it}/\omega C_{ox}$ curve with $\bar{U}_S$ = -10 to be about $1x10^{-15}$ cm^{-2} and $6.0x10^{11}$ $cm^{-2}eV^{-1}$, respectively. The onset of "pinning" is approximately estimated at $\bar{U}_S$ = -14 which gives R=50, a relatively low value. The result of the gated-diode I-V measurement is shown in Fig.3-10. The contribution due to the generation via interface states is significant ($2.5x10^{-10}$ Amp. for the device area of $6.7x10^{-3}$ cm^{-2}) due to the high value of D_{it} and the low value of R as discussed in the next section.

3.4 Discussion

Experimental results clearly indicate the "pinning" of the normalized trap conductance as the electron Fermi-level is moved below the "pinning" trap level U_{TP}. This result comfirms that the hole emission from the interface states is indeed the dominant process for the states located in the lower half of the bandgap in n-type Si. The "pinning" action is in contrast with a MOS capacitor where the peak in the trap conductance vs. frequency curve increases while the width of the curve decreases in the weak inversion region [19].

The high value of $\sigma_n(2.5x10^{-14}cm^2)$ with a low value of D_{it} ($7.0x10^{10}eV^{-1}.cm^{-2}$) and the lower value of σ_n ($1x10^{-15}cm^2$) with the higher value of D_{it} ($6.0x10^{11}eV^{-1}.cm^{-2}$) are obtained in the two

samples studied. The results corresponding to the two samples are plotted in Figs.3-11 and 3-12 and summarized in Table 1. It is interesting to note that the value of R decreases with increase in D_{it}. A similar dependence was reported by Cooper and Schwartz with an equilibrium technique [19]. However, more data is required on samples with different annealing conditions to substantiate this result.

The values of R, σ_n and D_{it} determine the surface leakage currents as discussed in section 2.5.2. Using $U_S=2U_F$ (strong "pinning") in equation (2.25) and assuming that only the traps located within +kT around U_{TP} contribute to pair generation (see Fig. 2-6), total generation current I_{gen} can be estimated by

$$I_{gen} = (q\sigma_n v_{th} n_i / 2\sqrt{R}) 2kT \, D_{it} A$$

$$= q v_{th} n_i kT D_{it} A \sqrt{\sigma_n \sigma_p} \quad (3.1)$$

where A is the device area and kT = 0.025 eV at room temperature. Equation (3.1) is plotted in Fig.3-13 with R as an independent variable and σ_n, D_{it} as device parameters. It is noted that $I_{gen}=5\times10^{-11}$A is negligibly small for the sample with $\sigma_n = 2.5\times10^{-14}cm^2$, $D_{it} = 7\times10^{10}eV^{-1}cm^{-2}$ and $R = 10^4$ as observed in the gated-diode I-V measurement (see Fig.3-8). However, $I_{gen}=2.2\times10^{-10}$A for the second sample with $\sigma_n = 1\times10^{-15}cm^2$, $D_{it} = 6.0\times10^{11}eV^{-1}cm^{-2}$ and R = 50. This agrees well with the experimental results of

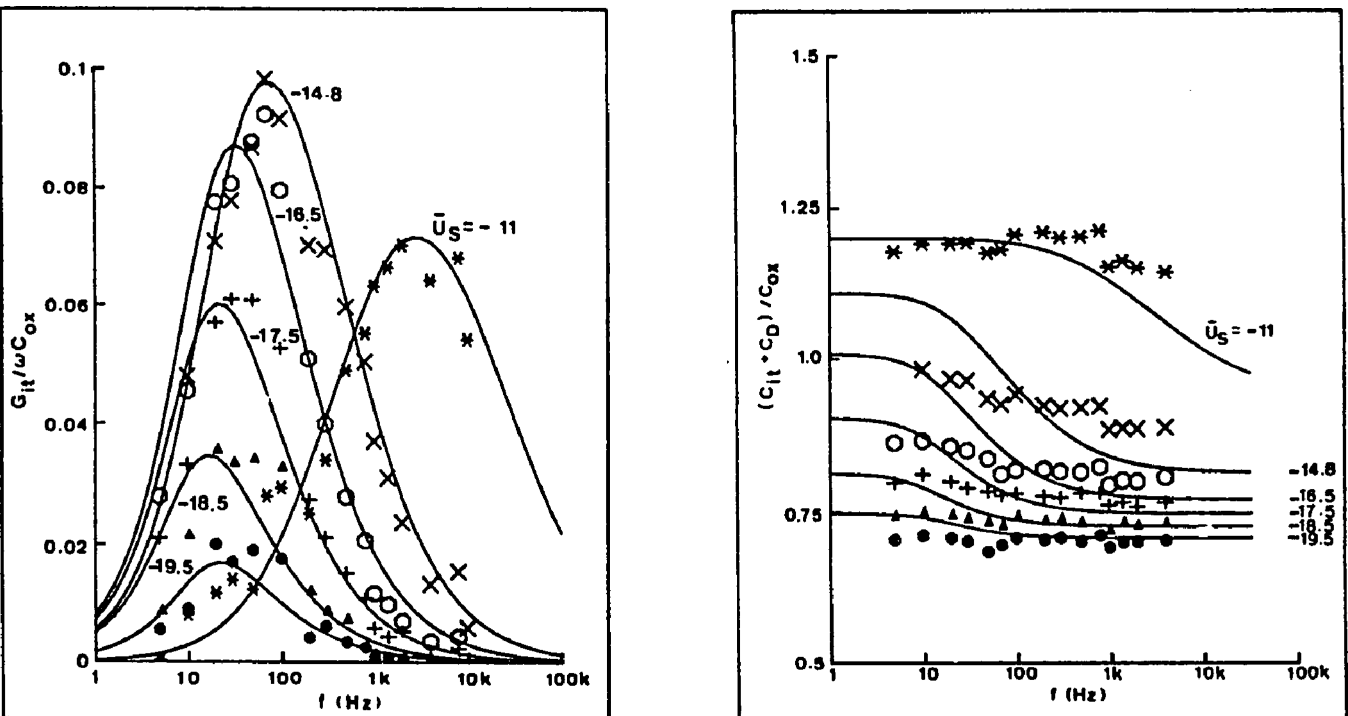


Fig. 3.3 Plots of measured (a)normalized trap conductance $G_{it}/\omega C_{ox}$, and (b)normalized trap capacitance C_{it}/C_{ox} vs frequency data points (*,X,O,+,▲,●) and the theoretical fits (solid line) using Eq.(2.24). Device parameters used are U_F=-12.0(N_D=2.3x$10^{15}$$cm^{-3}$), R=$10^4$, σ_s=1.2, C_{ox}=2.79x10^{-8} Fcm^{-2}, T=295K, σ_n and D_{it} are plotted as functions of the trap energy in Figs. 3.6 and 3.7.

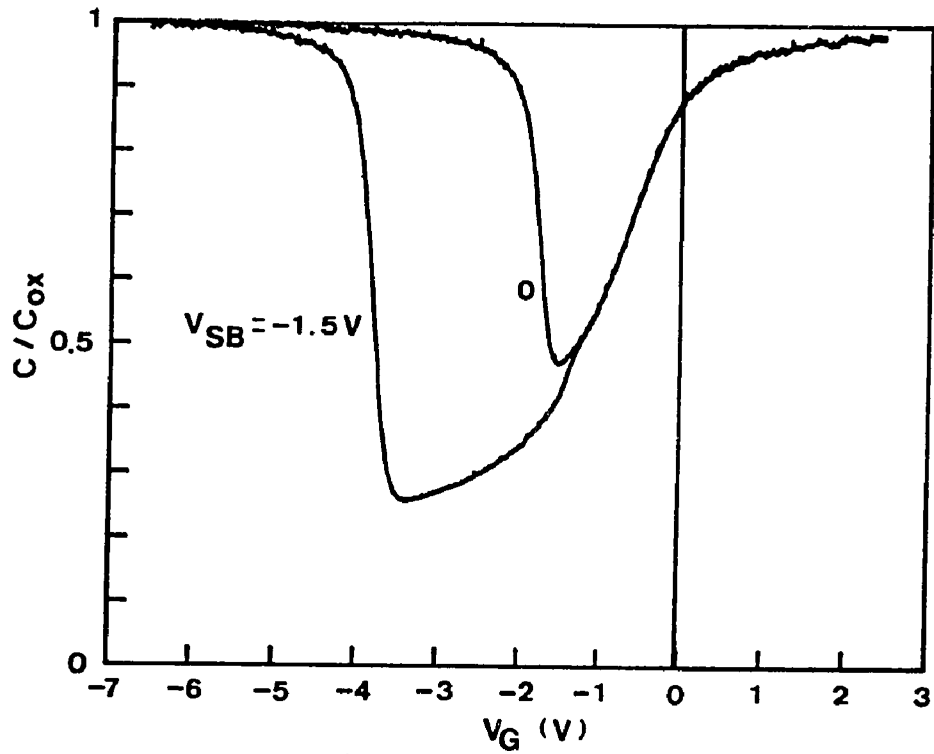


Fig. 3.4 Quasi-static curves with V_{SB}=0 and −1.5V. V_G denotes the bias applied to the gate relative to the bulk.

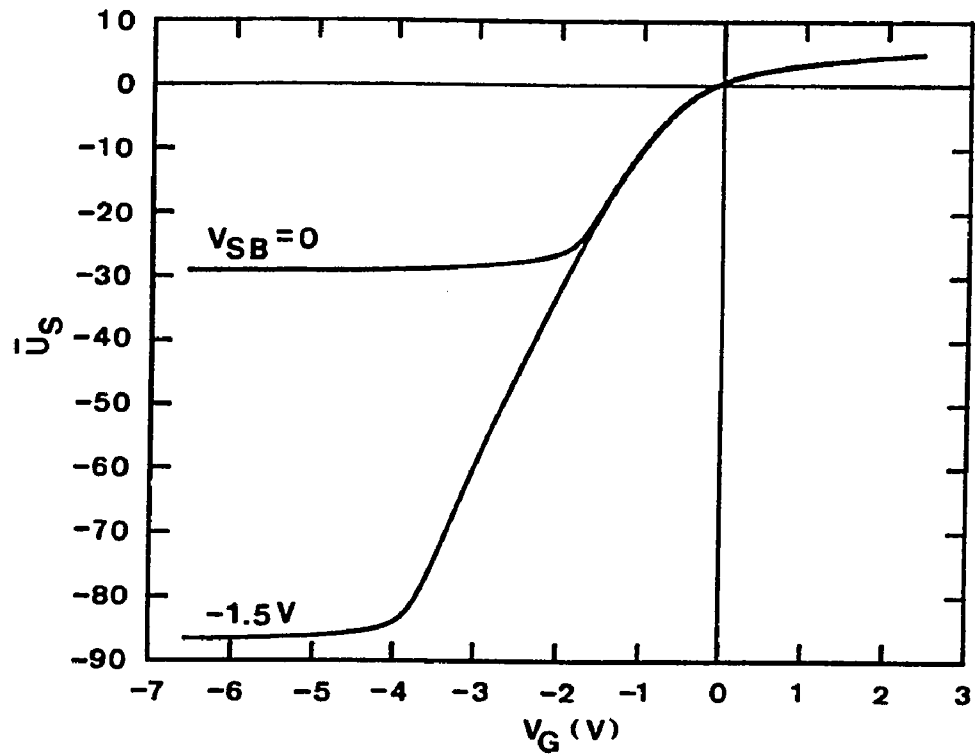


Fig. 3.5 Normalized mean surface potential as a function of the gate bias V_G for V_{SB}=0 and -1.5V

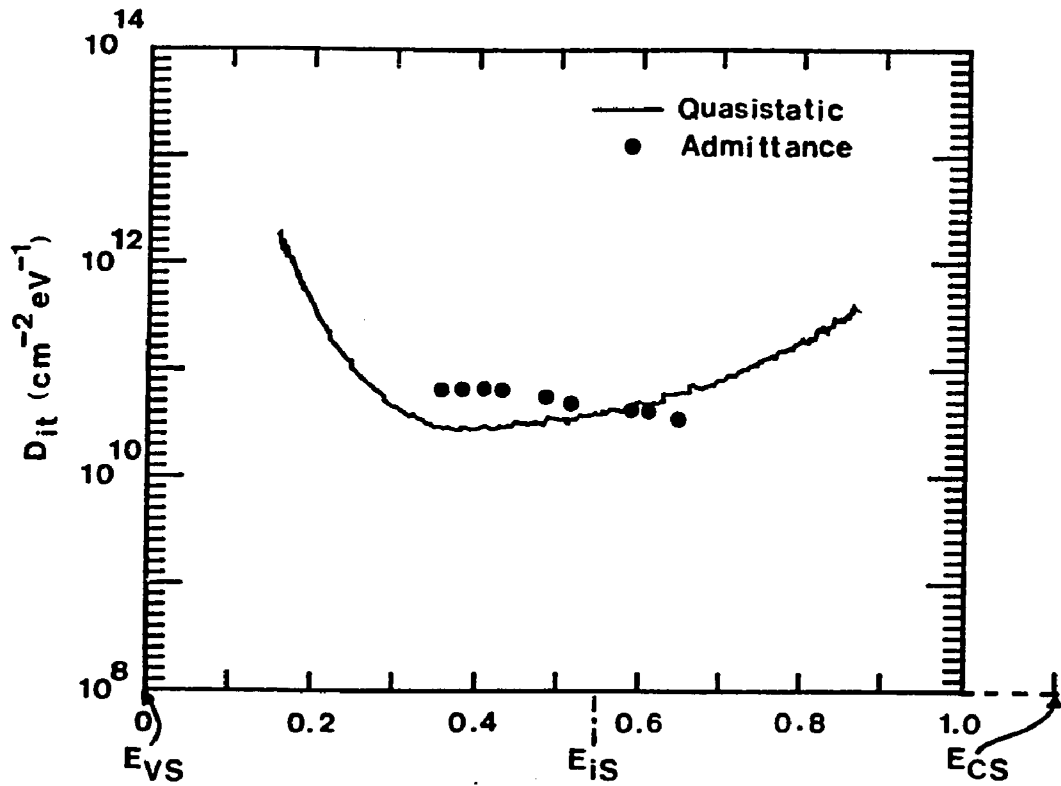


Fig. 3.6 Distribution of the interface trap density D_{it} as a function of the trap energy. The solid line represents the distribution calculated from the quasi-static curve with $V_{SB}=0$ and the points indicate the distribution obtained by the admittance measurements.

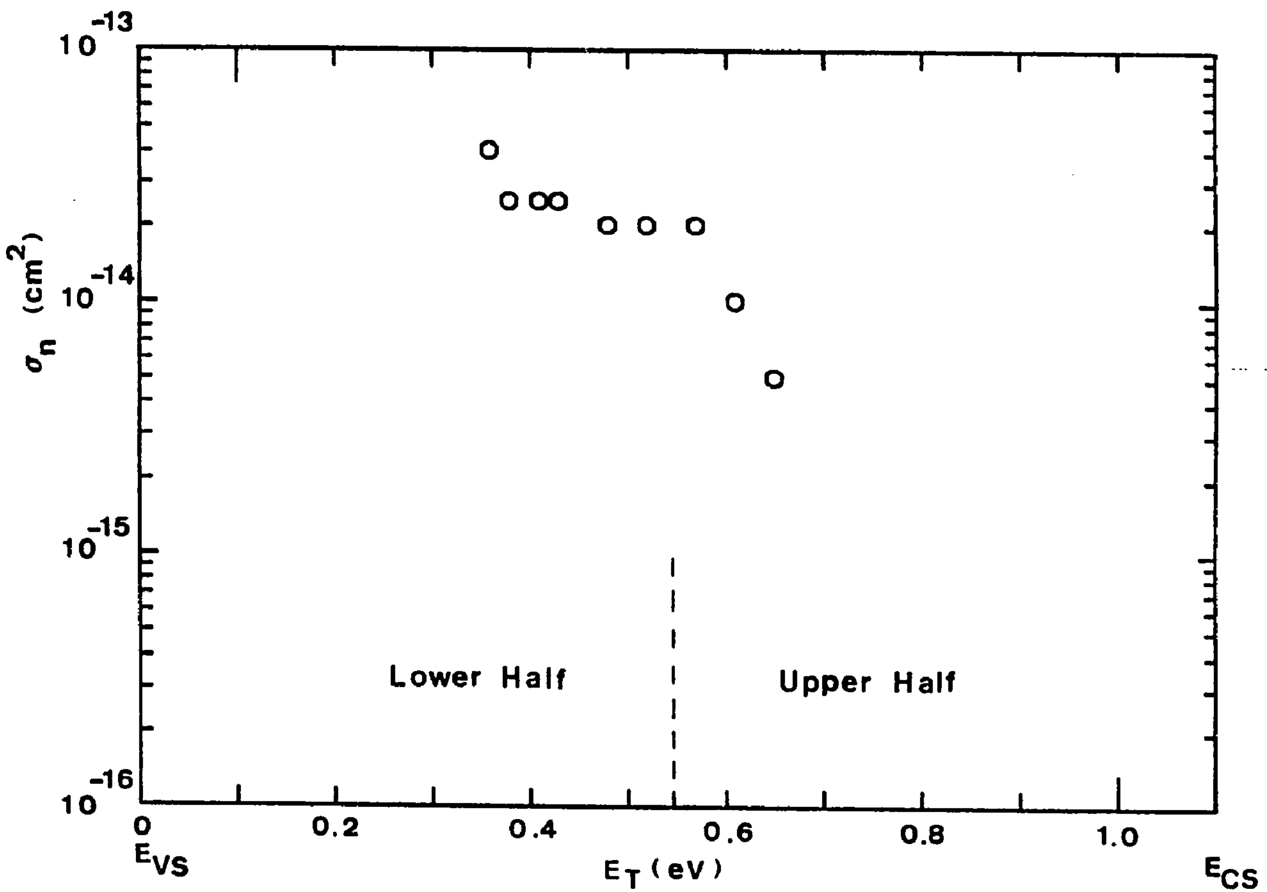


Fig. 3.7 Electron capture cross-section as a function of energy in the midgap region at room temperature.

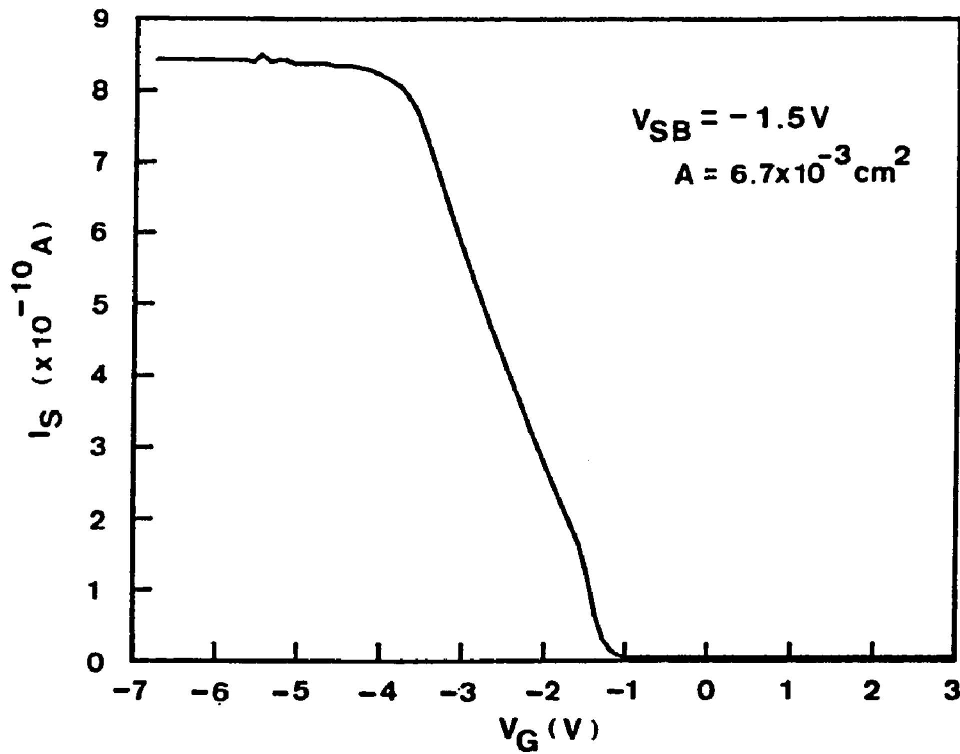


Fig. 3.8 Source current as a function of the gate bias V_G(relative to bulk) with the source/bulk junction reverse biased at -1.5V. Note that the semiconductor under the gate region inverts at about V_G=-4V as evident from Figs. 3.4 and 3.5.

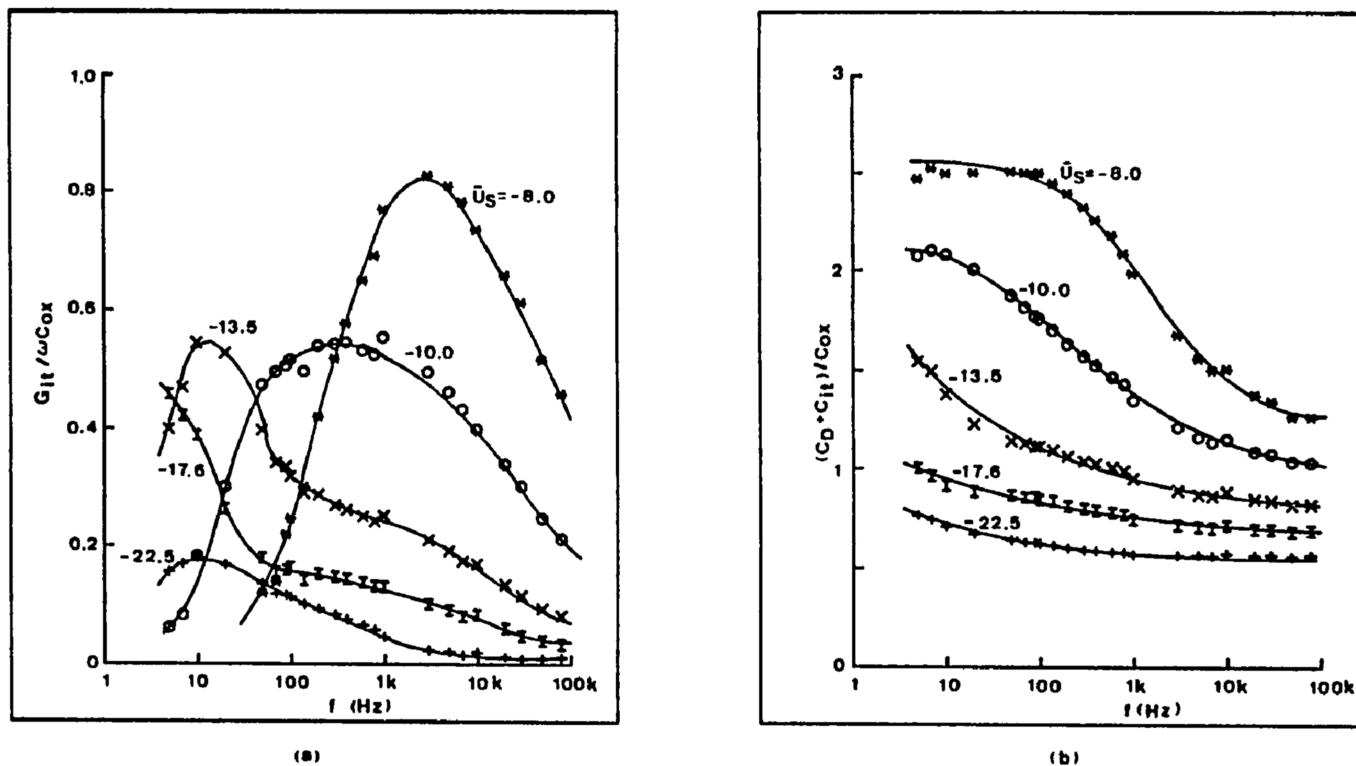


Fig. 3.9 Plots of measured (a) normalized trap conductance, and (b) capacitance vs frequency data points (*,O,X,I,+) for the sample with relatively high value of the interface density. The solid lines connecting different points are meant to distinguish between different curves and do not represent theoretical fits.

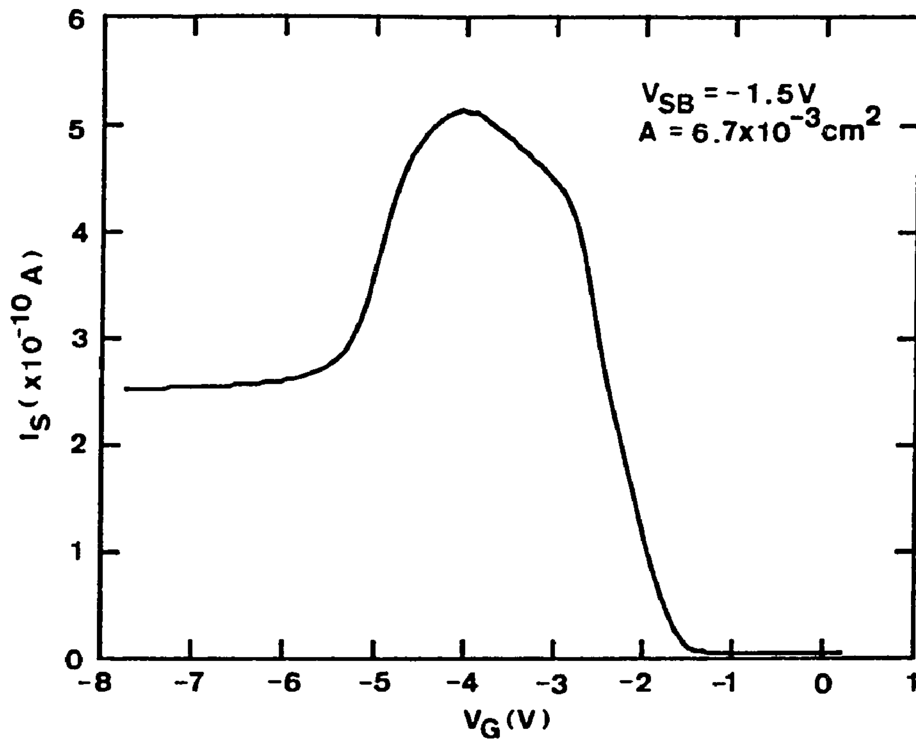


Fig. 3.10 Source current as a function of gate bias V_G (relative to bulk) with the source/bulk junction reverse biased at -1.5V. Note that there is a contribution of about 2.5×10^{-10}A due to the generation via interface traps.

Fig.3-10. We can conclude that the parameters obtained by fitting the admittance curves are consistent with the surface leakage current data.

3.5 Summary

In this chapter, the experimental setup for performing the small signal admittance measurements is described. Two samples (<100>, n-substrate) with no intentional variation in processing are studied. "Pinning" of the conductance curves is experimentally confirmed for both the samples. By carefully fitting the normalized trap conductance and capacitance vs frequency curves, midgap trap parameters D_{it}, σ_n, R, $\overline{U}_S$ and σ_S are extracted. In addition, leakage current measurements are also recorded. The trap parameters obtained by fitting the admittance data are found to be consistent with the surface leakage measurements.

The most striking result of this study is the high value of the ratio $R=\sigma_n/\sigma_p=10^4$ corresponding to a low value of $D_{it}=7.0\times10^{10}$ $eV^{-1}cm^{-2}$ and vice-versa for the interface traps located in the lower half bandgap of the n-type Si. This is in agreement with the observations reported in reference [19]. The fact, that the ratio R is high, indicates donor like traps in the lower half bandgap for n-type Si. This agrees with the interface model of trivalent Si alongwith a distribution of stretched and distorted bonds [45].

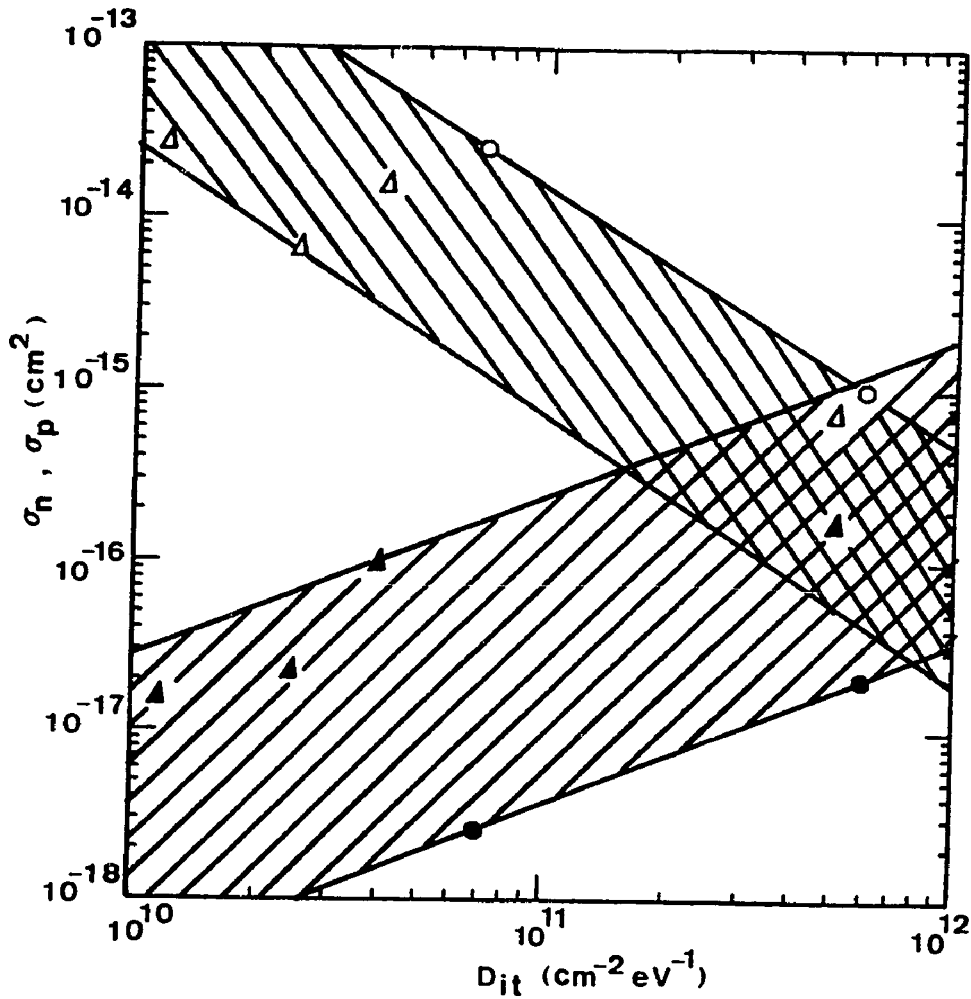


Fig. 3.11 Electron and hole capture cross-sections vs interface trap density around midgap for the two samples discussed in this work (o- σ_n, •- σ_p). Also included are the points reported by Cooper and Schwartz (Δ- σ_n, ▲- σ_p) in reference [19]. The spread in the results is indicated by the shading.

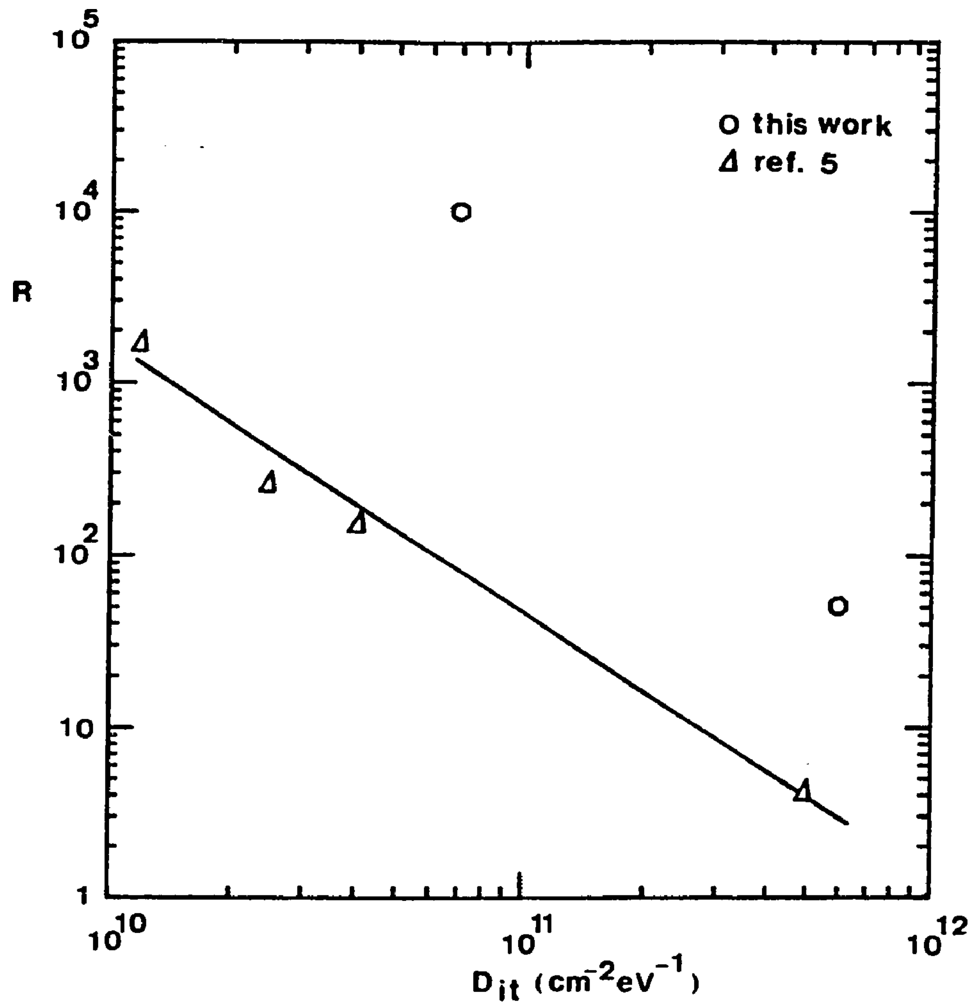


Fig. 3.12 Capture cross-section ratio $R=\sigma_n/\sigma_p$ vs interface trap density D_{it} around midgap.

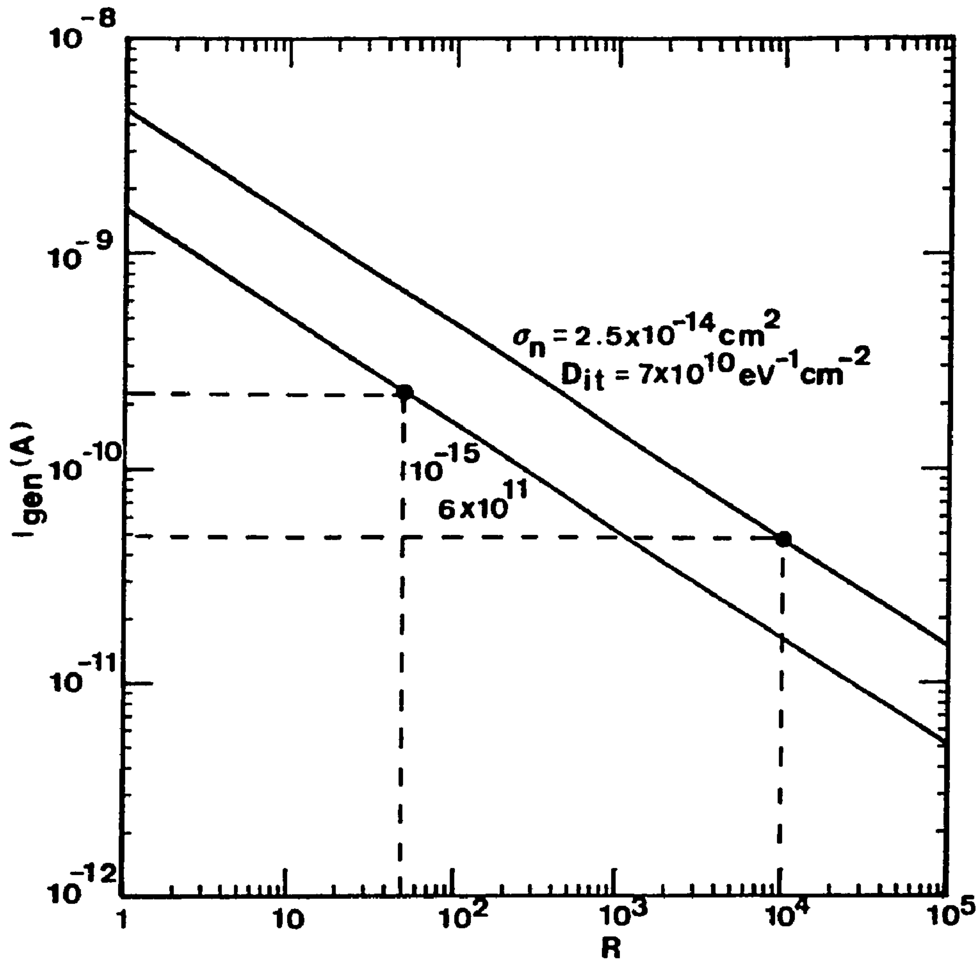


Fig. 3.13 Calculated generation current [eq. (3.1)] due to the interface traps vs R for different values of σ_n and D_{it}. The two points indicate the results for the two samples discussed in the text.

TABLE I. Interface trap parameters around midgap for the two samples as measured by the admittance technique.

Sample	D_{it} ($cm^{-2}eV^{-1}$)	σ_n (cm^2)	σ_p (cm^2)	R
1	7.0×10^{10}	2.5×10^{-14}	2.5×10^{-18}	10^4
2	6.0×10^{11}	1.0×10^{-15}	2.0×10^{-17}	50

CHAPTER IV

INTRODUCTION TO MULTI-DIELECTRIC STRUCTURES

4.1 Device Structure and the Principle of Operation

The conventional Metal-Nitride-Semiconductor (MNOS) sandwich and its ideal energy band diagram are shown in Fig4-1. The nitride film is usually about 500 A thick and is deposited by the Low Pressure Chemical Vapor Deposition(LPCVD) technique. The oxide is thin, about 10-30 A, and is either formed by an initial preclean and insitu reaction within the LPCVD chamber [40] or thermally grown at about 900^0C by an O_2/N_2 partial- pressure method [46].

The operation of the device is illustrated in Fig4-2. A voltage, positive with respect to the silicon, is applied to the gate of the MNOS device(Fig4-1a). This causes a high field, ξ_{oxt}, in the thin oxide layer. The Fig4-2 is drawn to the scale with ξ_{oxt}=8MV/cm and ξ_N=5MV/cm assuming no charge in the nitride layer. As a consequence of this action, several processes may take place: (a) direct (or modified FN) band to band tunneling of electrons from the semiconductor to the conduction band of Si_3N_4 and subsequent capture by the relatively deep traps in the nitride layer, (b) direct band to trap tunneling of electrons from the semiconductor to the traps in the nitride layer, (c) direct trap to band tunneling of

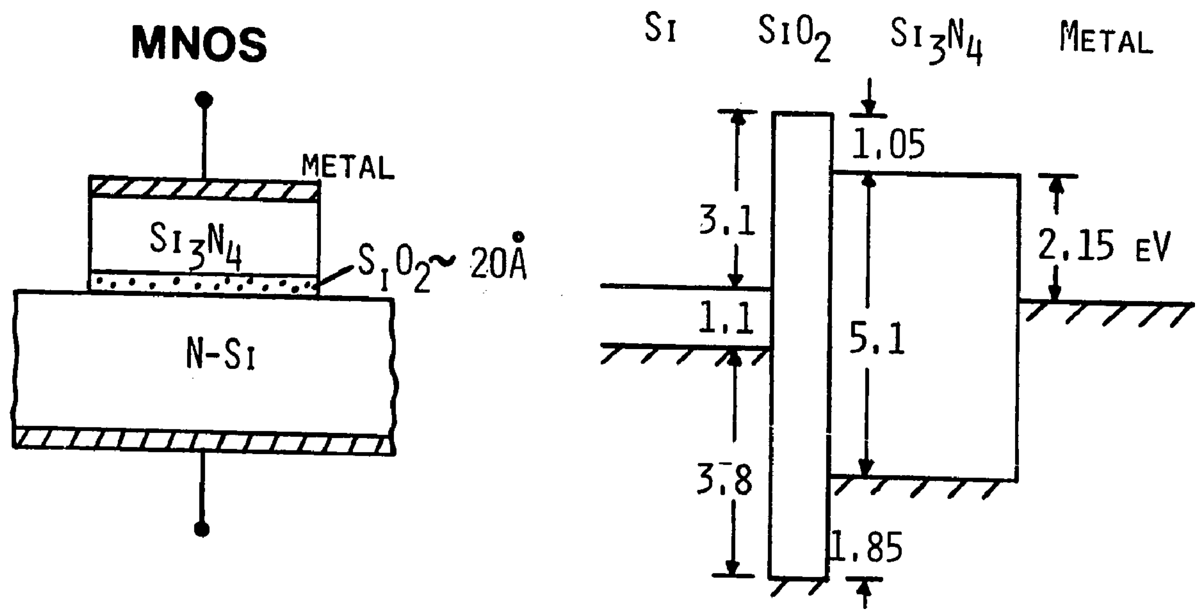
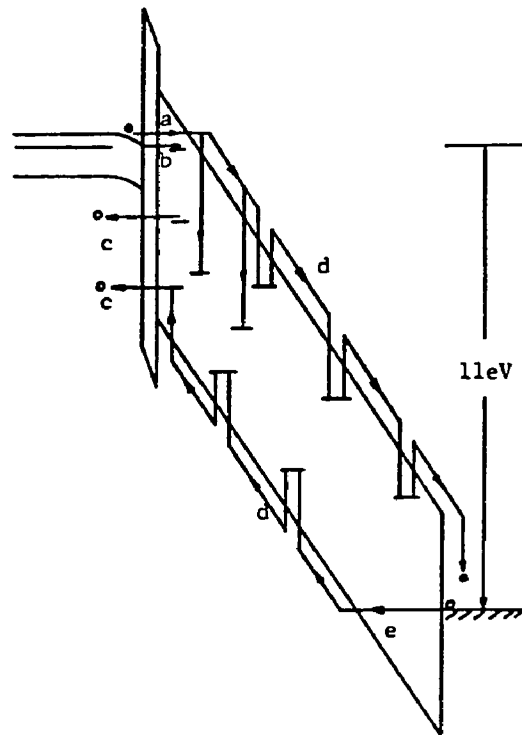


Fig. 4.1 (a) The conventional MNOS sandwich, and (b) its ideal energy band diagram.

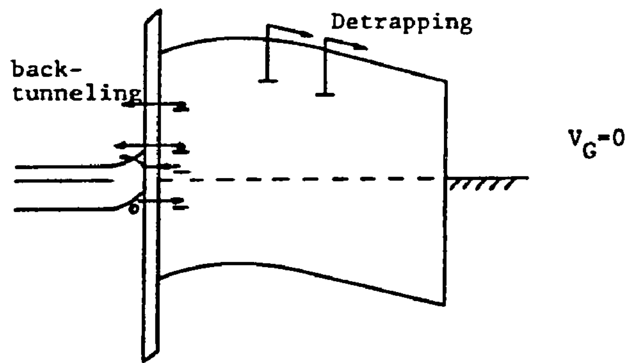
holes from the traps in Si_3N_4 to the valence band of Si, (d) Poole-Frenkel conduction of electrons and holes in Si_3N_4 via shallow traps, and (e) hole injection from the metal into the valence band of the nitride. Provided that the last process is not dominant, the net result is a buildup of negative charge in the nitride layer which shifts the flatband voltage of the device in the positive direction. This is called the 'ERASE' operation for the p-channel device. When the bias at the gate electrode is returned to the substrate potential, the internal electrostatic fields are set up due to the stored charge. The energy band diagram for this case is shown in Fig4-2b. The field in the tunneling oxide and in the part of the nitride that is close to the oxide-nitride interface favors "back-tunneling" (trap to band) of the previously injected electrons from the traps to the semiconductor. In addition, the field deeper into the nitride tends to favor "detrapping" of electrons which moves the centroid of the negative charge away from the oxide-nitride interface. Both of these mechanisms cause a deterioration in the charge "retention" of the device.

The 'WRITE' operation of the p-channel device involves the application of negative bias on the gate relative to bulk. In the Write mode, all the processes mentioned above for the Erase mode are inverted and the flatband voltage shifts in the negative direction.

The 500 A nitride layer necessiates the application of large



(a)



(b)

Fig. 4.2 Operation of the conventional MNOS structure: (a) positive bias on the gate,and (b) zero bias on the gate relative to bulk. Various current components are explained in the text.

voltages(excess of 25V) to write or erase the device in a short time(10 μs). The basic considerations for scaling down of the MNOS structure are discussed by Hampton and Cricchi [47]. It is estimated that a nitride thickness of 190 A is nearly the lower limit for the scaled down MNOS memory devices because of the charge centroid considerations [47]. A 10-V programmable MNOS device with 195 A thick Si_3N_4 has been reported [48].

An attractive approach to scaling has been proposed by Suzuki et al [46]. It involves the deposition of an extra layer of the blocking oxide(40~60 A) on top of the Si_3N_4 layer. The resulting Metal-Oxide-Nitride-Oxide-Semiconductor(MONOS) structure and its ideal energy band diagram are shown in Fig4-3. Using a 22 A tunneling oxide, 30 A Si_3N_4 layer and 33 A of blocking oxide, ±6 V operation was demonstrated for the first time [46].

As shown in Fig4-3, in the scaled-down MONOS structure, the thin nitride is confined by the top oxide and the tunnel oxide. Consequently charge, injected from the Si through the tunnel oxide, is blocked at the top oxide-nitride interface(Fig.4-4a) resulting in an increase in the trapping efficiency of the injected charge. The blocking action of the top oxide removes any limitation on reduction of the nitride thickness. The scaled-down MONOS structures enjoy several advantages. First of all, injection from the gate electrode is minimized for both the gate polarities. This fact helps in two

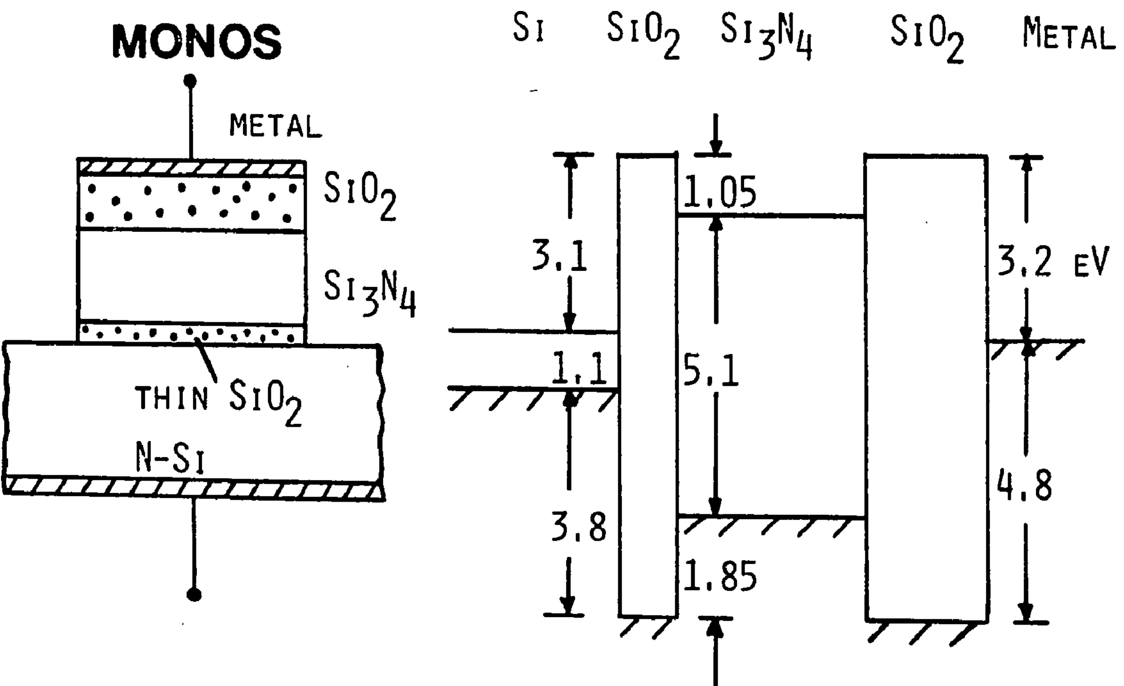


Fig. 4.3 (a) The scaled down MONOS structure, and (b) its ideal energy band diagram.

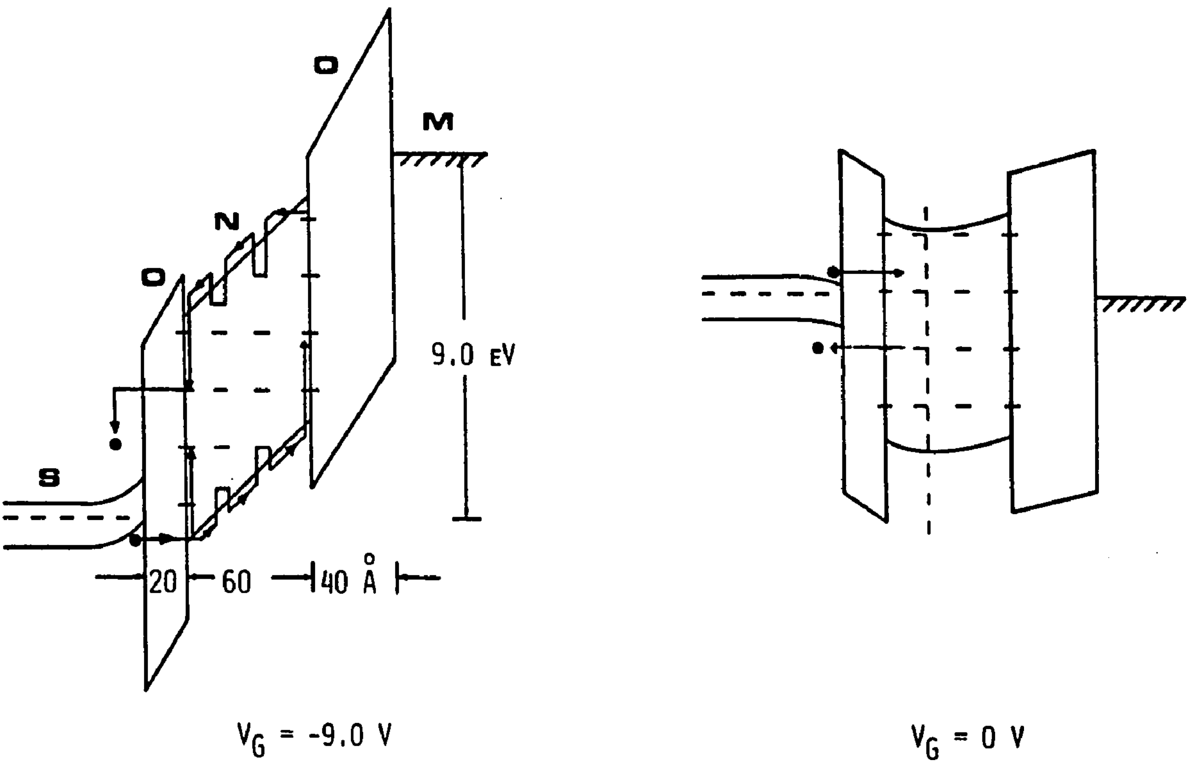


Fig. 4.4 Operation of the MONOS structure with no injection from the gate electrode.

ways: (a) injection from the gate tends to reduce the flatband shift and therefore, the memory window is widened by reducing this injection; (b) the generation of traps in the nitride layer and the interface traps at the $Si-SiO_2$ interface due to the excess charge transport through the sandwich is minimized [49, 50]. This improves the endurance of the MONOS device. Next, a large number of interface traps attributed to Oxygen are located at the top oxide-nitride interface. These deep interface traps serve as efficient memory sites in the scaled down MONOS structure. Therefore, the structure provides a large memory window inspite of decrease in the nitride thickness.

4.2 Historical Perspective

Reviews of the historical developments of the MNOS device have been given by Chang [4] and Verwey [5]. After the deposition technique for silicon-nitride and the potential of using it as a fabrication tool was reported by Doo et al [51, 52, 53], the instability of the metal-insulator-semiconductor(MIS) structure involving silicon-nitride was observed and investigated by many [54, 55, 56, 57, 58]. The memory potential of the MNOS structure was first recognized by Wegener et al [59] and later investigated by Pao et al [60].

4.2.1 Theories

Continuity Equation Models

To model the charging and discharging behavior of the MNOS device, Frohman-Bentchkowsky and Lenzlinger [61] used the one-dimensional continuity equation. They assumed that the charge is trapped only at the oxide-nitride interface and treated the thick-oxide(>50 A) case when the oxide current is dominated by the Fowler-Nordheim tunneling. The nitride current was assumed to be due to the Poole-Frenkel conduction. The continuity equation was numerically integrated to yield the switching and charge-retention characteristics. Neglecting the nitride conduction current, Sevensson and Lundstrom [62] also used the continuity equation to calculate the transient behavior of MNOS devices.

Direct Band-Trap Tunneling Models

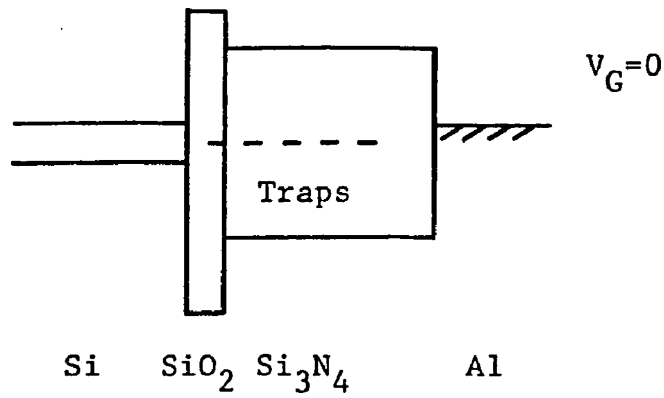
Direct-tunneling models apply only to thin-oxide(<50 A) devices in which the oxide tunneling current is much larger than the nitride conduction current. To interpret the MNOS data obtained by Wallmark and Scott [63], a direct-tunneling model was developed by Ross and Wallmark [64]. They assumed that the traps are spatially uniform and located at one energy level near the intrinsic Fermi level of the silicon(Fig4-5a). Under positive gate bias, holes are assumed to tunnel from the trap sites to the Si valence band(Fig4-5b). With negative gate bias, electrons tunnel from the trapping site into the

conduction band of Si(Fig4-5c). Ross and Wallmark also assumed that the tunneling probability is similar to that for tunneling through a rectangular potential barrier and independent of time. They found that the flat-band voltage shift increases roughly with the logarithm of the switching voltage pulse-width. This is in general agreement with the experimental obervations except the fact that the theory does not predict saturation as the pulse-width approaches infinity.

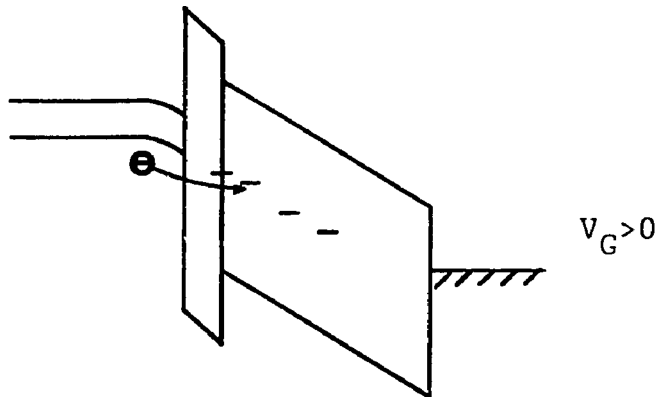
Dorda and Pulver [65] have obtained a good fit to the experimental data by assuming a trap distribution shown in Fig4-6 and neglecting nitride conduction. Their theory includes the effect of the transferred charge on the oxide field and the transition probability. The model predicts the logarithmic behavior of charging and discharging and eventual saturation.

To treat a more realistic distribution, Ferris-Prabhu [66, 67, 68] has extended the direct-tunneling theory to cover both spatially and energetically distributed trapping centers(Fig4-7). His series solutions assume time independent tunneling probability and predict the time dependence of the total charge transfer to be initially linear, then logarithmic and finally reaching saturation.

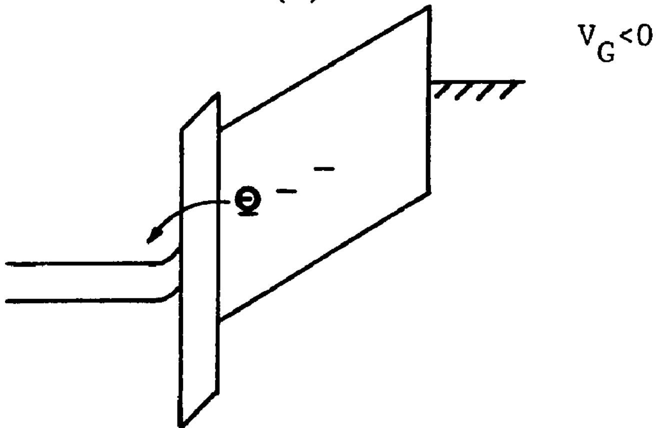
White and Cricchi [69] treated the direct tunneling from the single deep monoenergetic donor and acceptor trap close to the



(a)



(b)



(c)

Fig. 4.5 Energy band diagrams of the switching mechanism used in Ross and Wallmark's model.

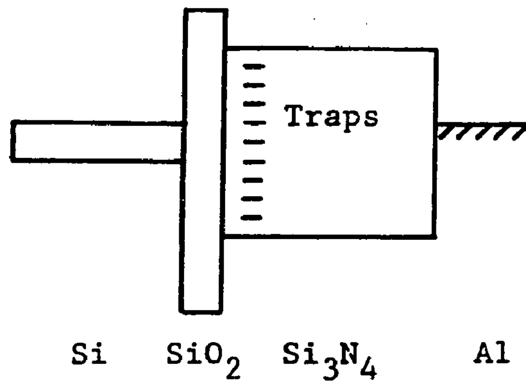


Fig. 4.6 The energetic distribution and location of traps assumed in Dorda and Pulver's model.

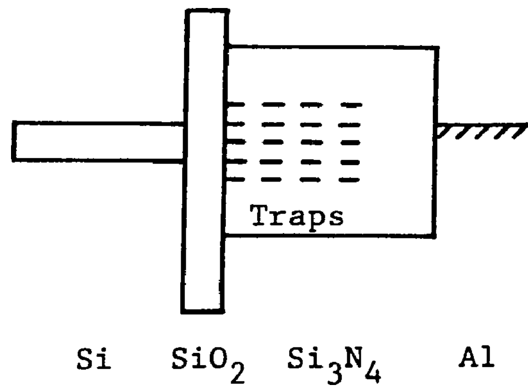


Fig. 4.7 The spatial and energetic distribution of traps assumed in Ferris-Prabhu's model.

oxide/nitride interface to the silicon conduction band and from the silicon valence band to the trap(Fig4-8). They neglected the time dependence of the nitride field and current but included the dependence of the tunneling current on the thin-oxide field strength and the initial and final densities of the states on both sides. They obtained a threshold voltage change that depends linearly on the pulse-width for small duration, then shows logarithmic dependence and finally reaches saturation.

Band to Band Tunneling

Lundstrom and Svensson [70] have proposed a model in which, with a positive gate bias voltage, electrons are assumed to tunnel from the silicon conduction band into the nitride conduction band whereas with a negative gate bias, holes tunnel from the silicon valence band to the nitride valence band [Fig.4-9]. Once the carriers are injected into the nitride bands, they are trapped by the relatively deep traps in the nitride layer. The charge propagation in the nitride is neglected with the assumption that the charge is firmly held practically at the oxide-nitride interface. They call this type of charge transport "modified Fowler-Nordheim(FN) tunneling." This mechanism is contrary to the assumptions used in all previous direct band-to-trap tunneling models [64]- [69] which assume that the electrons are injected from the silicon valence band and tunnel back to the silicon conduction band.

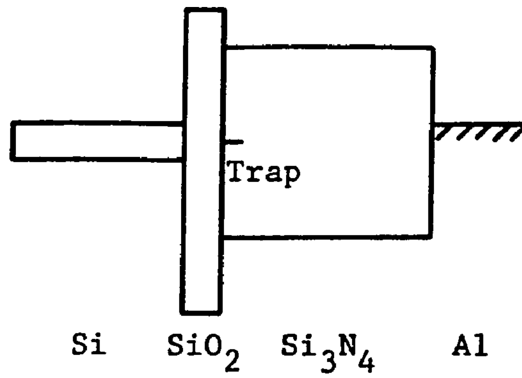


Fig. 4.8 The model considered by White and Cricchi.

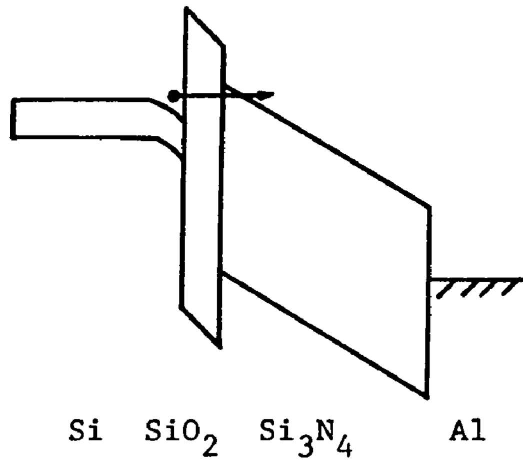


Fig. 4.9 Modified Fowler-Nordheim tunneling (band to band) proposed by Lundstrom and Svensson.

Chang [71] has also argued strongly in favor of the band-to-band tunneling. He considers a number of tunneling components as shown in Fig4-10 and formulates each tunneling component by considering an appropriate rectangular potential barrier similar to Ross and Wallmark [63]. The carrier distribution in the silicon bands is taken into account. For positive gate bias, the electrons (holes for negative gate bias) tunnel into the nitride conduction band and fall into traps locally or drift away due to the nitride field and are trapped in the nitride bulk. The resulting distribution of the trapped carriers is formulated in terms of the theory presented by Arnett [32]. The main result is that the tunneling process is initially dominated by direct band to band tunneling. As the nitride bulk is charged up, the field in the tunneling oxide decreases and tunneling by Modified FN mechanism becomes important. The characteristics reflecting logarithmic charging and eventual saturation are obtained.

Discussion

Regardless of the assumptions made concerning the traps or about the tunneling probability, all the models, discussed, predict the logarithmic behavior (ΔV_{th} vs pulse width) of charging and discharging. Extending these models to treat two carrier mechanisms is not likely to lead to different results. There is a considerable uncertainity in the values of many parameters.

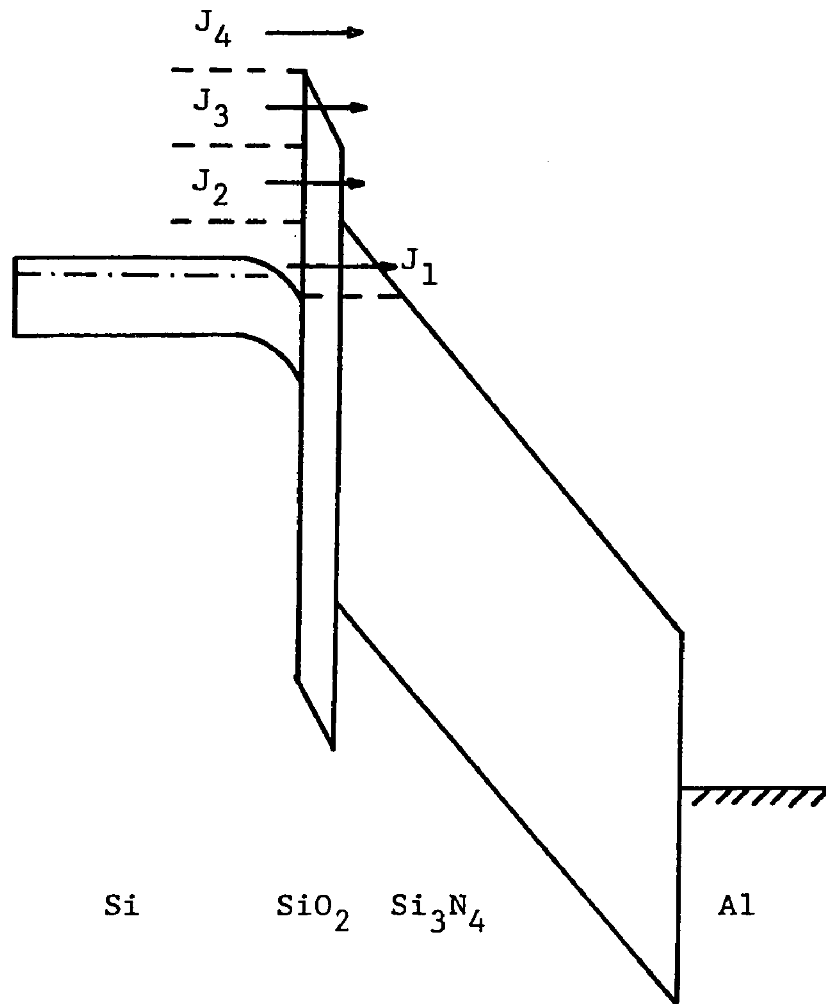


Fig. 4.10 Various tunneling current components in Chang's model.

It appears reasonable that before any refinements in the analytical models are attempted, experimental evidence should be carefully analyzed to narrow down the number of modeling possibilities. In what follows, a summary of various measurement techniques and the resulting conclusions is presented.

4.2.2 Experiments

V_{th} vs Pulse-width Measurements

Historically, this measurment has been most widely reported and used for confirmation of various models [63]- [71]. In this measurement scheme, voltage pulses of both polarities, various amplitudes and pulse-widths are applied to the device and the resulting shift in the threshold voltage V_{th}, is measured. One such example [69] is shown in Fig4-11. Such curves are generally linear for short pulse widths, logarithmic for longer pulse duration and saturated for very long pulses. Since all the models depend upon the tunneling mechanism of one carriers or the other with two or three unknown(therefore adjustable) parameters, good fit with the experimental data is usually obtained. Not much information about the physical mechanisms can be gained from such measurements.

Charge-Centroid Measurement

This measurement technique, developed by Yun [72, 73, 74], is illustrated in Fig4-12. The initial bias on the device is adjusted to the initial flatband voltage, V_{FB1}, which corresponds to, say, a

Fig. 4.11 Typical erase/write characteristics of an n-channel transistor [69].

nitride charge of Q_{N1}. At some time, a pulse of desired magnitude, polarity and pulse-width is applied to the gate relative to the bulk. After the pulse, the bias at the gate is changed to reach the new flatband value V_{FB2}, corresponding to the new charge Q_{N2}, in the nitride. The gate charge is monitored during this entire exercise. Considering the charge balance before and after the pulse, the change in the trapped nitride charge, ΔQ_N, can be determined as negative of the change in the gate charge, ΔQ_G, provided no charge injection occurs from the gate electrode. Knowledge of ΔQ_N along with ΔV_{th} then gives the position of the charge centroid of the injected charge [Eq. (6.14)].

By using this technique, Yun observed the following [73, 74]:

1. The charge centroid of a typical 500 A nitride device may be 100 A or more inside the nitride as measured from the SiO_2/Si_3N_4 interface. Thus the injected charge has penetrated 200 A or more into the nitride.
2. In the thin oxide(X_{ox}<25 A) device, as much as 80% of the injected charged tunnels back into the silicon before the conventional C-V measurement of the flatband voltage can be made.

The fact that the charge is trapped in the nitride bulk, not merely at the oxide-nitride interface, is very significant and has

recently been confirmed by Taylor and Simmons [75], Wei and Simmons [76], Maes and Van Overstraeten [77] and Lehovec [78].

Yun's method restores the sample to its initial charge-free state before the next charging pulse is applied. The memory charge is removed by applying pulses of opposite polarity until V_{FB} reaches a pre-determined value believed to be characteristic of the memory charge-free state. Reverse polarity may not only remove the injected charge (say, negative) by the preceding pulse, but also inject positive charge. Since V_{FB} depends upon the superposition of the positive and negative charge distributions, the method does not guarantee achievement of the charge-free state. To overcome this problem, Lehovec et al [78] have developed the technique of staircase charging. In this technique, each successive pulse is superimposed on the flatband voltage corresponding to the accumulated memory charge distribution resulting from the preceding pulses.

Although, the charge-centroid technique can determine the net charge injected into the insulator, it cannot determine the sign of the charges crossing the tunneling oxide-nitride interface. In addition, significant back-tunneling occurs before the new state of the device can be determined.

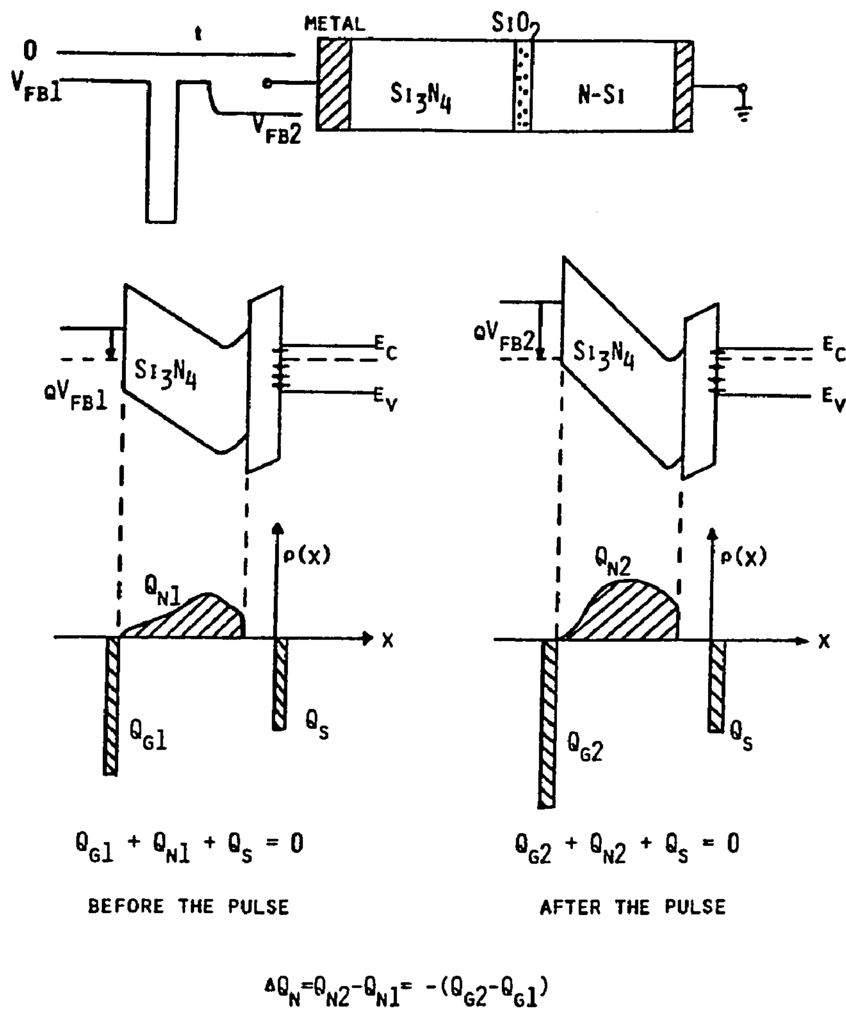


Fig. 4.12 Flatband-tracking technique of Yun[72].

Retention and Endurance Measurements

Retention deals with the ability of Si_3N_4 to retain charge when the Writing(or Erasing) pulse returns to the bulk potential. Usually, it is measured as the change in the threshold voltage as a function of time after the voltage pulse returns to the bulk potential [69, 79, 80, 81]. Neugebauer and Burgess [80] have identified short term decay and long term decay rates in the retention curves[Fig4-13]. The breakdown point from the short term decay to long term decay is function of the device cycling. The mechanism for short term decay is thought to be back-tunneling assisted by interface traps(see Fig4-2b) [79, 80]. The long term decay is postulated to be due to the rearrangment of charge within the nitride. The space charge close to the SiO_2-Si_3N_4 interface tends to move closer to it whereas the charge further away, moves closer to the gate electrode. Since the conductivity of the nitride increases with cycling [79, 80], this explains the observation that the breakdown point moves to shorter times as the device is cycled.

Endurance is the ability to maintain threshold window with Write/Erase cycling [69, 79, 80, 81]. It is usually measured in terms of the threshold voltage for Write and Erase operation vs number of Write/Erase cycles[Fig4-14]. The degradation in the threshold window is thought to be related to the change in the nitride conductivity. The data shown in Fig.4-14 shows degradation

in the memory window after 10^4 cycles. This degradation primarily arises due to the very high fields(in excess of $7Vcm^{-1}$) that have been used to inject charges into the nitride. However, White et al [81] observe no degradation upto 10^9 cycles by keeping the electrostatic field in the nitride below 5 $MVcm^{-1}$ as shown in Fig.4-15.

Other Techniques

Among many techniques, the Thermal Dielectric Relaxation Technique(TDRC) used by Wei and Simmons [76] seems to be useful in characterizing various trapping parameters in the nitride layer. This technique, when combined with charge centroid measurement, has the potential for determining bulk traps in Si_3N_4 and their role in the operation of memory devices.

Recently, Ginovker et al [79], Weinberg et al [37, 38, 39] and Schroder and White [40] have used the charge separation technique to investigate the steady-state charge transport through the MNOS structure. The main conclusion from this kind of study is that the nitride current is carried almost exclusively by holes in steady-state [41]. However, to obtain any steady-state current in the good quality (relatively trap-free) nitride layers, the field has to exceed 6 MV/cm which is quite close to the breakdown field of the nitride. In any case, the conclusions derived in steady-state under high field conditions are not necessarily valid for the actual

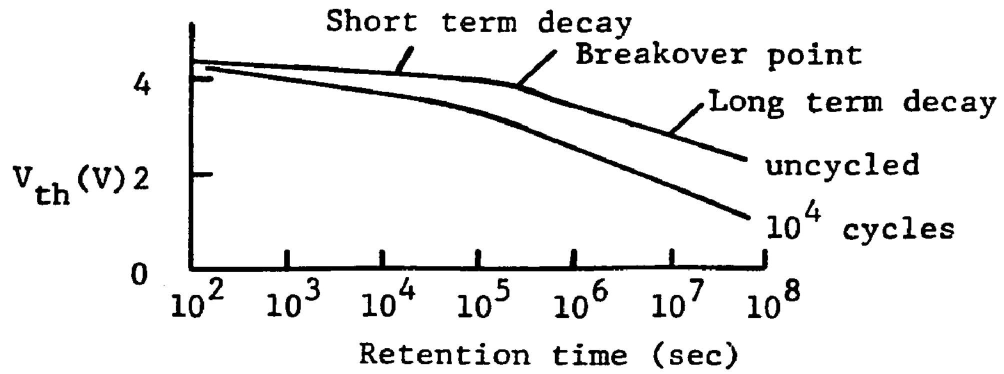


Fig. 4.13 Typical retention curves of MNOS memory transistor as a function of cycling.

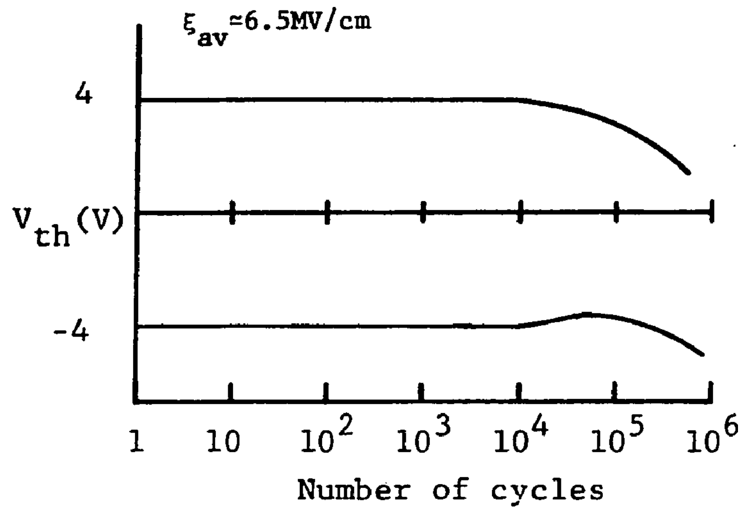


Fig. 4.14 Typical endurance curves of the MNOS memory transistor.

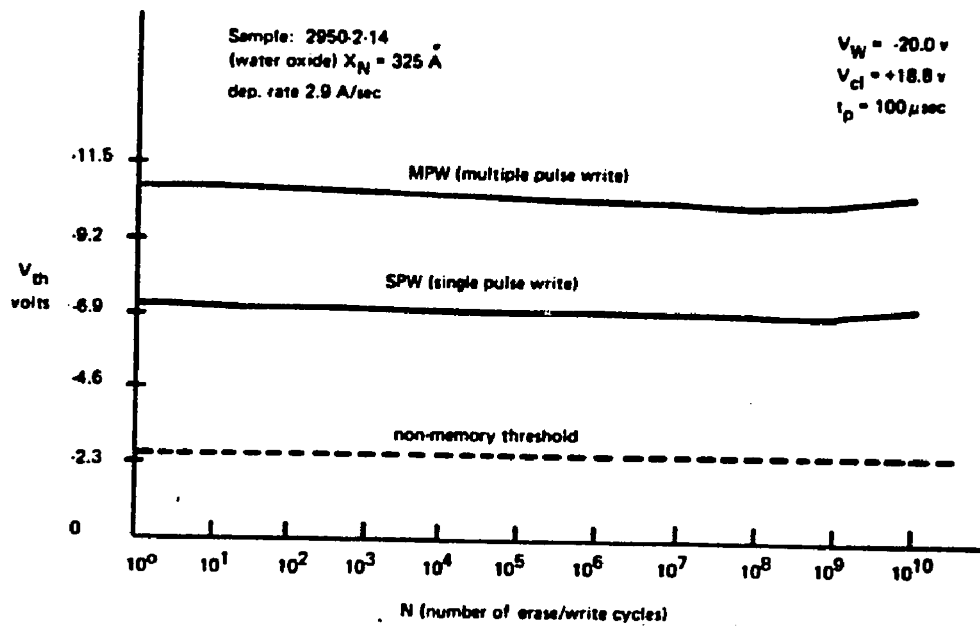


Fig.4.15 Endurance data of White et al [81].

charging of the device in memory application.

4.3 Purpose of This Work

As can be seen from the preceding paragraphs, considerable progress has been made in recent years toward understanding and theoretically modeling the operation of the MNOS device. However, despite the voluminous literature on MNOS, the understanding is far from being thorough. The biggest obstacle is the lack of knowledge about the various tunneling mechanisms indicated in Fig4-2.

The purpose of this work is to identify the dominant tunneling mechanism for the Write and Erase operation of the MNOS/MONOS devices and to study the trapping process in the nitride layer. Specifically, the charge separation technique of Ginovker et al [36] is extended and modified to apply in the non-steady-state mode of operation. The modified technique presented in the next chapter not only achieves the charge separation in the silicon but in addition can also monitor the net injected or collected charge vs the flatband voltage shift. It combines the charge centroid technique of Yun [72] with the charge separation technique of Ginovker et al [36] and thus presents a comprehensive picture of the device operation. This technique also minimizes the back tunneling of the injected charge before the final flatband voltage of the device can be determined.

CHAPTER V

A NEW EXPERIMENTAL TECHNIQUE

5.1 Linear Voltage Ramp Technique

Traditionally, the slow linear voltage ramp technique has been used to study the properties of the Si-SiO_2 interface in the MOS system [82]. The technique involves measuring the displacement gate current vs gate voltage respose of a MOS capacitor subjected to a slowly varying linear voltage ramp (ramp rate α=10mv/sec) so that the device remains in quasi-equilibrium. The resulting quasi-static characteristic is shown by curve 1 in Fig.5-1.

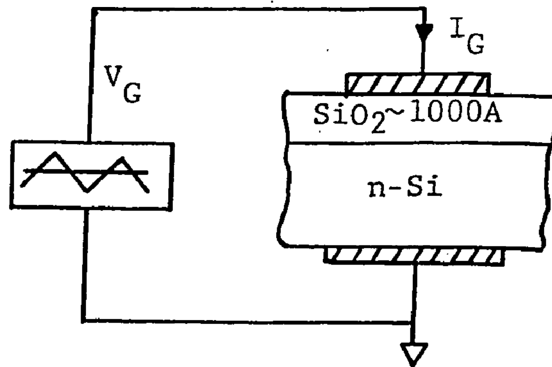
Recently, relatively fast ramps(α=0.1 to 1V/sec) have been employed to study the properties of the bulk semiconductor [83, 84, 85, 86, 87]. If the ramp rate is sufficiently high, the build up of the inversion charge cannot follow the bias and consequently, the semiconductor is partially deep-depleted. This results in a nonequilibrium situation in which the depletion width exceeds the equilibrium value. The electron-hole pair generation processes via bulk traps in the space-charge region and the interface states produce a generation current which is augmented by the depletion current due to the expanding depletion region. The displacement current at the gate, then, reflects these two

components of the current as shown by curve 2 in Fig5-1. Furthermore, if the device is cooled down to a low temperature , the pair generation via bulk and the interface traps decreases exponentially to a negligible level. Since, there is no other source of minority carriers in the MOS capacitor, the semiconductor is deep-depleted as shown by curve 3 in Fig5-1. A large fraction of the bias then appears across the semiconductor and only a small fraction is applied across the insulator.

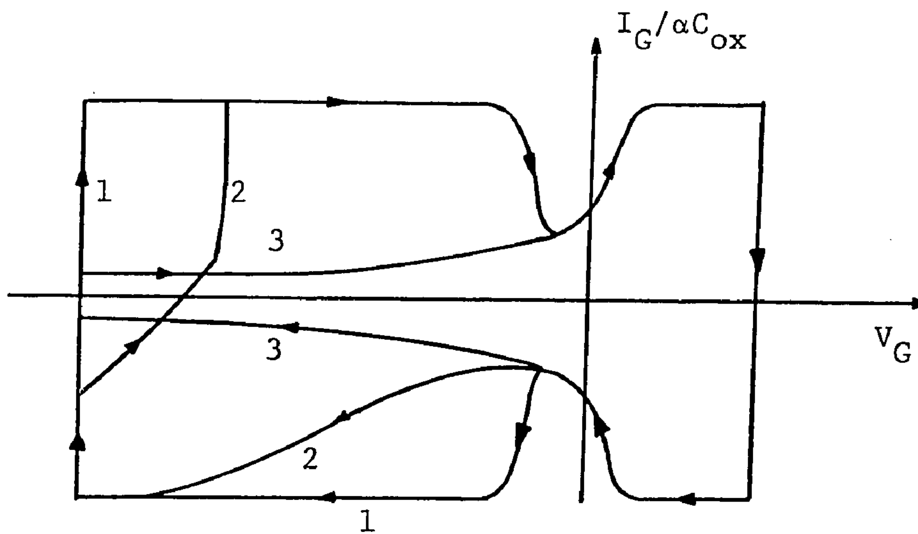
5.2 Charge Separation in the Semiconductor

Consider the case when the MOS capacitor is replaced with a gated-diode or a p-channel MOSFET with Drain and source electrode shorted together(Fig5-2). Source to bulk junction is reverse biased with a battery $V_{SB}(<0)$ and the gate electrode bias V_G is ramped with a function generator(α=0.1 to 1 V/sec) with the bulk terminal grounded. The device is cooled to about 100 K to eliminate thermal pair generation in the semiconductor if $V_{SB}<0$. The bulk current I_B and the source current I_S are measured as functions of V_G. The gate current I_G can be determined by $I_G=I_B+I_S$.

Figure5-3 shows the calculated current components for V_{SB}=0 and -1V and small values of V_G to reduce tunneling across the insulator. As the device is ramped from accumulation to depletion, only majority carriers(electrons) contribute to the current. Thus, I_S=0 and $I_G=I_B$. Since the p^+/n junction is reverse-biased by V_{SB}, the



(a)



(b)

Fig. 5.1 (a) Schematic of the ramp measurement, (b) Various $I_G/\alpha C_{ox}$ vs V_G curves with the semiconductor in quasi-equilibrium (curve 1), partial deep-depletion (curve 2) and deep-depletion (curve 3).

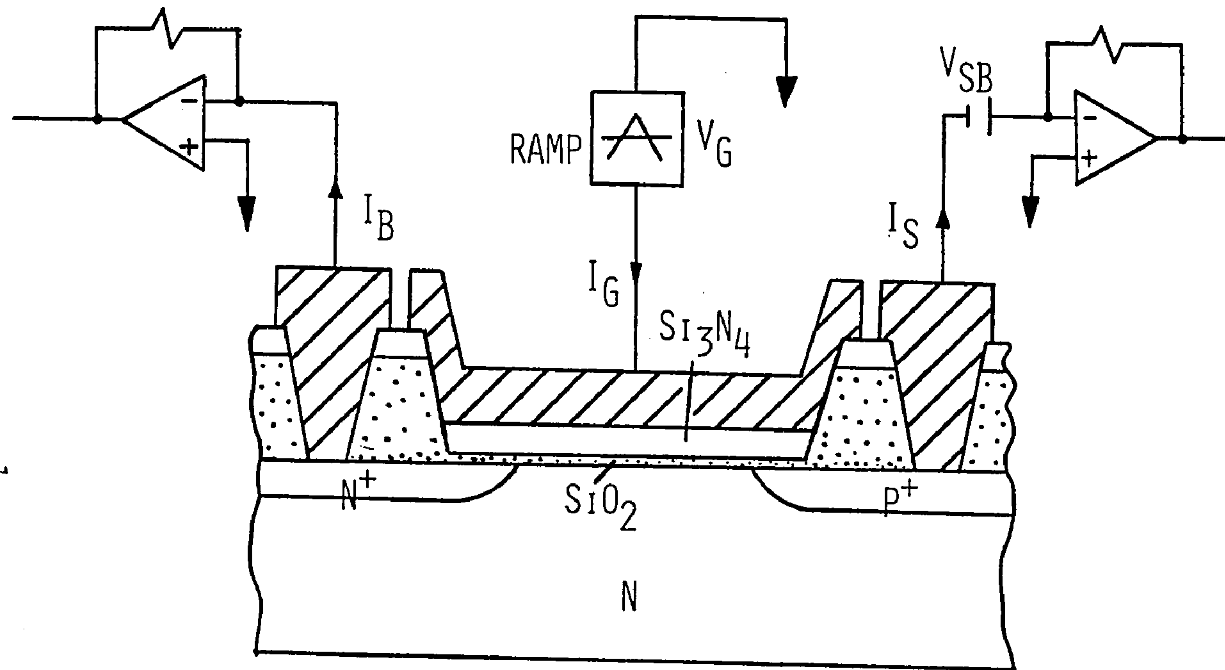


Fig. 5.2 The schematic of the device structure and the linear voltage ramp technique with charge separation.

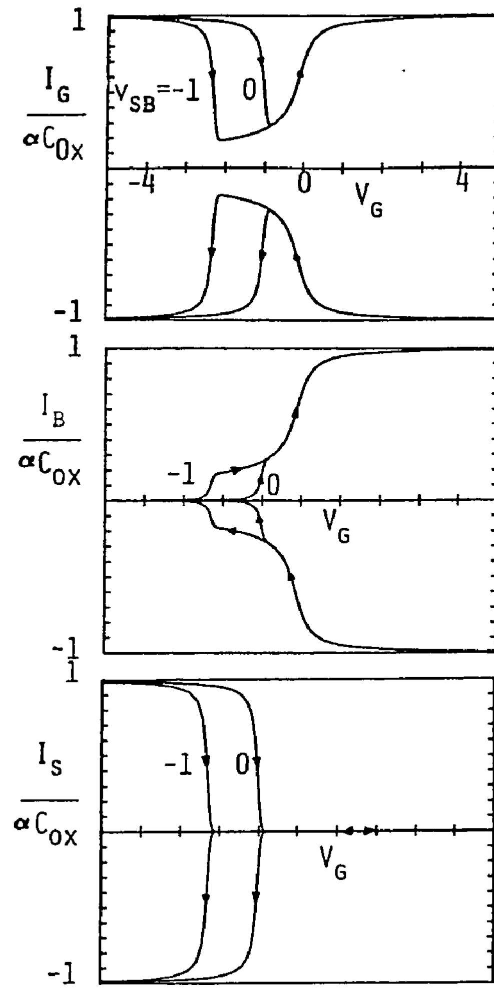


Fig. 5.3 The calculated current components in the ramp technique for the non-injection case.

hole Fermi-Level E_{Fp} is raised above the *electron* Fermi-Level E_{Fn} by qV_{SB} in the space-charge region as shown in Fig5-4(a). As a consequence of this separation, the device moves into deep-depletion. When the surface potential $|\psi_s|$ exceeds $|V_{SB}+\phi_B|$, E_{Fp} scans the lower half of the bandgap at the Si-SiO_2 interface (Fig5-4b). As a result, the p^+ region injects minority carriers (holes) into the space-charge region and I_S rises to $-\alpha C_{ox}$ with I_B decaying to zero. When $|\psi_s|=|V_{SB}+2\phi_B|$, the device is in strong inversion. The depletion region ceases to expand(I_B) and all the current is supplied by the p^+ source ($I_G=I_S$). During the reverse sweep, we obtain mirror images of the curves.

A set of experimental curves is shown in Fig5-5 for a p-channel device with $V_{SB}=0$. The various current components are similar to the ones shown in Fig5-3 with two exceptions: (a) The non-zero I_B in inversion region is due to the MOS capacitor underneath the gate bonding pad, C_{pad}, which does not invert due to the thick field oxide (see Fig5-6), (b) The non-zero I_S in the accumulation region is due to the gate to junction overlap capacitance, $C_{overlap}$, as shown in Fig5-6. Both of these components may be made negligibly small by suitably designing the structure (by reducing the bonding pad area, increasing the field oxide thickness and using self-aligned gate technology). In this study, these components are simply substracted from the acquired data. It is interesting to note that this technique provides a very unique method of measuring

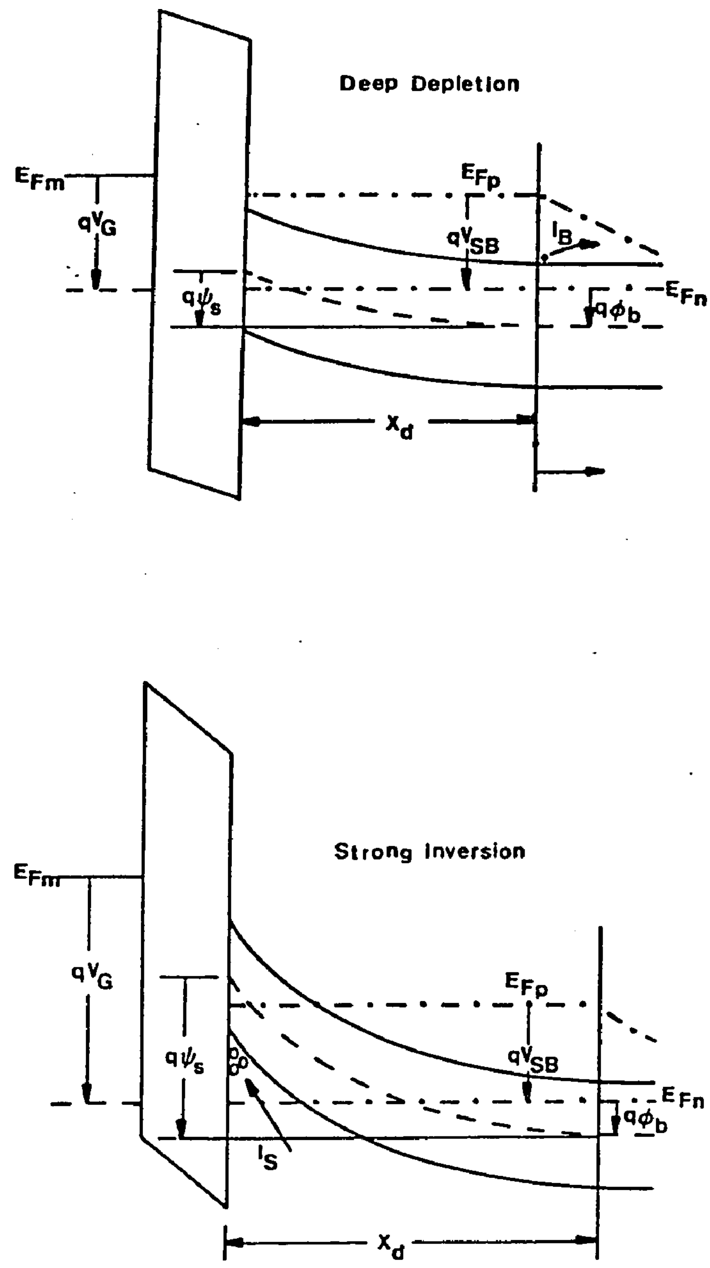


Fig. 5.4 The energy band diagrams in deep-depletion and strong inversion conditions in silicon.

the overlap capacitance in small geometry structures where overlap capacitance may be as high as the gate capacitance, itself.

5.3 Application to the Multi-Dielectric Structure

The linear voltage ramp technique can be applied to multi-dielectric structures provided there is no significant steady-state leakage between the gate and bulk electrodes. This condition is easily satisfied in practical MNOS and MONOS devices. Consider the p-channel MONOS device that was used to obtain the experimental curves shown in Fig5-5. If a large positive bias (+20) is applied on the gate for a large period (say, 1 min), the electrons will tunnel from the semiconductor into the nitride conduction band and will be captured by the traps in the nitride film (see Fig5-7a). In addition, some holes may tunnel from the traps in nitride to the semiconductor. The overall result would be a net negative charge Q_{N1} in the nitride film correspoding to a positive flatband voltage V_1 (+8V). Next, the gate bias is swept from weak accumulation (+12 V) to a large negative bias (-20V) into strong inversion and back to the starting value (+12 V). The various current components are shown in Fig5-8.

As the device is swept from V_1, which corresponds to the initial flatband voltage, to strong inversion at large values of the negative bias, two processes may take place (Fig5-7b) : (1) holes that are supplied by the diffused regions via current I_S tunnel

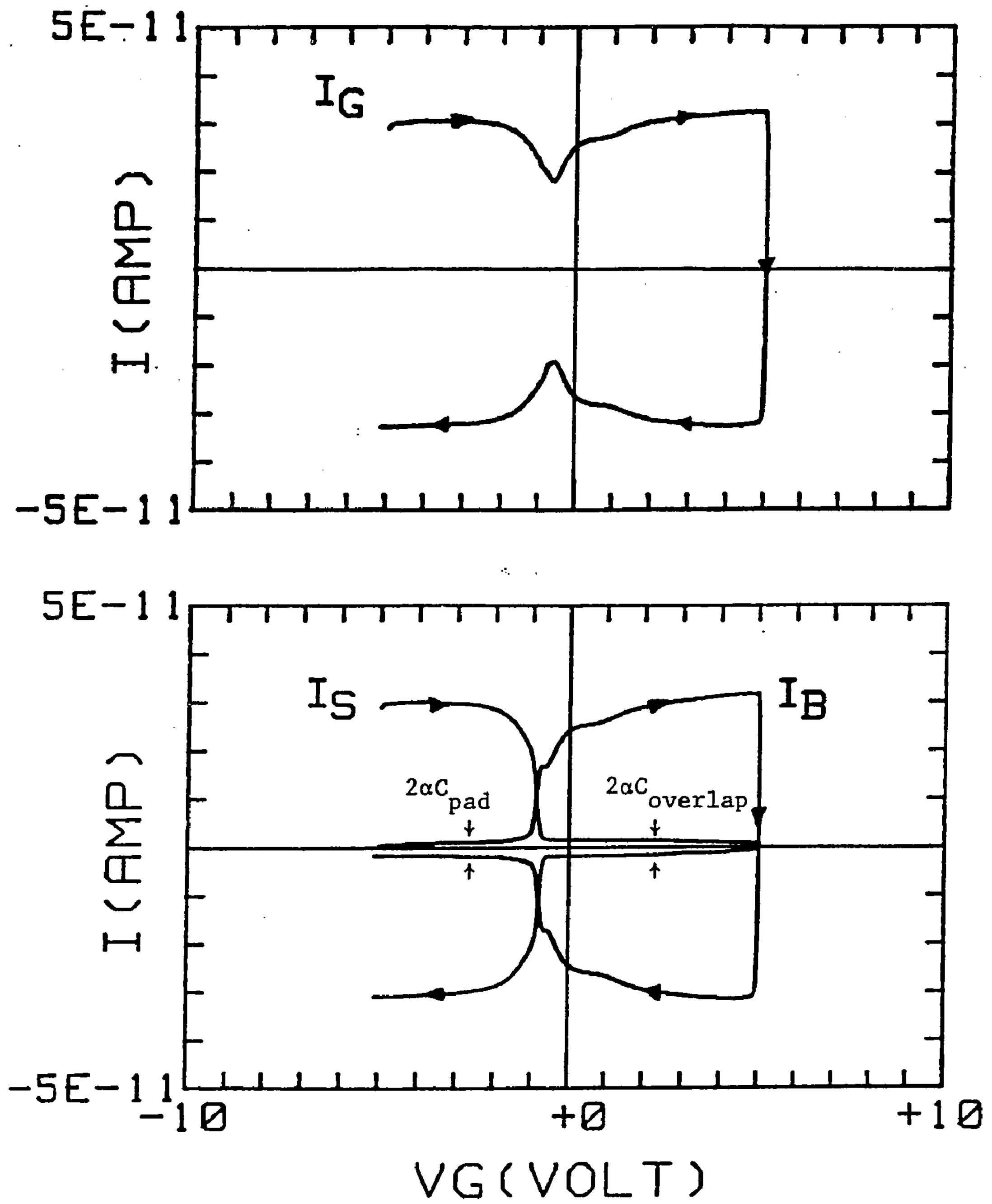


Fig. 5.5 Experimental curves for a p-channel MONOS structure ramped at 0.2V/s.

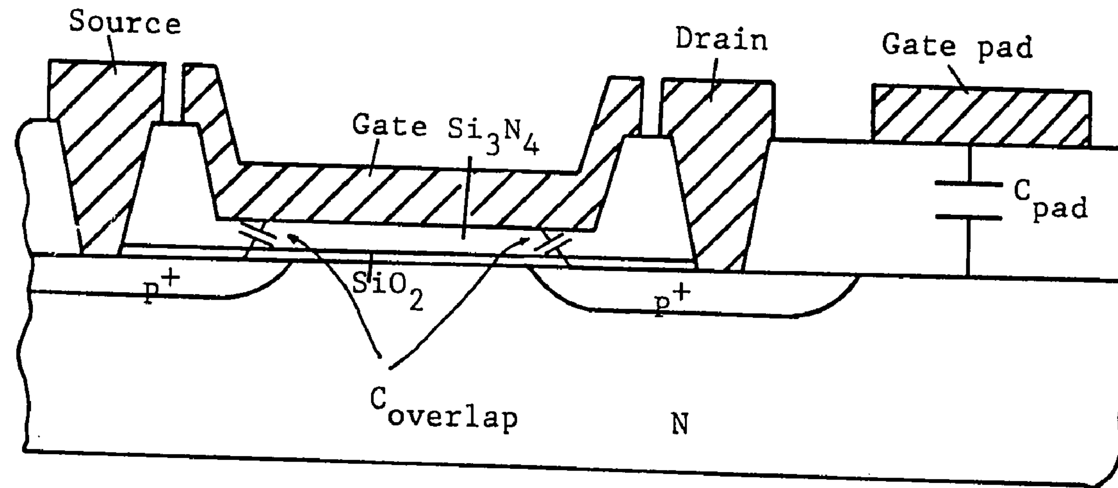


Fig. 5.6 Schematic of the MOSFET (p-channel) structure indicating parasitic capacitances.

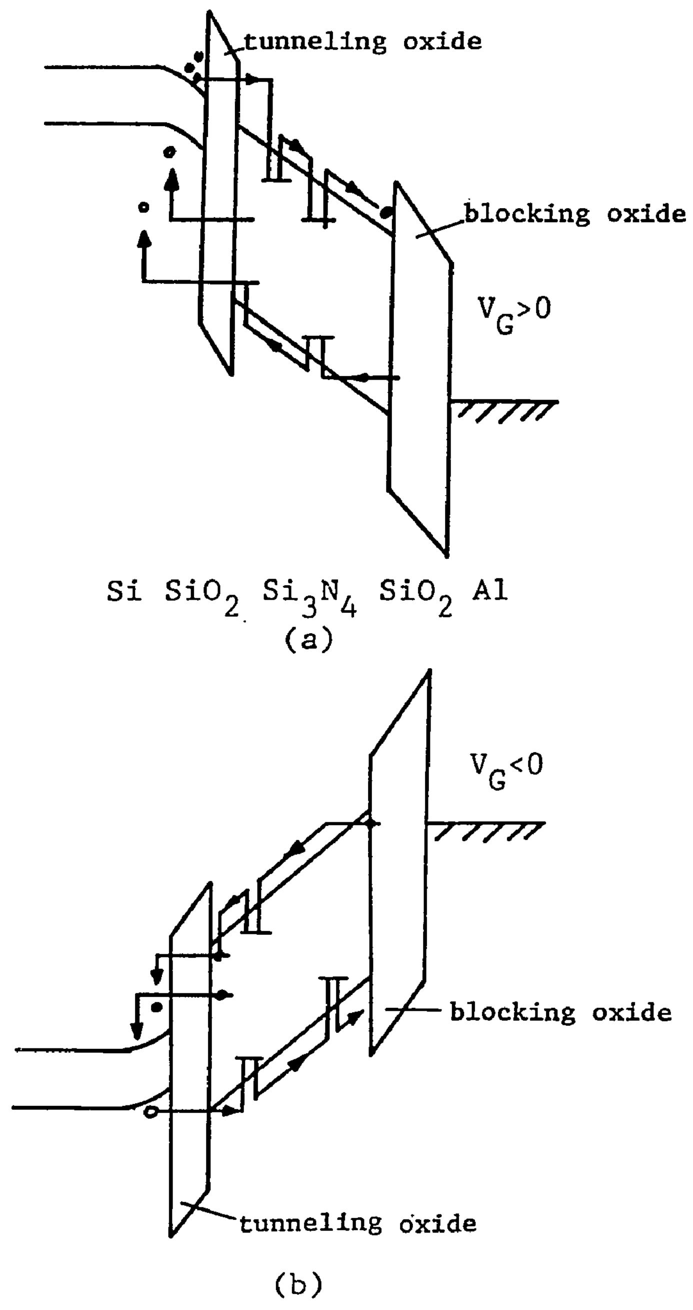


Fig. 5.7 Various tunneling current components in the MONOS sandwich under: (a) positive gate bias (relative to bulk), and (b) negative gate bias.

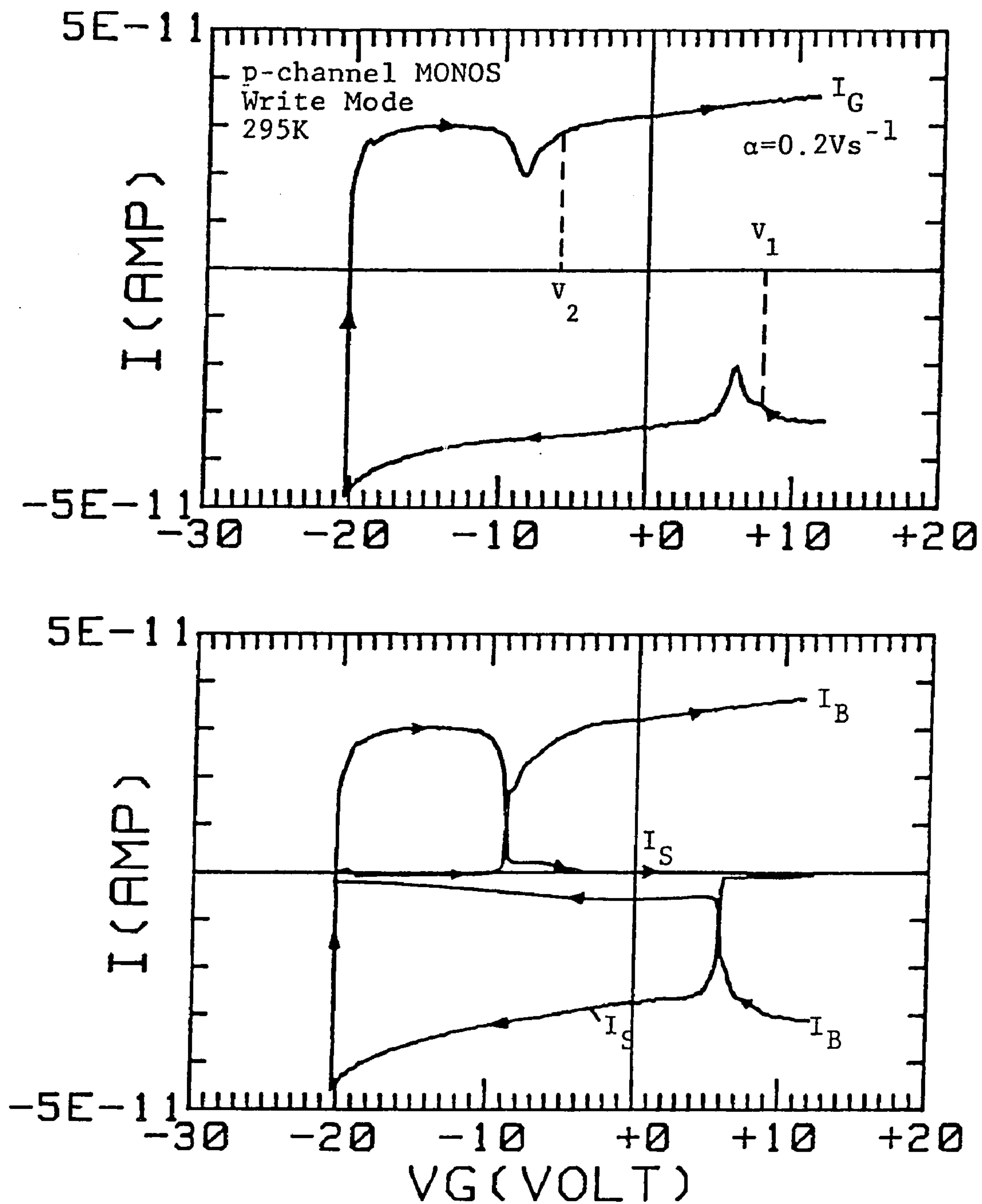


Fig. 5.8 Various current components in the Write mode of the p-channel MONOS structure. The device dimensions are: X_{oxt}=10A(tunnel oxide thickness);X_N=530A(nitride thickness), X_{oxB}=75A(blocking oxide thickness)and device-area=1.51x10^{-3}cm^2.

through the thin oxide and are trapped in the nitride film, resulting in an increase in I_S over and above $- \alpha C_{eff}$, (2) electrons that are previously trapped in the nitride film tunnel directly into the silicon conduction band, drift through the depleted region and are collected by the bulk electrode as I_B. This results in an excess component in I_B in the inversion region as can be seen by comparing I_B in Figs5-8 and 5-5. During the reverse sweep, holes are removed from the inversion layer via I_S which is positive and flows out of the source electrode. Finally, the bias is swept past the new flatband voltage, V_2, which is now shifted to a negative value (-7V).

Let us look at the sweep more closely. Fig.5-9 shows the bulk and source current components along with the energy band diagrams at various points on the sweep. The sweep begins in accumulation with the net nitride charge being negative. It can be seen from the energy band diagram in accumulation (Fig5-9) that the field in the thin tunneling is not favorable for the injected electrons to tunnel back to the silicon. This field becomes favorable when the device is swept into depletion and then inversion, past the flatband voltage. Electrons that are trapped close to the tunneling-oxide/nitride interface and are within a tunneling distance, tunnel into the semiconductor conduction band are and are collected as the bulk current I_B. In addition, as shown in Fig5-9 holes can also tunnel from semiconductor. Initially, at small fields in the

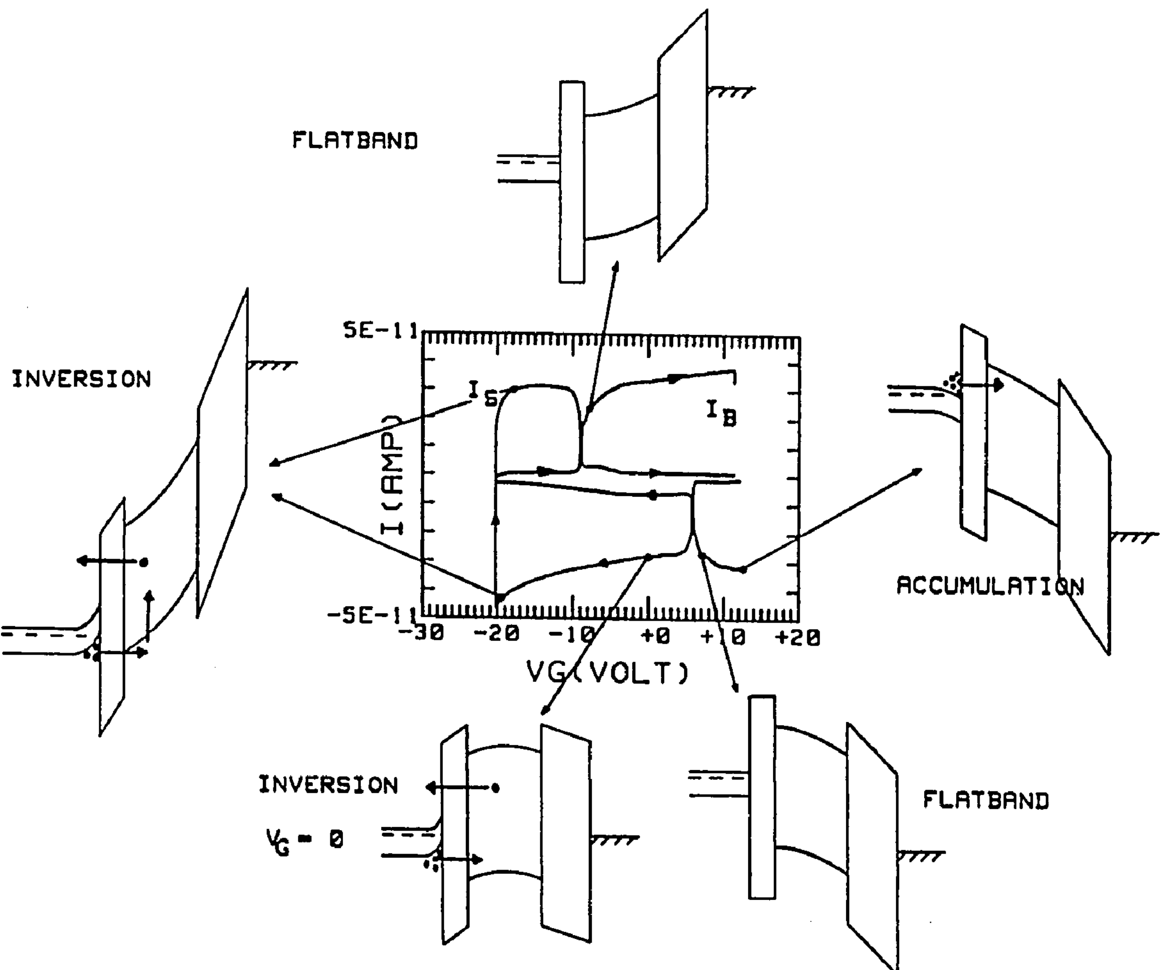


Fig. 5.9 The energy band diagrams at various points in the voltage sweep on a p-channel device.

tunneling oxide, holes tunnel directly to the traps. At higher field, band to band tunneling of holes becomes more probable as shown in the energy band diagram at -20 V in Fig5-9. During the reverse sweep, field in the tunneling oxide is not favorable for the injected holes to back-tunnel into the semiconductor before the new flatband voltage is reached. Thus, this technique monitors the tunneling of electrons from the nitride to the semiconductor and also monitors the hole tunneling from the semiconductor to the nitride. In addition, the back-tunneling of the newly injected holes is prevented during the reverse sweep due to the unfavorable field in the tunneling oxide before the new flatband voltage may be measured.

The bulk and source current components of Fig5-9 are redrawn in Fig5-10. We only consider the sweep between the initial and final voltages V_1 and V_2, respectively. Since the gate bias is proportional to time, the area under the I-V curve is proportional to the charge. The semiconductor charge is the same at points V_1 and V_2 and thus, we conclude that in the absence of any injection from the gate electrode, the hatched areas $(A_1-A_2)/\alpha A$ under I_B and I_S curves represent ΔQ_N (coul cm^{-2}) the net electron charge released from the insulator into Si and ΔQ_p(coul.cm^{-2}), the net hole charge trapped into the insulator, respectively. A represents the device area. In addition, $\Delta V_{FB}=V_2-V_1$ can be read directly from the I_B vs V_G curve.

The reproducibility of the measurement is illustrated in Fig5-11. The device is prepared by applying +20 V for 1 min and then swept from +12 V to different values of the negative voltage. After each sweep, the device is prepared again in the identical manner. The higher the value of the negative bias, the higher is the shift in flatband voltage.

Similar experiments can be performed with the opposite gate polarity where an initial large negative bias is applied and the device is swept from a small negative bias to a large positive bias and back. The flatband voltage shifts to more positive value and ΔQ_n represents the net electron charge which tunnels from the semiconductor to the silicon nitride and is trapped (Fig.5-7a) while ΔQ_p is the hole charge released from the traps in the nitride and collected by the source electrode as I_S. This is the 'ERASE' mode of operation for a p-channel device. There is a serious problem with charge separation in this mode of operation. Holes that are released from the traps in nitride, tunnel into the valence band of the semiconductor (Fig.5-7a). However, in order to be collected by the source region which may be 50-100 μm away, holes have to drift in the space-charge region over such long distances. There is a very high probability of recombination with electrons that are present in high numbers in the accumulation layer. Therefore, this technique would not detect any significant hole charge in the Erase mode of p-channel devices. This situation should be contrasted with

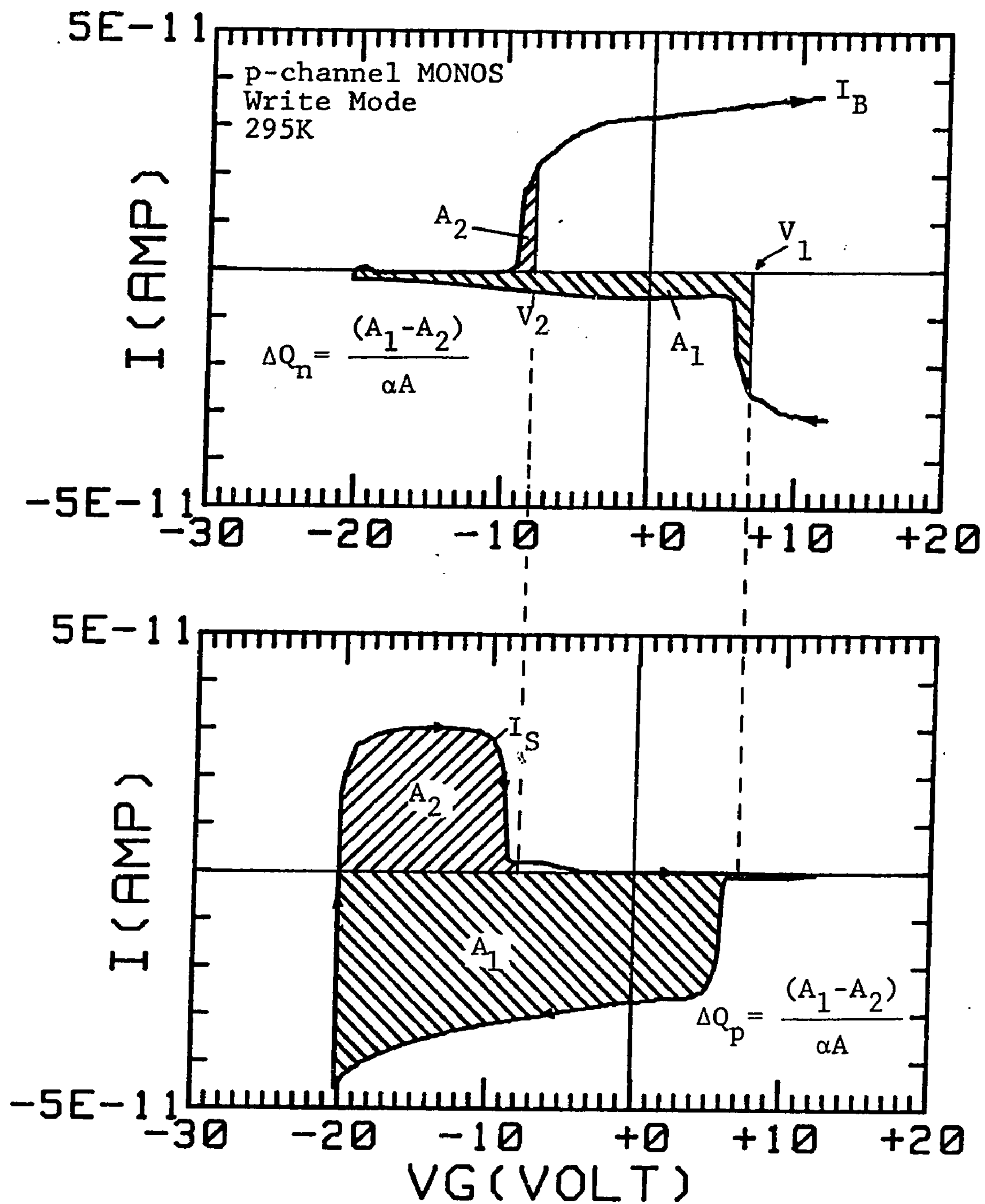


Fig. 5.10 Integration of the I-V curves to obtain the values of the injected and released (from nitride into silicon) charges in the ramp technique. The device dimensions are: $X_{oxt} \approx 10A$; $X_N = 530A$; $X_{oxB} = 75A$; $A = 1.51 \times 10^{-3}$ cm^2.

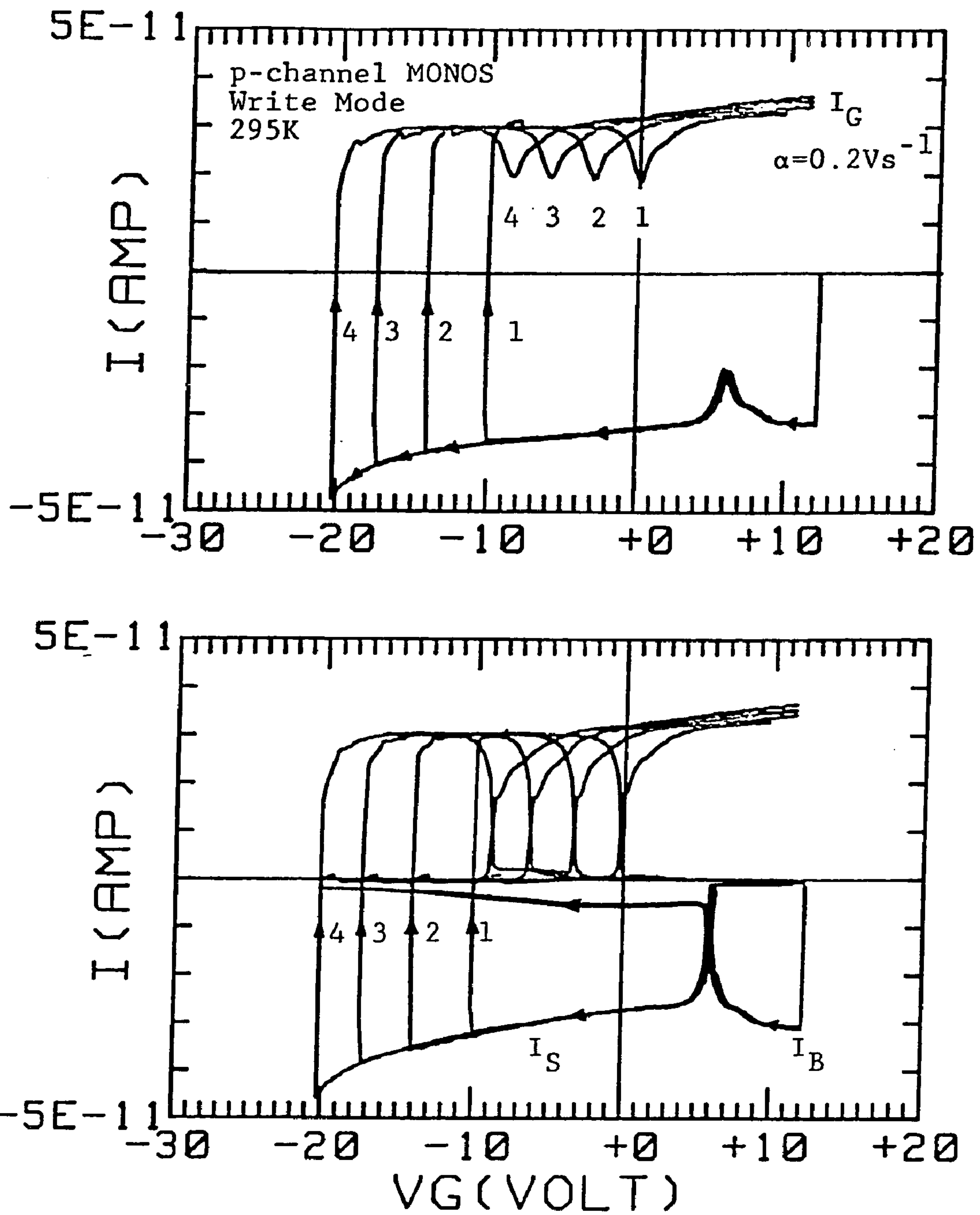


Fig.5.11 . Various current components with different magnitudes of the Writing voltage applied to the p-channel MONOS device.

the Write operation on the p-channel devices where holes, supplied by the source region, drift in the highly conducting inversion layer while the electrons, that are released from the nitride layer, tunnel into the conduction band of silicon and drift towards the bulk contact through the space-charge region which is only of the order of 1μm.

The technique would likewise work for the Write operation of the n-channel devices(V_G>0) and not work for the Erase operation. Since the Erase mode for n-channel devices is the same as Write mode of p-channel, we should be able to get the complete picture by using both types of structures.

CHAPTER VI

THEORY OF LINEAR VOLTAGE RAMP TECHNIQUE APPLIED TO THE MONOS STRUCTURE

In this chapter, the nature of the ramp I-V curves is analysed in terms of charge tunneling across the thin oxide and trapping in the nitride. The displacement current at the gate is a sensitive indicator of what happens within the gate insulator. With a careful interpretation, certain mechanisms within the insulator may be ruled out. To this end, we begin with the electrostatics of the MONOS structure.

6.1 Electrostatics of the MONOS Structure

Consider a MONOS sandwich with n-type semiconductor in accumulation. The energy band diagram, the charge and the electrostatic field distributions are shown in Fig.6-1. Neglecting the charge in the interface traps, we can write

$$Q_G + Q_N + Q_S = 0 \tag{6.1}$$

and $$V_G = V_{oxB} + V_N + V_{oxt} + \psi_S + \phi_{ms}, \tag{6.2}$$

where, $Q_G(>0)$, $Q_N(<0)$ and $Q_S(<0)$ denote the charge per unit area at the gate electrode, in the silicon nitride and the silicon, respectively. The gate voltage $V_G(>0)$ is the sum of the potential drops across the blocking oxide $V_{oxB}(>0)$, nitride $V_N(>0)$, tunneling

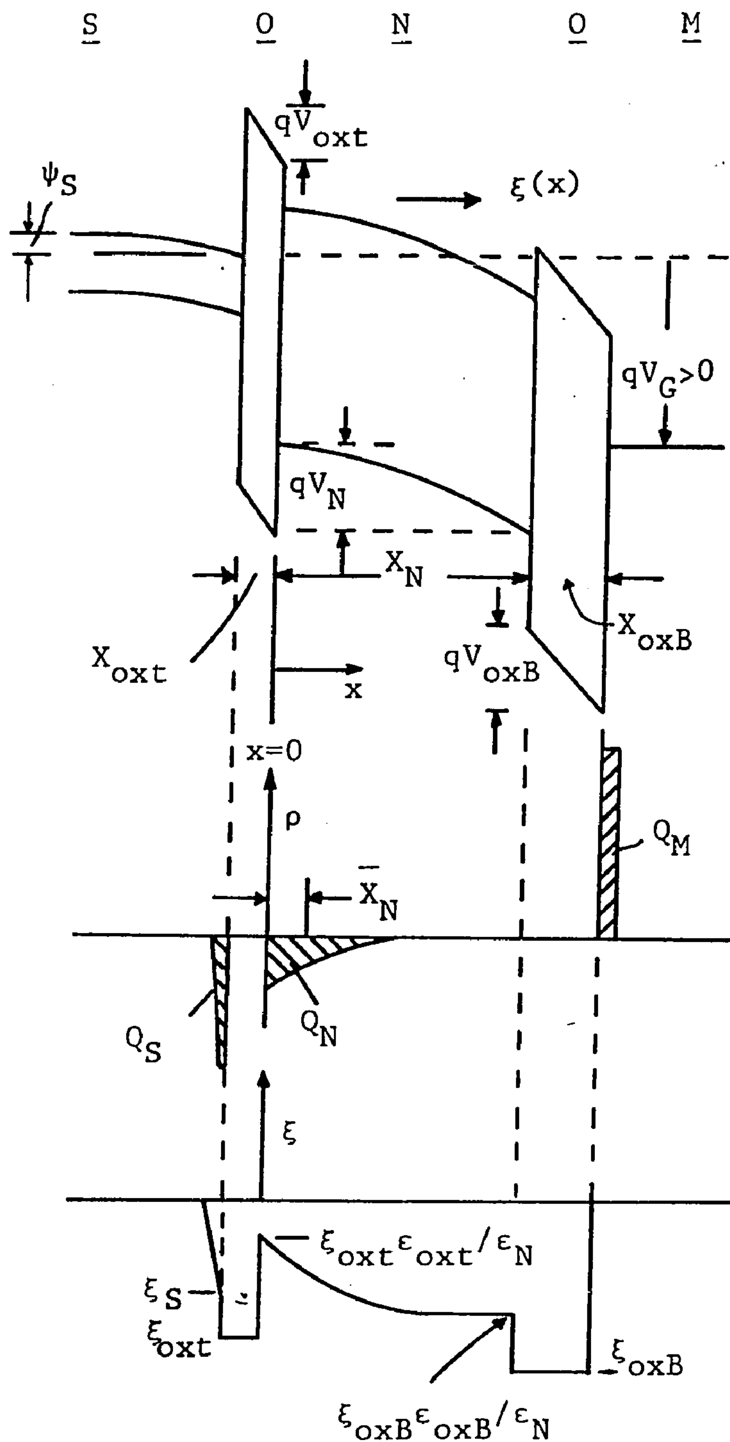


Fig. 6.1 The energy band diagram, charge density ρ (x) and the electrostatic field distribution ξ(x) in the MONOS structure biased positively at the gate.

oxide $V_{oxt}(>0)$ and the surface potential $\psi_s(>0)$. The charge in the nitride Q_N, can be written as

$$Q_N = \int_0^{X_N} \rho_N(x)\, dx \tag{6.3}$$

where, $\rho_N(<0)$ is the charge density [Ccm^{-3}] in the nitride. The semiconductor charge Q_S is related to the surface potential ψ_s for the n-type semiconductor [22]:

$$\begin{aligned} Q_s &= \frac{-\sqrt{2}\varepsilon_s}{\beta L_D} [(e^{\beta\psi_s} - \beta\psi_s - 1) + \frac{n_i^2}{N_D^2}(e^{-\beta\psi_s} + \beta\psi_s - 1)]^{1/2} \\ &\simeq \frac{\sqrt{2}\varepsilon_s}{\beta L_D} e^{\beta\psi_s/2}, \text{ (accumulation)} \end{aligned} \tag{6.4}$$

where, $L_D = \sqrt{[2\varepsilon_S/qN_D\beta]}$ is the extrinsic Debye length, and $\beta = q/kT$.

Assuming no charge in the oxide, the potential drops V_{oxB} and V_{oxt} are given by

$$V_{oxB} = -\xi_{oxB}X_{oxB} = Q_G X_{oxB}/\varepsilon_{oxB}, \tag{6.5}$$

$$V_{oxt} = -\xi_{oxt}X_{oxt} = -Q_S X_{oxt}/\varepsilon_{oxt}, \tag{6.6}$$

where, ξ_{oxB} and ξ_{oxt} are the electrostatic fields [$V \cdot cm^{-1}$] in the blocking and tunneling oxides, respectively. The blocking oxide is characterized by its thickness X_{oxB} and permittivity ε_{oxB} and the tunneling oxide by X_{oxt} and ε_{oxt}. The electrostatic field in the nitride is obtained by integrating Poisson's equation:

$$d\xi_N(x)/dx = \rho_N(x)/\varepsilon_N \; , \qquad (6.7)$$

which gives

$$\xi_N(x) = \int_0^x [\rho_N(x)/\ \varepsilon_N]dx + \xi_N(0) \; .$$

From Gauss's law, $\xi_N(0) = Q_S/\varepsilon_N$.(neglecting interface state charge).

Thus,

$$\xi_N(x) = [Q_S + \int_0^x \rho_N(x)dx]/\varepsilon_N \; . \qquad (6.8)$$

Using $Q_S = -(Q_G + Q_N)$ and integrating $\xi_N(x)$ over the nitride layer gives V_N:

$$V_N = -\int_0^{X_N} \xi_N(x)dx = [Q_G X_N + \int_0^{X_N} x\rho_N(x)dx]/\varepsilon_N \; . \qquad (6.9)$$

Equations (6.3)-(6.6) and (6.9) describe the various terms in eqs. (6.1) and (6.2). Equation (6.2) may be re-written in terms of the flatband voltage, V_{FB} as

$$V_G = V_{FB} - Q_S/C_{eff} + \psi_S \; , \qquad (6.10)$$

with $V_{FB} = -Q_N[X_{oxB}/\varepsilon_{oxB} + X_N/\varepsilon_N] + [\int_0^{X_N} x\,\rho_N(x)\,dx]/\varepsilon_N + \phi_{ms}$ (6.11)

and $1/C_{eff} = [X_{oxB}/\varepsilon_{oxB} + X_N/\varepsilon_N + X_{oxt}/\varepsilon_{oxt}]$. (6.12)

Note that the flatband voltage involves not only the total

nitride charge but also its distribution within the nitride layer. The charge distribution $\rho_N(x)$ is often characterized by its charge-centroid $\overline{X}_N$ measured from the tunneling oxide-nitride interface:

$$\overline{X}_N = [\int_0^{X_N} x\rho_N(x)dx]/Q_N \quad . \tag{6.13}$$

Flatband voltage V_{FB}, from eq.(6.11), may now be written in terms of the charge centroid $\overline{X}_N$ as

$$V_{FB} = -Q_N[x_{oxB}/\varepsilon_{oxB} + (X_N-\overline{X}_N)/\varepsilon_N] + \phi_{ms} \quad . \tag{6.14}$$

6.2 Dynamic Analysis of the Ramp Measurements

Consider the system described in the previous section subjected to a linear voltage ramp (ramp rate $\sim 0.2\ Vs^{-1}$) at the gate such that the gate bias relative to the bulk is positive and increases with time. This ramp will drive the n-type semiconductor towards an increasing degree of accumulation. Now, electrons will tunnel from the accumulation layer across the thin tunneling oxide into the nitride conduction band and are trapped within the nitride layer. This results in a change in the flatband voltage of the device with time. This dynamic situation is considered in this section and some useful relationships are derived.

The linear voltage ramp is described in terms of the ramp rate $\alpha[Vs^{-1}]$:

$$V_G(t) = V_0 + \alpha t \ . \tag{6.15}$$

Using eq.(6.15) in (6.10) and differentiating with respect to time,

$$\alpha = \dot{V}_{FB}(t) + \dot{\psi}_S(t) - \dot{Q}_S(t)/C_{eff} \ . \tag{6.16}$$

Using eq.(6.4) to obtain $\dot{Q}_S$ in terms of $\dot{\psi}_S$ and substituting in the above equation,

$$\dot{\psi}_s(t) = \frac{\alpha - \dot{V}_{FB}(t)}{1 + \dfrac{\varepsilon_s e^{\beta\psi_s/2}}{\sqrt{2}\, L_D C_{eff}}} \tag{6.17}$$

The above differential equation describes the rate of change of the surface potential $\dot{\psi}_S$, in terms of the ramp rate α, rate of change of the flatband voltage $\dot{V}_{FB}$ which is related to ψ_S (to be shown in the next section) and the ψ_S itself. In section 6.5, this equation is numerically solved in order to describe the gate current I_G, which may be written as

$$I_G(t) = \dot{Q}_G(t) = -\dot{Q}_N(t) - \dot{Q}_S(t) \ . \tag{6.18}$$

Eliminating $\dot{Q}_S(t)$ from eqs.(6.16) and (6.18) yields

$$I_G(t) = C_{eff}[\alpha - \dot{\psi}_S(t) - \dot{V}_{FB}(t)] - \dot{Q}_N(t) \ . \tag{6.19}$$

The above equation may be re-written by substituting for $\dot{V}_{FB}(t)$ from eq.(6.11),

$$I_G(t) = [\alpha-\dot{\psi}_s(t) - \dot{Q}_N(t) \frac{X_{oxt}}{\varepsilon_{oxt}} - \frac{1}{\varepsilon_N} \frac{d}{dt} \int_o^{X_N} x\rho_N(x,t)dx]C_{eff} \qquad (6.20)$$

It is clear from the above analysis that $I_G(t)$ contains information about $\rho_N(x,t)$ and the manner in which it changes with time. Note that $-\dot{Q}_N(t)$ is positive while $-C_{eff}\dot{V}_{FB}(t)$ is negative so that the two terms oppose each other. Using eqs.(6.12) and (6.14), it can be shown that $-\dot{Q}_N(t) > -C_{eff}\dot{V}_{FB}(t)$.

6.3 Trapping in the Silicon Nitride Layer

Trapping of the carriers in the silicon nitride film is a complex issue primarily because the nature of the traps within the film is not sufficiently understood. Thus, any serious modeling effort should address the problem of identifying those traps which store the memory charge. Due to the lack of such information at present, we begin by assuming what is needed to explain the observed data in this study. Hopefully, in spite of the heuristic nature of this approach, we may be able to limit the vast number of possibilities to a few judicious choices.

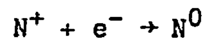
From the discussion in section (5.3) and the experimental results to be presented in Chapter 7, the following observations are made:

1. The net charge in the nitride film can be made either positive or negative by suitably biasing the MONOS structure (see Fig. 5.8). In addition, the net charge in a virgin device is small and usually positive (see Fig. 5.5).
2. The net negative charge in the nitride film is usually brought about by an injection of electrons from the silicon conduction band whereas the net positive charge is created via the hole-injection from the silicon valence band.
3. The displacement current at the gate electrode rises above αC_{eff} during the forward sweep (positive gate bias) and increases monotonically throughout the forward sweep (see Fig. 7.12).
4. The extent to which the gate current rises above αC_{eff} during the forward sweep and the slope of the rise decrease with decrease in the temperature (see Fig. 7.15).
5. The displacement current at the gate returns to $-\alpha C_{eff}$ during the reverse sweep (see Fig. 7.12).

The observations (1) and (2) indicate two possibilities. (a) the film contains an "amphoteric" trap that can exist in negative (N^-), positive (N^+) or neutral (N^0) charge states. The natural charge state in a virgin film is primarily the neutral state. With charge injection, the trap can undergo the following reactions.

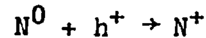
electron injection:

$$N^0 + e^- \rightarrow N^-$$

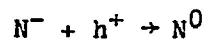


$$N^+ + e^- \rightarrow N^0$$

hole injection:



$$N^0 + h^+ \rightarrow N^+$$



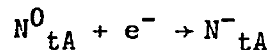
$$N^- + h^+ \rightarrow N^0$$

For a given injection condition, one of the capture processes involves a charged state and is characterized by a relatively larger capture cross-section compared to the other capture process by the neutral charge state.

(b) The nitride film contains two non-interacting sets of traps. One donor like (N_{tD}) which can exist in positive or neutral charge states and the other, acceptor like (N_{tA}) with negative and neutral charge states. During charge injection, the following capture processes may take place.

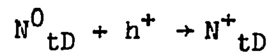
electron injection:

$$N^+_{tD} + e^- \rightarrow N^0_{tD}$$

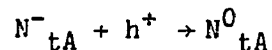


$$N^0_{tA} + e^- \rightarrow N^-_{tA}$$

hole injection:



$$N^0_{tD} + h^+ \rightarrow N^+_{tD}$$



$$N^-_{tA} + h^+ \rightarrow N^0_{tA}$$

Again, one of the capture processes involves a coulombic trap and is

characterized by a relatively large capture cross-section compared to the capture process by the neutral charge state.

The observation (3) implies that the centroid of the trapped charge in the nitride, as defined by the eq.(6.13), moves deeper into the nitride with increasing amount of the total injected charge, Q_N. To see this, let us consider the case of n-type semiconductor being driven into accumulation and assume that $\bar{X}_N$ remains constant. The surface potential ψ_S will increase with the degree of accumulation. This increases the field in the tunneling oxide (see Fig.6-1) resulting in an increase in the tunneling current. This increases Q_N and V_{FB}. Since $\bar{X}_N$ remains constant, $V_{FB} \sim Q_N$ [see eq.(6.14)]. At some value of ψ_S, $\dot{V}_{FB}(t)$ equals α and $\dot{\psi}_S$ becomes zero [eq.(6.17)]. Thus, the flatband voltage changes as fast as the gate voltage is ramped with constant values of ψ_S and Q_S in eq.(6.10). This results in [eqs.(6.14), (6.19)],

$$I_G(t) = -\dot{Q}_N(t) = \frac{\dot{V}_{FB}(t)}{\left[\frac{X_{oxB}}{\varepsilon_{oxB}} + \frac{X_N - \bar{X}_N}{\varepsilon_N}\right]}$$

$$= \alpha C_{eff} \frac{\left[\frac{X_{oxB}}{\varepsilon_{oxB}} + \frac{X_N}{\varepsilon_N} + \frac{X_{oxt}}{\varepsilon_{oxt}}\right]}{\left[\frac{X_{oxB}}{\varepsilon_{oxB}} + \frac{X_N - \bar{X}_N}{\varepsilon_N}\right]}$$

$$= \alpha C_{eff} K \tag{6.21}$$

We see that $K > 1$. As shown in Fig.6-2, $I_G/\alpha C_{eff}$ rises from 1 to K and levels off. This is quite different from the observed monotonic rise in current as shown by the broken line. Thus, the extent to which $I_G/\alpha C_{eff}$ rises above 1, is an indication of the depth of the centroid. The observation (4), then, implies that, at a lower temperature, the charge centroid is relatively closer to the tunneling oxide-nitride interface and it moves deeper at higher temperatures. In section 6.6.1, we will see that this can be explained by the temperature dependence of the capture cross-section provided detrapping is negligible.

The observation (5) has an important consequence. Initially, when the sweep reverses its direction, the sign of the ramp rate α in Eq.(6.19) changes from positive to negative. Although, $\dot{\psi}_S(t)$ becomes negative and the surface potential begins to decrease with the returning ramp, the field across the tunneling oxide is still enough to make the term $-\dot{Q}_N(t)-\dot{V}_{FB}(t)C_{eff}$ [eq.(6.19)] finite and positive which, then, keeps the gate current above $-\alpha C_{eff}$. However, when the gate voltage reduces by few volts, the field in the tunneling oxide reduces to cut off any charge injection into the nitride. This makes $-\dot{Q}_N(t) = 0$. Now, if there is no detrapping in the nitride film due to the high field, the charge distribution

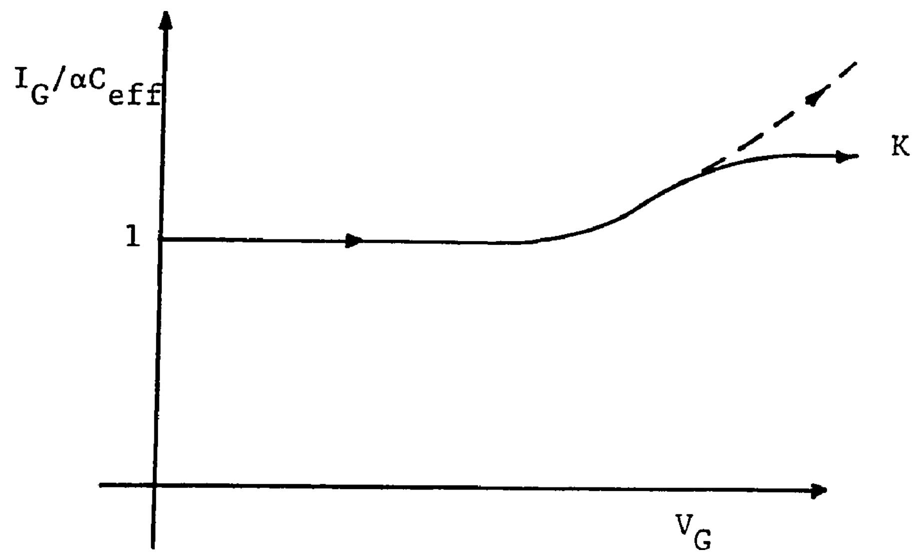


Fig. 6.2 The normalized gate current vs the gate voltage curve. The solid line is described by eq.(6.21) which is a consequence of assuming constant centroid $\bar{X}_N$. The broken line represents the observed behavior.

within the film freezes and $\dot{V}_{FB}(t)$ becomes zero. Consequently, $I_G(t)$ returns to $- \alpha C_{eff}$. Thus the fact, that $I_G(t)$ returns to $- \alpha C_{eff}$ during the reverse sweep, implies that detrapping in the film is not present even at the room temperature. This fact results in a considerable simplification of the equations which have to be solved.

Trapping Model

In order to describe what happens to the electrons as a function of space and time after they are injected in the nitride conduction band, we use the results of the analysis by Arnett [32]. The salient features of Arnett's analysis are repeated here.

The presence of the donor and acceptor like traps is assumed. In addition, it is assumed that initially all donors are empty and thus positively charged, whereas, acceptors are negatively charged and their occupancy is not altered during the electron injection. This runs against the spirit of the model we discussed before where the donors are neutralized and acceptors are negatively charged during the electron injection. However, the intent, here, is not to be precise but to gain greater physical insight. By restricting the capture process to donors only, the closed form solutions are possible to obtain. The extension to include the trapping by acceptors would not change the results qualitatively.

The basic equations used to describe the transient conduction are [see Fig.6-3]

$$\frac{\partial J_c(x,t)}{\partial x} = - \frac{\partial \rho_N(x,t)}{\partial t} , \tag{6.22}$$

$$J_c(x,t) = - qn_c(x,t)v_d, \tag{6.23}$$

$$\rho_N(x,t) = q[N_{tD}(1-f_{tD}) - N_{tA} - n_c] , \tag{6.24}$$

and $$\frac{\partial f_{tD}(x,t)}{\partial t} = n_c(x,t)v_{th}\sigma_{nD}[1-f_{tD}(x,t)]. \tag{6.25}$$

Equation (6.22) is the continuity equation with J_c, ρ_N, x and t representing the current in the conduction band, charge density, distance measured from the tunneling oxide-nitride interface and time, respectively. Equation (6.23) describes the drift current in the conduction band of the nitride with n_c the electron concentration (cm^{-3}) in the conduction band and v_d the saturated drift velocity of electrons. At high fields (>1 $MVcm^{-1}$), diffusion currents are negligible and the drift velocity saturates [88]. The charge density ρ_N (cm^{-3}) in the nitride film is given by eq.(6.24). The donor (N_{tD}) and Acceptor (N_{tA}) charge densities (cm^{-3}) are assumed to be about equal. The occupancy function of the donor like trap is represented by f_{tD}. The nitride film is assumed to be neutral prior to charge injection. Donor like centers are

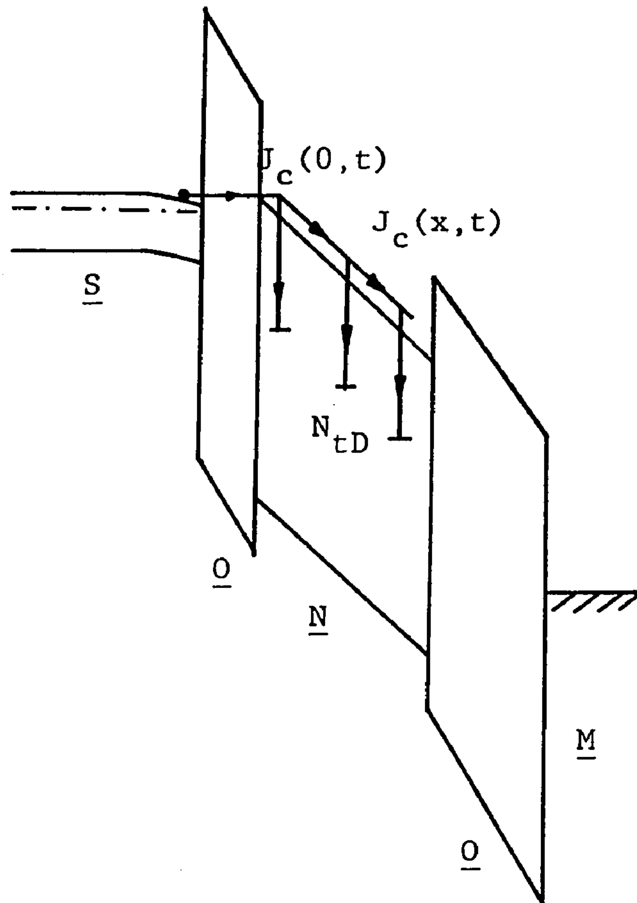


Fig. 6.3 The energy band diagram showing electrons being injected from the semiconductor into the nitride layer with subsequent capture by the positively charged donor traps.

positively charged and are compensated by the filled acceptor sites whose charge state is not altered during the charge injection. Thus, $f_{tD}(x,0) = 0$, and $N_{tD} = N_{tA}$ in eq.(6.24). In addition, for the practical values of the injection current $J_c(0,t)$ encountered in the ramp measurements, n_c may be neglected in eq.(6.24). Equation (6.25) states that the time rate of change of the occupancy of the donor like traps is equal to the capture rate of the electrons. σ_{nD} (cm^2) is the thermal capture cross-section of the donor like traps. Note that in view of the previous discussion, emission term is not included in the eq.(6.25).

Equations (6.22) to (6.25) are now solved for $f_{tD}(x,t)$. Combining eqs.(6.22) to (6.24), we get

$$\frac{\partial n_c(x,t)}{\partial x} = -\frac{N_{tD}}{v_d}\left[\frac{\partial}{\partial t} f_{tD}(x,t)\right]. \tag{6.26}$$

Solving eq.(6.25) for $n_c(x,t)$ and substituting in eq.(6.26), we get

$$\frac{\partial}{\partial t}\left[x_t \frac{\partial}{\partial x} \ln\{[1-f_{tD}(x,t)]^{-1}\} + f_{tD}(x,t)\right] = 0 \ , \tag{6.27}$$

$$\text{where, } x_t \doteq \text{trapping length } \frac{1}{N_{tD}\hat{\sigma}_{nD}} = \frac{1}{N_{tD}(\sigma_{nD}\frac{v_{th}}{v_d})} \ . \tag{6.28}$$

Integrating eq.(6.27) once and using $f_{tD}(x,0) = \frac{\partial f_{tD}}{\partial x}(x,o) = 0$, yields

$$\frac{\partial f_{tD}(x,t)}{\partial x} + \frac{f_{tD}(x,t)[1-f_{tD}(x,t)]}{x_t} = 0. \tag{6.29}$$

The solution of eq.(6.29) is given by (assuming x_t = constant)

$$f_{tD}(x,t) = [1 + A(t)e^{x/x_t}]^{-1} \tag{6.30}$$

where, A(t) is independent of x and only function of time. To determine A(t), consider eqs.(6.23) and (6.25) at x=0:

$$\frac{\partial f_{tD}(o,t)}{\partial t} = -\frac{J_c(o,t)}{q}\hat{\sigma}_{nD}[1-f_{tD}(o,t)]$$

Solving the above equation subject to the initial condition $f_{tD}(x,0)=0$, yields

$$f_{tD}(o,t) = 1 - e^{-\frac{\hat{\sigma}_{nD}}{q}\int_0^t -J_c(0,\tau)d\tau}.$$

Substituting x=0 in eq.(6.30) gives

$$f_{tD}(0,t) = [1 + A(t)]^{-1}.$$

Thus, $A(t) = [e^{\hat{\sigma}_{nD}Q_{inj}(t)/q} - 1]^{-1}$,

with $Q_{inj}(t) = \int_o^t - J_c(o,\tau)d\tau.$ (6.31)

Substituting the expression for A(t) in eq.(6.30), we get

$$f_{tD}(x,t) = \frac{1}{1 + \frac{e^{x/x_t}}{e^{\hat{\sigma}_{nD} Q_{inj}(t)/q} - 1}} . \tag{6.32}$$

The nitride charge may now be calculated by integrating $\rho_N(x,t)$ over the nitride thickness:

$$Q_N(t) = -qN_{tD}\int_o^{X_N} f_{tD}(x,t)dx$$

If $x_t << X_N$, then we can let $X_N \to \infty$

$$Q_N(t) = \lim_{X_N \to \infty}[-q\, N_{tD} \int_o^{X_N} f_{tD}(x,t)dx]$$

$$= -Q_{inj}(t) = \int_o^t J_c(o,\tau)d\tau. \tag{6.33}$$

differentiating $Q_N(t)$ yields

$$\dot{Q}_N(t) = J_c(o,t) . \tag{6.34}$$

Flatband voltage can be written as [eq.(6.11)]

$$V_{FB}(t) = Q_{inj}(t)(\frac{X_{oxB}}{\varepsilon_{oxB}} + \frac{X_N}{\varepsilon_N}) - \frac{qN_{tD}}{\varepsilon_N}\int_o^{X_N} xf_{tD}(x,t)dx. \tag{6.35}$$

The integral in the above equation can not be evaluated in closed form and has to be evaluated numerically. However, $\dot{V}_{FB}(t)$ can be expressed in a closed form:

$$\dot{V}_{FB}(t) = -J_c(0,t)\left[\frac{X_{oxB}}{\varepsilon_{oxB}} + \frac{X_N}{\varepsilon_N}\right] - \frac{qN_{tD}}{\varepsilon_N}\int_0^{X_N} x\dot{f}_{tD}(x,t)\,dx.$$

Letting $X_N \to \infty$, we obtain after some manipulation,

$$\dot{V}_{FB}(t) = -J_c(o,t)\left[\frac{X_N}{\varepsilon_N} + \frac{X_{oxB}}{\varepsilon_{oxB}} - \frac{Q_{inj}(t)}{\varepsilon_N qN_{tD}}\left\{1 + \frac{1}{e^{\frac{\hat{\sigma}_{ND}Q_{inj}(t)}{q}} - 1}\right\}\right]. \quad (6.36)$$

The trap occupancy function as described by eq.(6.32) is plotted as a function of x/x_t for various degrees of charge injection, $\hat{\sigma}_{nD}Q_{inj}/q$ in Fig.6-4. The circles represent the centroids [eqs.(6.13)] of the charge distributions corresponding the plots of the occupancy function. It is evident from the figure that for low level injection ($\hat{\sigma}_{nD}Q_{inj}/q < 1$), the centroid of the trapped charge is located at the trapping distance x_t from the tunneling oxide-nitride interface. As the degree of charge injection increases, more traps are occupied and the distribution close to the tunneling oxide-nitride interface, begins to saturate. Consequently, the centroid moves deeper into the nitride layer. For, $\hat{\sigma}_{nD}Q_{inj}/q = 5$ the centroid $\bar{X}_N$ is located at about $3x_t$. This movement of the centroid makes a given amount of the trapped charge less effective in shifting the flatband voltage [eq.(6.14)]. As explained earlier in this chapter [observation(3)], the extent to which the gate current rises above αC_{eff}, is determined by the movement of the centroid. This will be demonstrated more clearly in

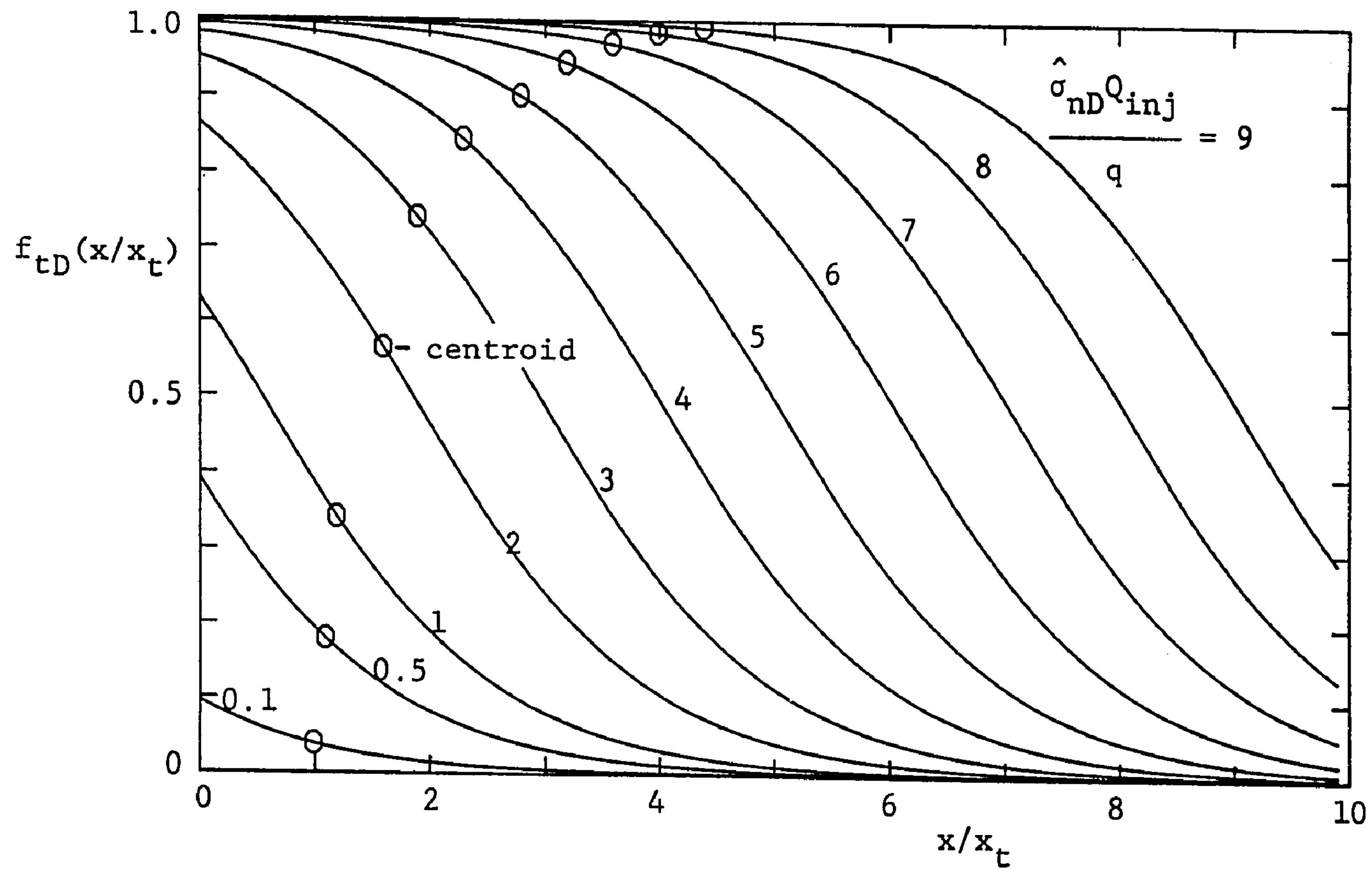


Fig. 6.4 The plots of the trap occupancy function $f_{tD}(x/x_t)$ vs x/x_t as described by eq.(6.32) with normalized charge injection as a parameter. "O" denotes the centroid of the various charge distributions as defined in the eq.(6.13). Trapping length x_t is given by $x_t = 1/(\hat{\sigma}_{nD}N_{tD})$

section 6.6.

6.4 Tunneling Model

Historically, the first tunneling model for the operation of the MNOS devices was given by Ross and Wallmark [64]. They considered direct tunneling of carriers between silicon energy bands and spatially distributed traps [Fig.6-5a]. Dorda and Pulver [65] treated band to trap tunneling for spatially fixed but energetically distributed traps [Fig.6-5b] while Ferris-Prabhu [66], [67], [68] considered traps with both spatial and energetic distribution [Fig.6-5c]. In all these treatments, the authors assumed rectangular potential barriers and neglected the effect of the reduced field in the tunneling oxide due to the charging of the traps. White and Cricchi [69] removed these assumptions but considered a single trap to obtain closed form solutions [Fig.6-5d]. However, all these theories involved two key assumptions: (1) the charge is assumed to be trapped at or near the tunneling oxide-nitride interface, and (2) the tunneling is assumed to be between the silicon energy bands and the traps in the insulator located energywise, within or near the silicon bandgap [Fig.6-5]. The first assumption is really a consequence of the second assumption. Band to trap tunneling requires small tunneling distances (< 40A) and, therefore, the traps are placed as close to the semiconductor as possible. Studies by Yun et al [72, 73, 74, 89] and Lehovec [90] have conclusively shown that the charge is distributed upto atleast

300A into the nitride layer. Explanations based on band to trap tunneling can not explain how the carriers, if they are in such deep traps, can propogate hundreds of angstorms into the nitride.

In view of the above discussion, we consider the Modified Fowler Nordheim band to band tunneling [Fig.6-6] as first proposed by Lundstrom and Svensson [70]. Chang [71] has also argued in favor of the band to band tunneling in these structures. However, he has formulated the tunneling probability through the oxide and nitride layers with two rectangular potential barriers of different heights. The resulting expression involves three fitting parameters and no attempt is made to relate these parameters with barrier heights and effective mass of the carriers.

The modified FN tunneling expression given by Lundstrom and Svensson [70] involves the following assumptions:

1. Single band model in SiO_2 and Si_3N_4.
2. Conservation of transverse momentum.
3. WKB approximation.
4. Effective masses of the electron in the oxide and nitride layer are identical.

The tunneling current $J_c(0,t)$ is given by

$$J_c(0,t) = C_{FN}\xi_{oxt}^2 \, P_{ox}P_N \tag{6.37}$$

$$\text{with } C_{FN} = q^3/[16\pi^2\hbar\phi_1], \tag{6.38}$$

$$P_{ox} = \exp\{-\frac{4\sqrt{2m^*}}{3\hbar q}[\phi_1^{3/2} - (\phi_1 - q\xi_{oxt}X_{oxt})^{3/2}]/\xi_{oxt}\},$$

and, $$P_N = \exp\{-\frac{4\sqrt{2m^*}}{3\hbar q}(\phi_1 - \phi_2 - q\xi_{oxt}X_{oxt})^{3/2}/\xi_N\}.$$

The time dependence of $J_c(0,t)$ comes from $\xi_{oxt}(t)$ and $\xi_N(x,t)$. For simplicity, we assume that ξ_N is constant in the tunneling region and is related to ξ_{oxt} in the following manner:

$$\xi_N = \xi_{oxt}\varepsilon_{oxt}/\varepsilon_N \ . \tag{6.39}$$

Using $\hbar=1.054\times10^{-34}$ joule.sec, $q=1.602\times10^{-19}$ coulomb and $\phi_1=3.1$ eV$=3.1\times1.602\times10^{-19}$ joules, C_{FN} is calculated to be 5.0×10^{-7} AV^{-2}. Using $\phi_2=1.05$ eV, J_c vs ψ_s is plotted in Fig.6-7 for several values of C_{FN}. The surface potential ψ_s uniquely determines the electric field ξ_{oxt} in the tunneling oxide via eqs.(6.4) and (6.6). It is noted that the choice of C_{FN} does not appreciably affect the value of ψ_s required to obtain the same value of the current density. To obtain 10^{-8} Acm^{-2}, a 4 order of magnitude change in C_{FN} is equivalent to a 12 mV change in ψ_s as shown in Fig. 6.7.

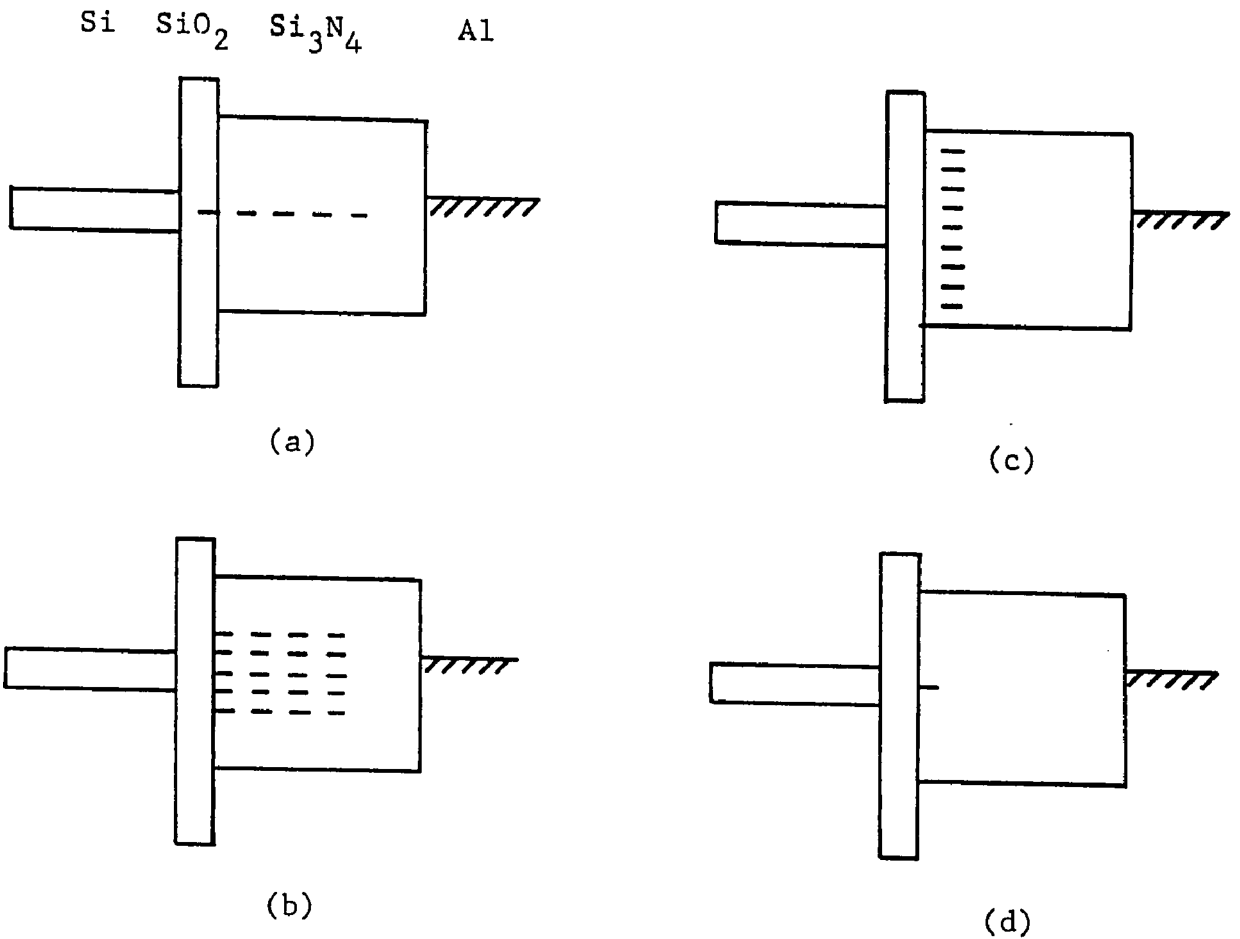


Fig. 6.5 The trap distributions used by previous authors: (a) Ross and Wallmark model, (b) Dorda and Pulver model, (c) Ferris-Prabhu model, and (d) White and Cricchi model.

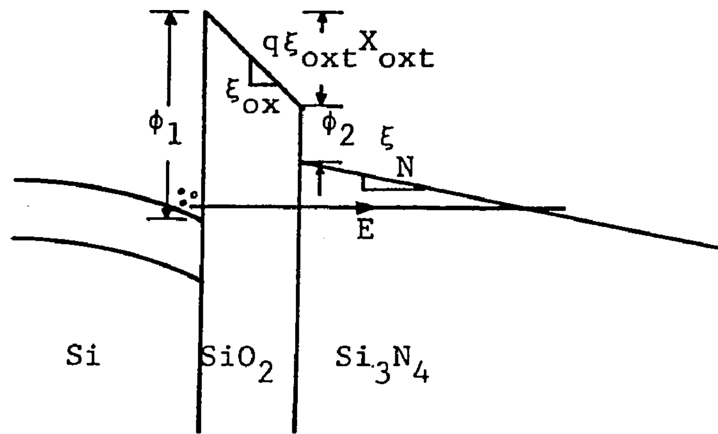


Fig. 6.6 The energy band diagram showing the Modified Fowler-Nordheim tunneling through the oxide and nitride barriers.

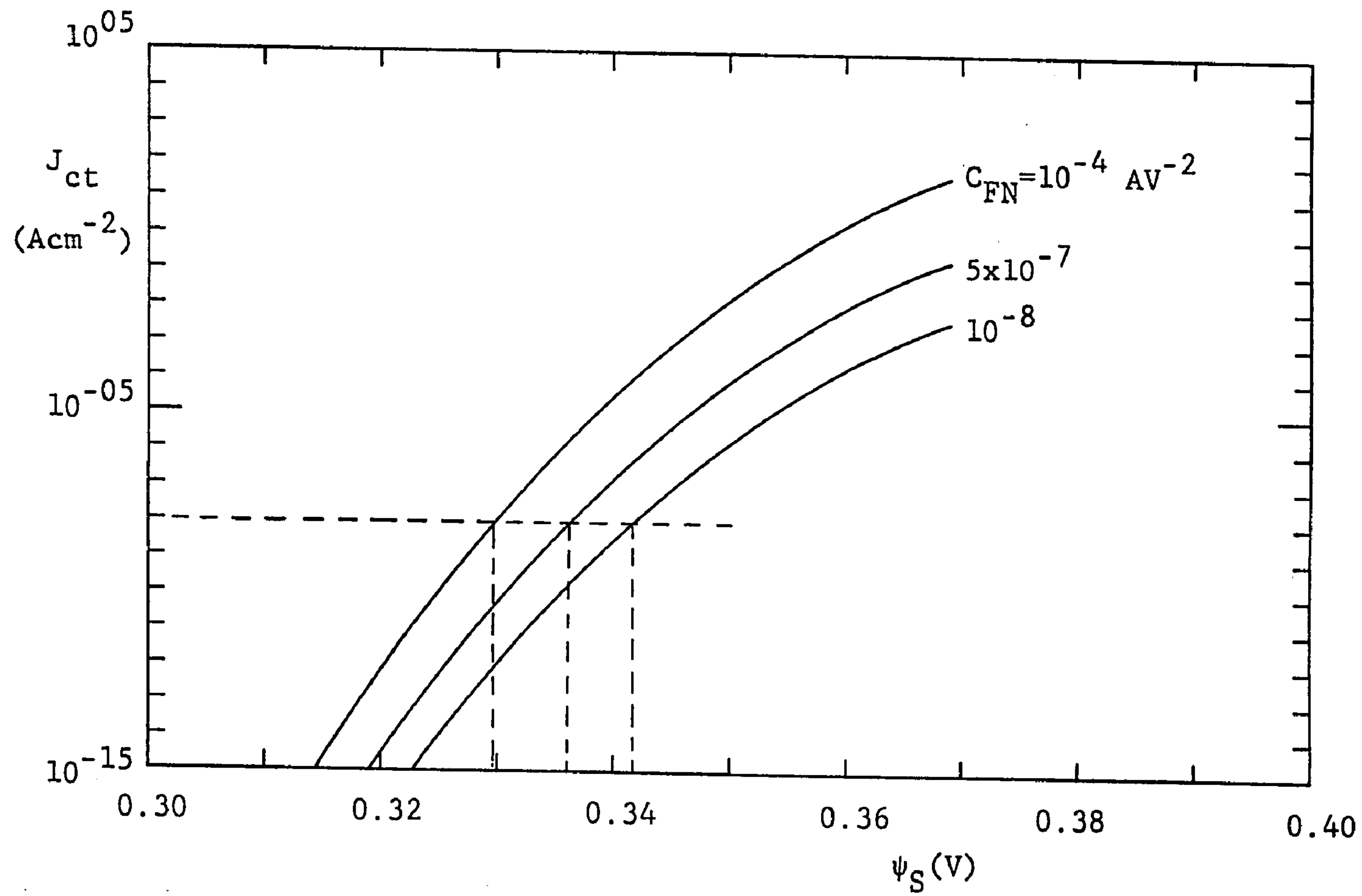


Fig. 6.7 Tunneling Current vs the surface potential as described by eqs. (6.4), (6.6) and (6.37) for several values of C_{FN}.

6.5 The Overall Model

In this section, we interconnect the tunneling model, the model for trapping in the nitride layer and, the results from the dynamic analysis of the linear voltage ramp. The goal, here, is to describe the gate current vs gate voltage characteristic as function of the tunneling and trap parameters. No attempt is made to extract any device parameters because there is a large number of them and none can be independently measured. The emphasis is placed on the trends in the gate current vs gate voltage characteristics as some of the parameters are changed.

We begin by observing that $\dot{\psi}_S(t)$ in eq.(6.17) depends upon $\psi_S(t)$ and $\dot{V}_{FB}(t)$ once the device dimensions and the ramp rate α, are fixed. The differential equation (6.17) in $\psi_S(t)$ can be numerically solved provided $\dot{V}_{FB}(t)$ may be determined in terms of in terms of $\psi_S(t)$ alone. To this end, we note that in eq.(6.36), $\dot{V}_{FB}(t)$ is determined by the knowledge of $J_c(0,t)$ and $Q_{inj}(t)$ for a given choice of trap parameters $\hat{\sigma}_{nD}$ and N_{tD}. The current density $J_c(0,t)$ may be determined via eq.(6.37) in terms of ξ_{oxt} alone and then, $Q_{inj}(t)$ can be computed by integrating the past values of $J_c(0,t)$ as shown by eq.(6.31). Now, ξ_{oxt} is only dependent upon ψ_S as evident from eqs.(6.4) and (6.6). Thus, $\dot{V}_{FB}(t)$ can indeed be uniquely determined in terms of $\psi_S(t)$ alone, and the differential equation (6.17) can be solved for $\psi_S(t)$ starting with an initial value in accumulation, $\psi_S(0)$. Once $\psi_S(t)$ is known, $I_G(t)$ [eq.(6.19)] and

$V_G(t)$ [eq.(6.10)] are easily calculated in terms of $\psi_S(t)$ and its past values. A detailed flow chart of this calculation is shown in Fig.6-8.

The 4th order Runge-Kutta technique [91] is used to solve eq.(6.17). This technique is briefly summarized below and the details may be found in the reference cited or in any text on the finite-difference methods.

Fourth Order Runge-Kutta Technique

A function y(t) that can be expressed by a differential equation of the form, dy/dt = f(y,t), may be solved for provided $y_0=y(t=t_0)$ is known. The entire time duration, in which the solution is desired, is devided into N intervals of size "h" and denoted by $t_0,t_1,\ldots,t_n,\ldots$ as shown in Fig.6-9. The value of the function at the end of each time interval is denoted by $y_0,y_1,\ldots,y_n,\ldots$ etc. The value of y_{n+1}, at time t_{n+1}, may be determined from y_n by the following algorithm:

$$y_{n+1} = y_n+(K_0+2K_1+2K_2+K_3)/6 \qquad (6.40)$$

with, $K_0 = hf(y_n,t_n)$,

$K_1 = hf(y_n+K_0/2,\ t_n+h/2)$,

$K_2 = hf(y_n+K_1/2,\ t_n+h/2)$,

and, $K_3 = hf(y_n+K_2,\ t_n+h)$.

It is evident that the algorithm is self-starting if y_0 at $t=t_0$ is

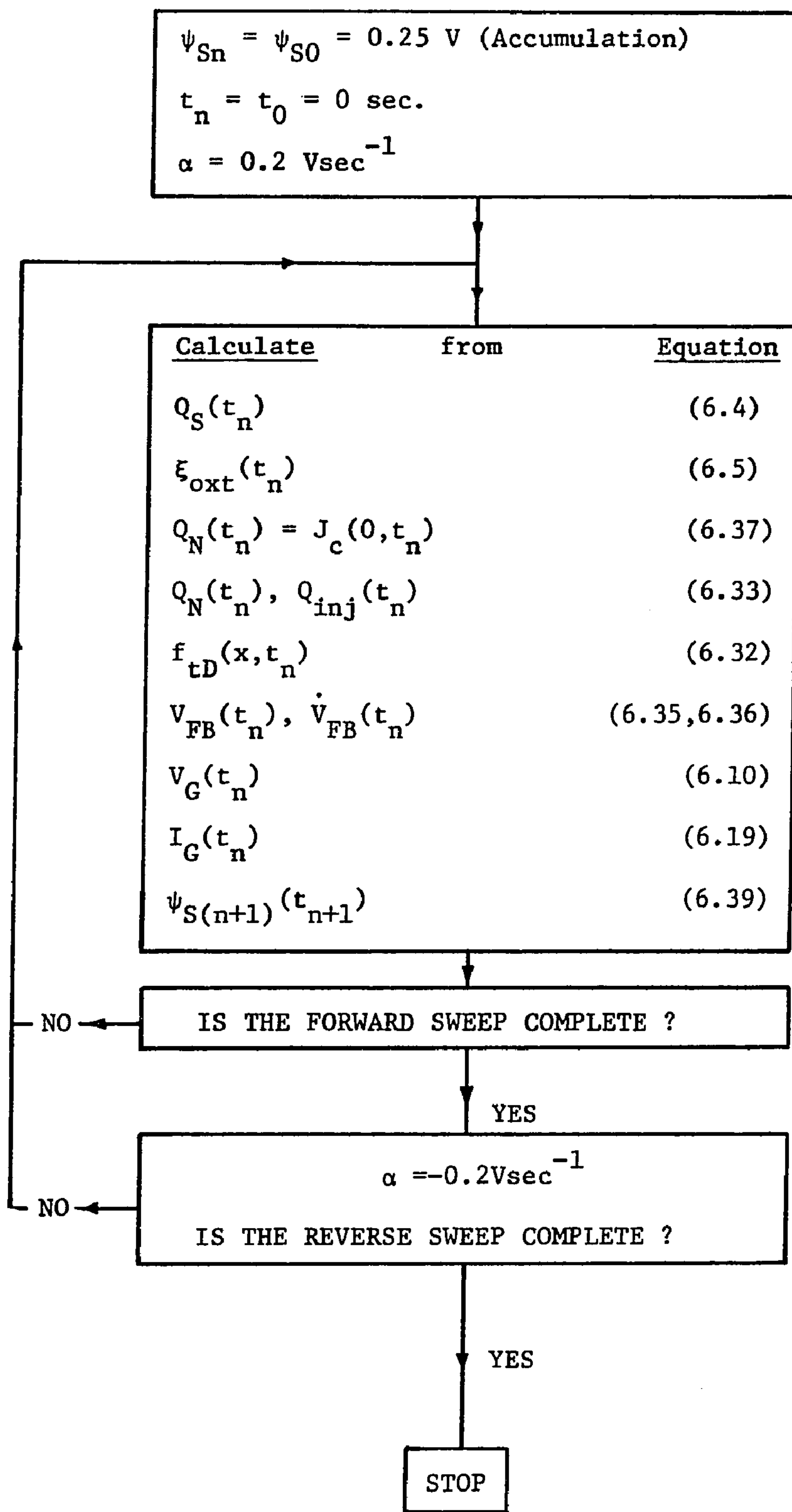


Fig. 6.8 Flow-chart for the numerical simulation of the overall model.

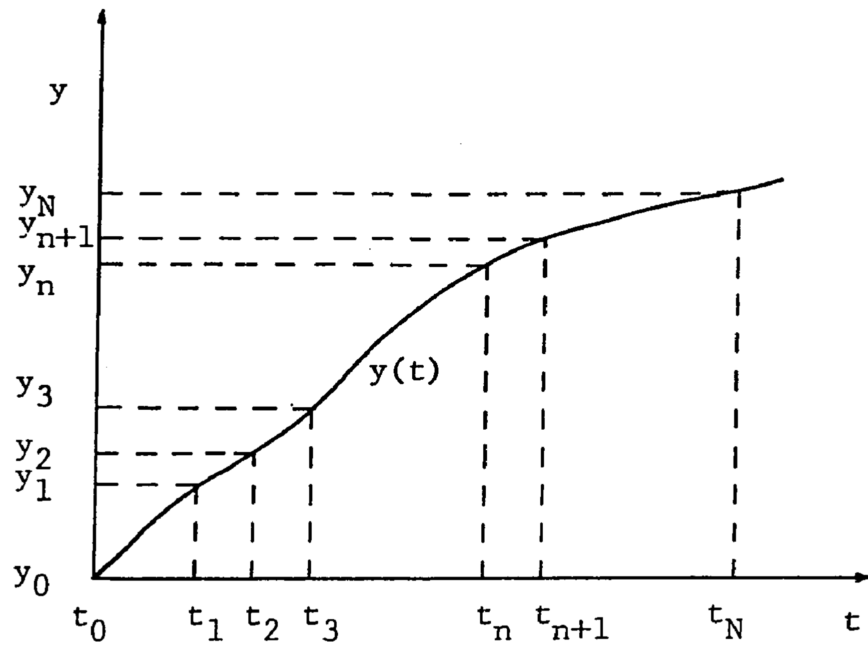


Fig. 6.9 The time sequence in the fourth order Runge Kutta Technique.

known.

6.6 Results of the Numerical Simulation

The results of the numerical simulation on a typical device are shown in Fig.6-10. The device dimensions and the trap parameters chosen for the simulation are indicated in the figure caption. The device dimensions are representative of the p-channel sample chosen for the measurements in this study. The trap parameters N_{tD} = $5x10^{18}cm^{-3}$ and $\hat{\sigma}_{nD}$ = $5x10^{-13}cm^2$ (x_t=40A) are very close to the values suggested in reference [5]. With this choice of the parameters, the condition $x_t << X_N$ is satisfied. The two curves in Fig.6-10a represent the normalized gate current vs gate bias characteristics as the gate bias is ramped from 0V to +24V (curve 1), to +26V (curve 2) and back. The semiconductor (n type) is driven into accumulation during the forward sweep. The normalized gate current, $I_G/\alpha C_{eff}$ exceeds unity during the forward sweep around V_G~18V due to the electron injection from the semiconductor into the nitride film. During the reverse sweep, it returns to -1 since the detrapping has been neglected due to the experimental observations reported in chapters 5 and 7.

The plots of the normalized charge $Q_n\hat{\sigma}_{nD}/q$ vs V_G are shown in Fig.6-10b. By the end of the forward sweep in curve 2, $Q_n\hat{\sigma}_{nD}/q$ ~ 2.7 which indicates [Fig.6-4] that the charge centroid has moved to about $2x_t$=80A for a flatband shift of about 7.5V [Fig.6-10c]. As

expected, the onset of the shift in Q_N and V_{FB} (around V_G=18V) coincides with V_G at which $I_G/\alpha C_{eff}$ exceeds unity. During the reverse sweep, $Q_N\hat{\sigma}_{nD}/q$ and V_{FB} do not change any more because of a rapid decrease in ψ_S which cuts off the electric field ξ_{oxt} and the tunneling current J_{ct}.

Figure 6-10d illustrates that, during the forward sweep, the surface potential ψ_S increases upto about 0.338V at the onset of charge injection (V_G~18V) and then increases very slowly (less than 1mV) during the charge injection. This checks against Fig.6-7 [with $C_{FN}=5\times10^{-7}$ AV^{-2} and ψ_S~0.338V, the tunneling current J_{ct} is little over $10^{-8}Acm^{-2}$]. Note that, in this case, αC_{eff} is about $2\times10^{-8}Acm^{-2}$. Less than a 1mV change in ψ_S is sufficient to supply the excess current, $(I_G-\alpha C_{eff})$. During the reverse sweep, ψ_S decreases and cuts off the charge injection. The electric field in the tunneling oxide, which is determined by the surface potential alone, changes from less than $1MVcm^{-1}$ to about $7MVcm^{-1}$ and then levels off at the onset of charge injection [Fig.6-10e]. Figure 6-10f illustrates the charge injected into the nitride, Q_N, as a function of the flatband voltage, V_{FB}. The curve has a little curvature at higher values of V_{FB} due to the centroid movement [eq.(6.14)].

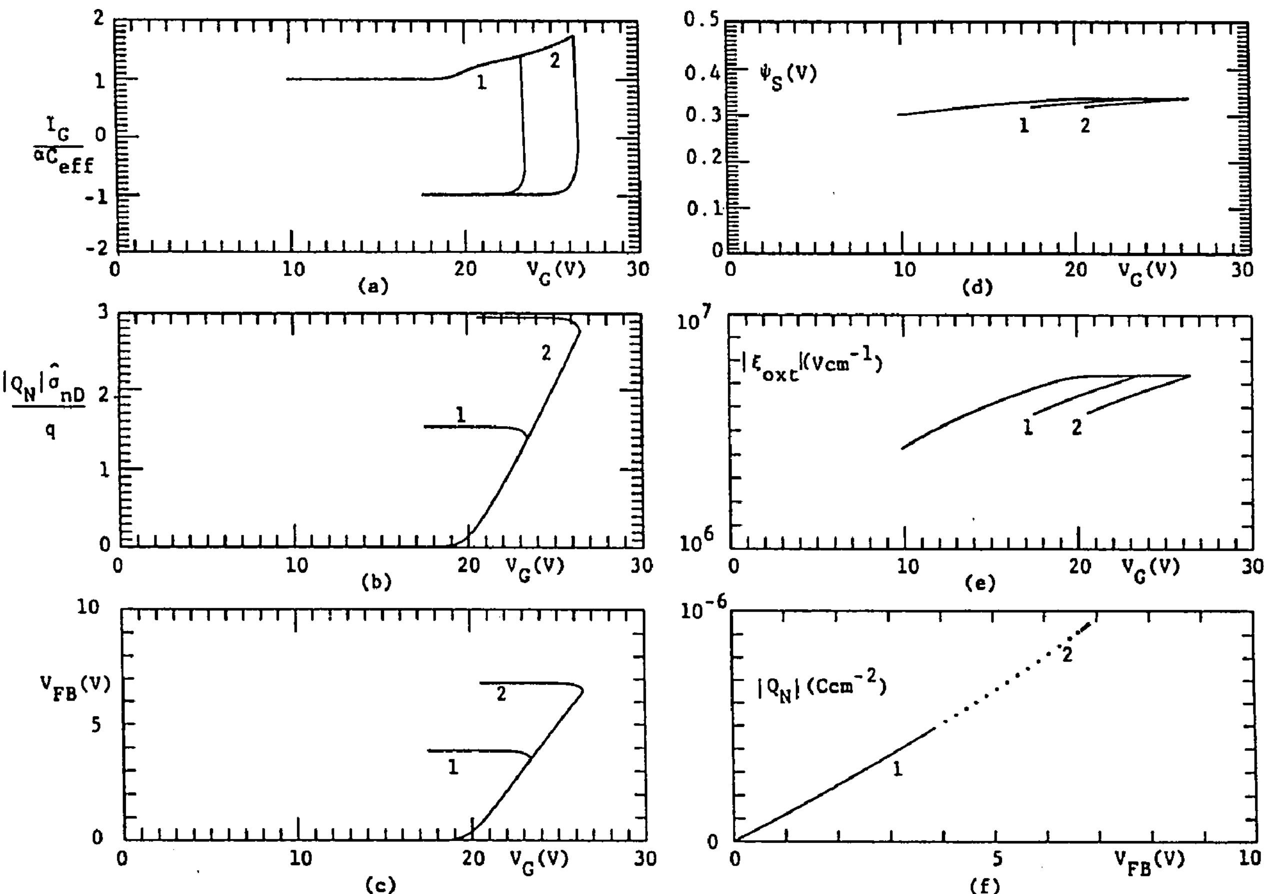


Fig.6.10 Numerical simulation of the ramp measurement on the MONOS structure with n-type silicon and positive bias on the gate. Curves 1 and 2 are computed for different maximum positive values of the bias. The device parameters used are X_{oxt}=20A, X_{oxB}=40A, X_N=500A, $N_{tD}=5x10^{18}$ cm^{-3}, $\sigma_{nD}=5x10^{-13}cm^2$($x_t$=40A), $C_{FN}=5x10^{-7}A.V^{-2}$, ϕ_1=3.1eV, ϕ_2=1.05eV, $\varepsilon_{oxt}=\varepsilon_{oxB}=3.9\varepsilon_o$, $\varepsilon_N=6.5\varepsilon_o$ and $\alpha=0.2Vs^{-1}$.

6.6.1 Effect of Varying the Trap Capture Cross-section:

Figure 6-11 illustrates the various curves of interest as the capture cross-section is varied from $2.5x10^{-13}cm^2$ (x_t= 160A) to $10^{-12}cm^2$ (x_t=40A) with $N_{tD}=2.5x10^{18}cm^{-3}$. With a smaller value of the capture cross-section, trapping length x_t is larger and therefore, more charge has to be injected to produce the same flatband shift as demanded by the voltage ramp. This explains the trends in $I_G/\alpha C_{eff}$ vs V_G and Q_N vs V_{FB} curves with various values of $\hat{\sigma}_{nD}$. With a higher value of $\hat{\sigma}_{nD}$, $Q_N\hat{\sigma}_{nD}/q$ is higher implying that the traps closer to the tunneling oxide-nitride interface fill easily [Fig.(6.4)] but because of the smaller trapping length, x_t, charge does not penetrate as much.

The motivation behind studying the effect of varying the capture cross-section arises from the temperature dependence of attractive capture processes (for a review see [92], [93]). According to the theoretical models proposed by Lax [94], [95], an attractive center should have a capture cross-section that varies T^{-K}, where K=2 to 5. In chapter 7, ramp I-V measurements at the room temperature and 100K are reported. The observations are consistent with T^{-K} dependence. The value of K is not easy to determine because the capture cross-section of an attractive center has a strong field dependence which, in turn, depends on the temperature too (see [96]).

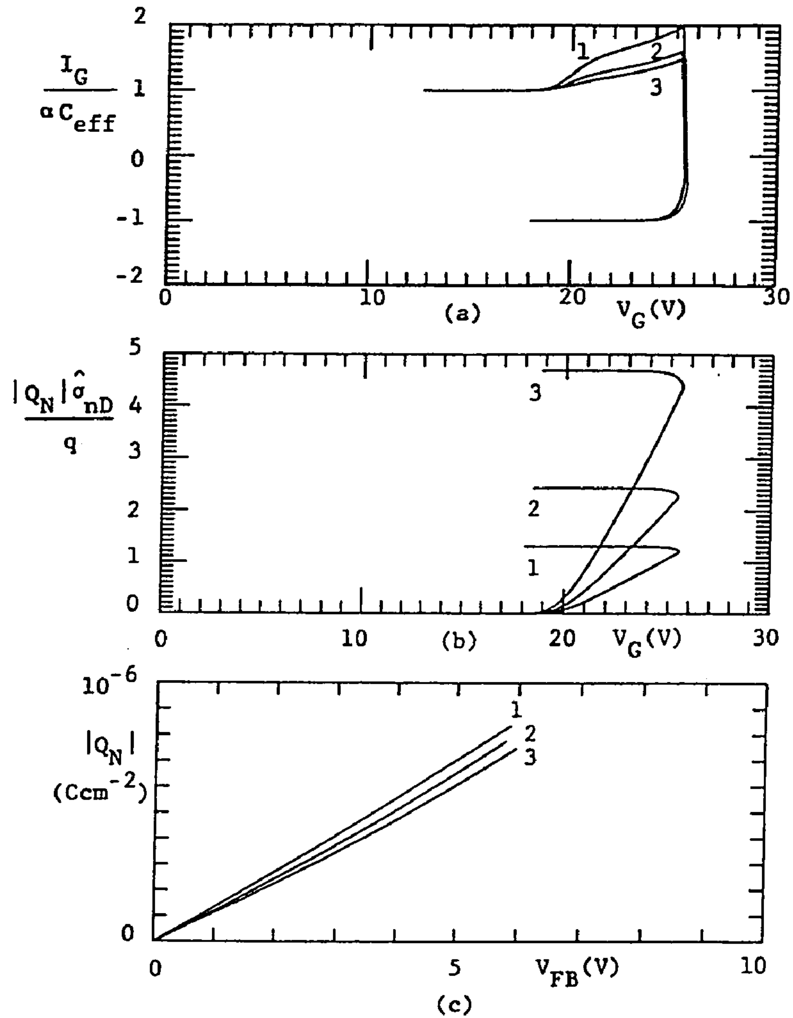


Fig.6.11 Plots of (a) $I_G/\alpha C_{eff}$ vs V_G, (b) $|Q_N|\hat{\sigma}_{nD}/q$ vs V_G, and (c) $|Q_N|$ vs V_{FB} for various values of $\hat{\sigma}_{nD}$ at a fixed value of N_{tD} = $2.5\text{x}10^{18}$ cm^{-3}. Curve 1 corresponds to $\hat{\sigma}_{nD}$=$2.5\text{x}10^{-13}$ cm^2 (x_t=160 A), curve 2 to $\hat{\sigma}_{nD}$=$5.0\text{x}10^{-13}$ cm^2 (x_t=80 A), and curve 3 to $\hat{\sigma}_{nD}$=$1.0\text{x}10^{-12}$ cm^2 (x_t=40 A). Other parameters are same as indicated in Fig. 6.10.

6.6.2 Effect of Varying the Trap Density:

The influence of the trap density is illustrated in Fig.6-12. At a fixed value of $\hat{\sigma}_{nD}=5x10^{-13}cm^2$, N_{tD} is varied from $2.5x10^{18}cm^{-3}(x_t=80A)$ to $10^{19}cm^{-3}(x_t=20A)$. As the trap density is decreased, the trapping length increases. Consequently, the charge spreads deeper into the nitride layer and therefore becomes less effective in shifting the flatband voltage. Therefore, to keep pace voltage ramp, more charge needs to be injected. This explains the increase in the excess gate current $(I_G-\alpha C_{eff})$ with decreasing trap density and the corresponding trends in the Q_N vs V_{FB} curves. In addition, for lower values of N_{tD}, traps fill up quickly and so the centroid moves further into the nitride layer. This explains why $\hat{\sigma}_{nD}Q_N/q$ is higher for lower values of N_{tD}.

The sensitivity of the gate current to the trap density can be fruitfully used to investigate various annealing conditions and the effect of cycling on the bulk trap density in the nitride.

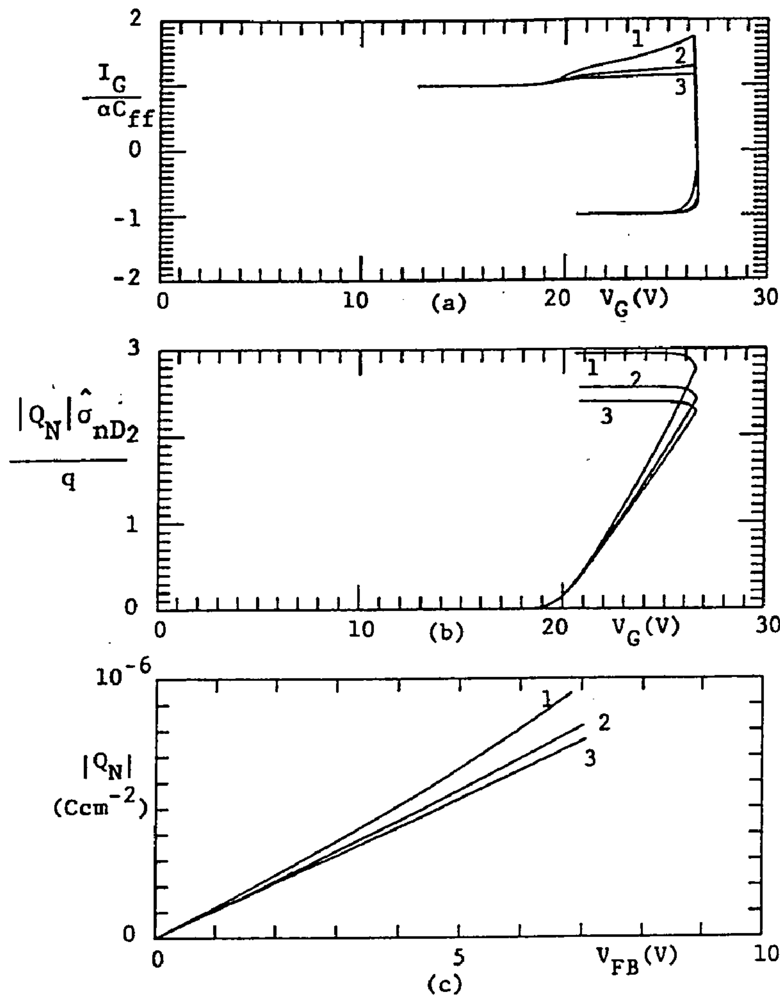


Fig.6.12 Plots of (a) $I_G/\alpha C_{eff}$ vs V_G, (b) $|Q_N|\hat{\sigma}_{nD}/q$ vs V_G, and (c) $|Q_N|$ vs V_{FB} for various values of N_{tD} at a fixed value of $\hat{\sigma}_{nD}=5.0\times10^{-13}cm^2$. Curve 1 corresponds to $N_{tD}=2.5\times10^{18}cm^{-3}$ (x_t=80A), curve 2 to $N_{tD}=5\times10^{18}cm^{-3}$ (x_t=40A), curve 3 to $N_{tD}=10^{19}cm^{-3}$ (x_t=20A). All other parameters are same as indicated in Fig. 6.10.

CHAPTER VII

EXPERIMENTAL RESULTS ON MNOS/MONOS STRUCTURES

7.1 Fabrication

Figure7-1 illustrates the test pattern designed to characterize the silicon nitride films and the MNOS/MONOS structures. These include large area capacitors and transistors and small geometry transistors of the conventional and the drain-source protected type [97]. We are primarily interested in transistors T_2 (n-channel) and T_4 (p-channel). The geometries of T_2, T_4 are identical and the details are shown in Fig7-2. The design is compatible with the CMOS processing and permits the fabrication of both n-channel and p-channel transistors on the same chip.

Process sequence employed in the fabrication of devices is discussed below. Cross-sections of the n-channel and p-channel devices at various points in the fabrication sequence are illustrated in Fig7-3.

(i) Wafer Cleaning

The starting material is phosphorous-doped n-Si<100> of resistivity 5-10 ohm cm. The wafer is degreased in warm trichloroethylene, acetone, and methanol for 5 minutes each and is then rinsed thoroughly in high resistivity (18 megaohm.cm) DI water.

It then goes through a boil in hot sulphuric acid for nearly 10 minutes and is rinsed again. This is followed by a boil in hot nitric acid and DI water rinse. The wafer is now etched in 10% HF (10 sec.) to remove any oxide and is rinsed thourghly.

(ii) Intrinsic Gettering

The cleaned wafer is loaded into an oxidation furnace to grow 6000-7000A thick oxide in wet oxygen at 1100°C. The wafer is then subjected to a nitrogen anneal at 1100°C for 16 hours. The purpose of this anneal is to anneal out the defects in the wafer through oxygen precipitation.

(iii) P^-well Formation

The wafer is coated with Hunt positive photoresist and windows are exposed in regions for the formation of P^- well, PR is stripped in acetone, rinsed, and the wafer is cleaned using hot sulphuric acid, hot nitric acid sequence. The wafer is etched in 10% HF for 10 sec. The wafer is then loaded into the diffusion furnace for Boron predeposition at 900°C for 15 minutes. After predeposition, the boron glass is removed in buffered HF and a silicon dioxide layer about 3000A is now grown in wet oxygen at 1100°C. This is followed by drive-in diffusion at 1150°C in nitrogen ambient for 24 hours. This would result in P^- well nearly 8-9 microns deep.

(iv) P^+ Diffusion

Windows are opened for isolation diffusion as well as P^+ contacts to the substrate. Boron is predeposited at 900°C for 30 minutes, glass stripped in buffered HF and an oxide is grown in wet oxygen at 1000°C.

(v) N^+ Diffusion

The wafer is coated with photoresist and windows are opened for N^+ source and drain diffusions as well as substrate contacts. Phosphorous glass is removed and the wafer is loaded for growing an oxide (wet) about 3000A.

(vi) Formation of Gate Oxide

Windows are now opened for the gate oxide formation, contact windows are opened also in order to reduce the etching time prior to metal evaporation. The oxide is about 1000A thick and is grown in dry O_2 at 1000°C.

(vii) Memory Oxide-Nitride

The growth of the thin silicon dioxide and the nitride layer is the most important step in the MNOS fabrication. Memory windows are opened for this purpose, the PR stripped, and the wafer is cleaned in warm organic solvents and rinsed in DI water. This is followed by a 10 sec. etch in 10% HF.

The wafers are now introduced into the LPCVD reactor for silicon nitride deposition where a nitride film 400-500A would be deposited from very high-purity dichlorosilane and ammonia (30:100) at a total reactor pressure below 1 torr at about 750°C. The wafer cross-section at this point is shown in Fig7-3e.

Alternatively, a thin oxide may be grown by just dipping the wafer in DI water prior to introducing the wafer into the LPCVD reactor.

(viii) Contact Formation

After the nitride is deposited the wafer is coated with a low temperature deposited silicon dioxide which serves as a masking layer for etching the silicon nitride in regions where contact windows are to be opened. First the deposited oxide is removed from regions for contact windows using buffered HF and the silicon nitride would then be etched in hot phosphoric acid. PR is removed before etching the nitride. After this step, the deposited oxide is removed and the oxide in the contact windows is etched with the contact window mask. After removing the photoresist, the wafer has the topography shown in Fig7-3f.

The wafers are cleaned again with hot H_2SO_4 and H_3NO_4 followed by the DI rinse and into the LPCVD furnace for depositing the blocking oxide. The furnace temperature is set at 700°C, pressure

at 1.0 Torr, ratio $N_2O:SiH_2Cl_2$=200:20 and deposition time at 40 min. This results in approximately 74A of the deposited oxide. With the contact window mask, the oxide in the contact windows is etched again. After removing the photoresist, the cross-section appears as shown in Fig7-3g.

After cleaning the wafer, Aluminum is evaporated in a high vacuum system and sintered in forming gas at 450°C for 20 minutes. The metal is now patterned using the metal mask. The final cross-section is illustrated in Fig7-3h. A photo-micrograph of the fabricated devices is shown in Fig7-4.

7.2 Ramp Measurement

In this section, the experimental results of the linear voltage ramp measurement on the n and p channel devices are described. The ramp rate, α, used for these studies is 0.2 $Vsec^{-1}$. The measurements have been performed at room temperature(295 K) and low temperature (100 K) for both Erase and Write modes of operation. Although the charge separation is not effective in the Erase mode due to a high probability of recombination as discussed in chapter 5, the displacement current at the gate electrode is still an indicator of the trapping processes in the insulator. Based on these measurements, ΔQ vs Δv_{FB} data is computed by the integration method as indicated in chapter 5. Steady-state currents vs gate bias curves are also presented to illustrate the effectiveness of

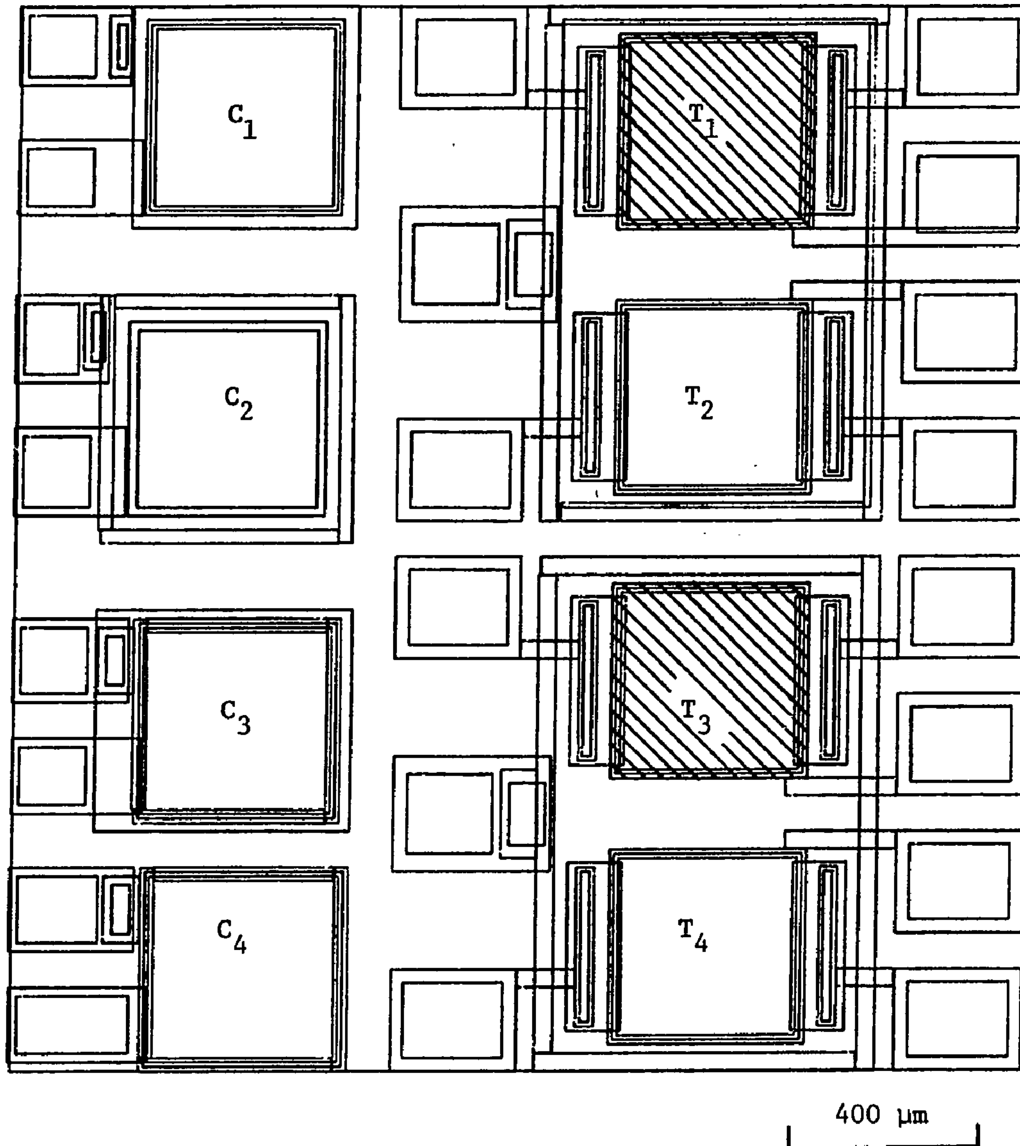


Fig.7.1 Various MNOS test structures designed for the present work. C_1 - C_4 represent MNOS capacitors on n^+, p^+, p and n type semiconductor, respectively. T_1-T_4 represent n-channel(buried) , n-channel (surface), p-channel(buried), p-channel(surface) transistors, respectively.

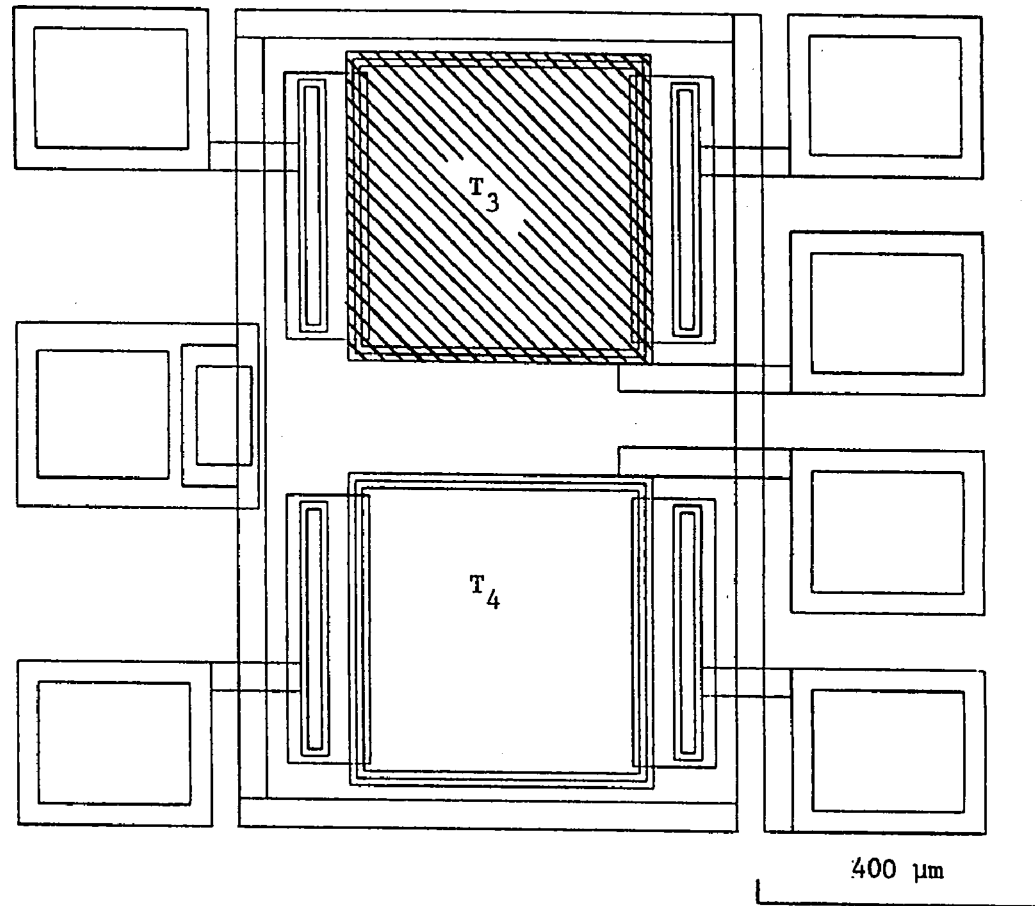


Fig. 7.2 Detailed picture of transistors T_3 (buried p-channel) and T_4 (surface p-channel).

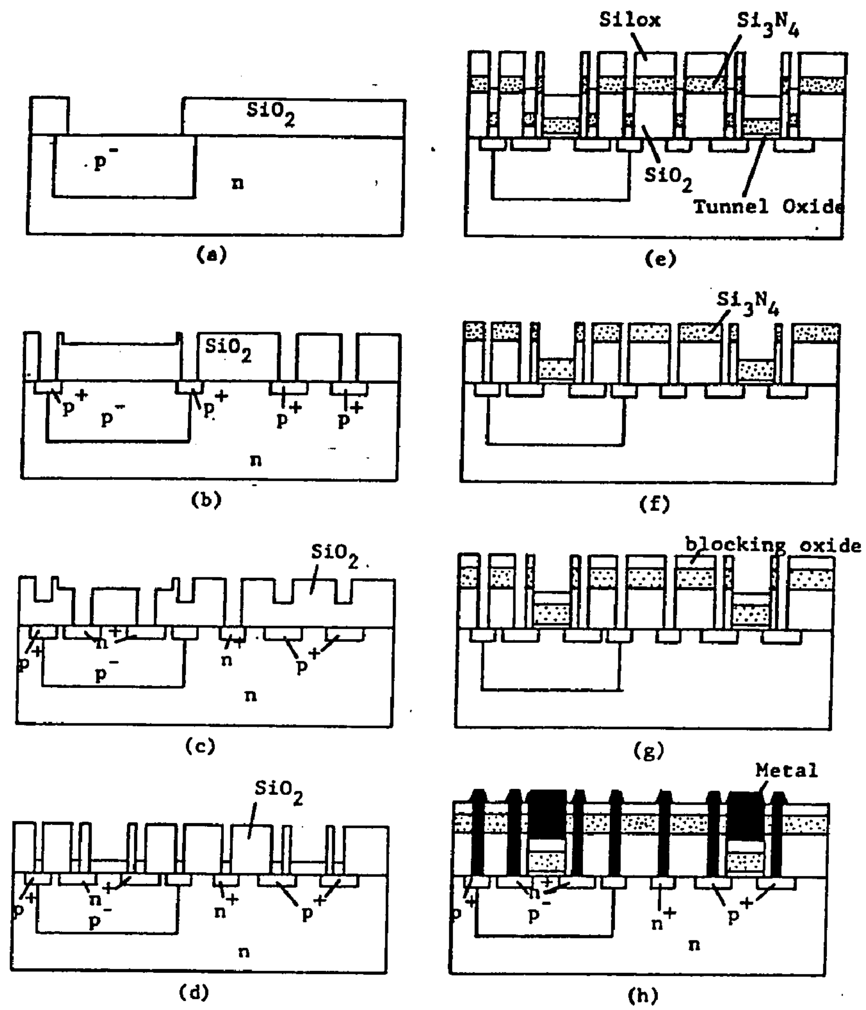


Fig. 7. 3 Schematic of the MONOS n and p-channel transistors at various stages in the fabrication sequence.

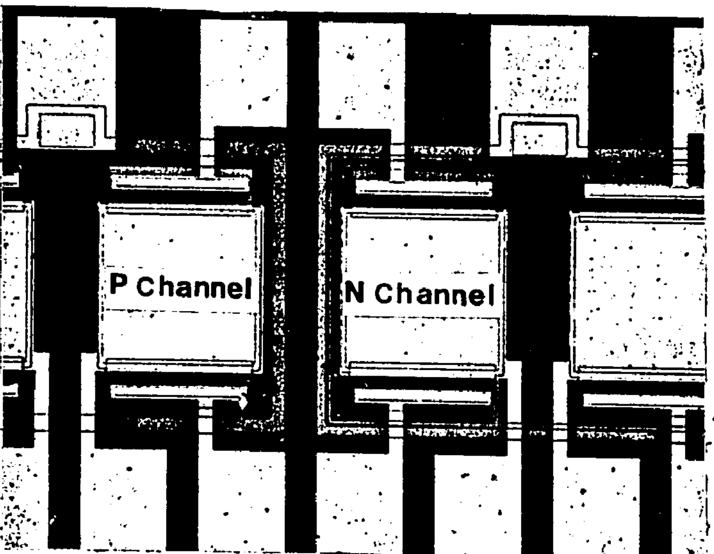


Fig. 7.4 The photomicrograph of the p and n-channel MONOS devices.

the blocking oxide.

7.2.1 The Measurement Setup

Figure 7-5 shows the block diagram of the computer-aided data-acquisition system used for measurements. It is based on the HP-9836 desktop computer with an IEEE 488 bus. The linear voltage ramp is supplied by the function generator in the form of a symmetrical triangular wave. The source and bulk currents are measured by the Keithley 616 electrometers. The analog outputs of the electrometers and the ramp voltage are sequentially sampled (5 ms interval) by the A/D converter. The device is mounted in a TO-5 header and cooled in the Janis Supervaritemp Cryostat with liquid nitrogen. It is very important to check the p/n junction (breakdown voltage etc) on a curve tracer before a device is used in ramp measurements. In addition, if the source to bulk junction is to be reverse biased, it should be done with a battery mounted in a Teflon holder to minimize the leakage. Device and the connectors should be kept in a dry ambient produced by the flow of dry grade nitrogen gas. Any humidity in the atmosphere produces noticeable leakage.

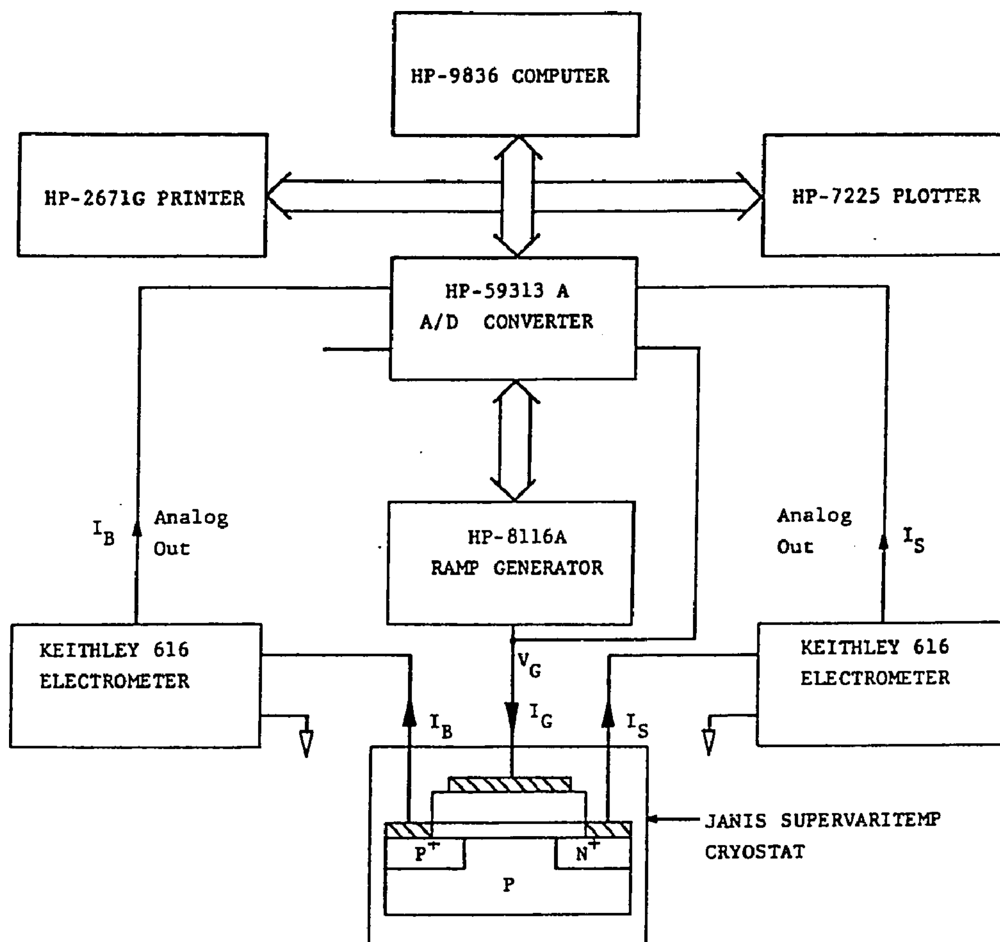


Fig. 7.5 The block diagram of the automatic data-acquisition system.

7.2.2 Results on the P-Channel Device

The p-channel MONOS devices were fabricated with the following parameters:

Substrate = <100> n type
Substrate resistivity = 5-10 Ωcm
Tunnel oxide thickness ≃ 10 A
nitride thickness = 530 A
Blocking oxide thickness = 75 A
Device area(memory window)=400 μmX378 μm=$1.512X10^{-3}cm^2$

Write Mode:

In this mode of operation, the device is first prepared by applying +20 V on the gate relative to the bulk for 1 min (saturated Erase). This results in the transfer of electrons from Si to the traps in nitride and holes from the traps to the Si valence band as shown in Fig (5-7a). The flatband voltage shifts to a positive value. Next, the device is swept from a small positive value of the gate bias (weak accumulation) to a large negative value (-10 to -20 V) so that the semiconductor is driven into strong inversion (Write Mode). This results in the tunneling of the holes from the semiconductor to the nitride and electrons from the traps in nitride to the semiconductor conduction band as illustrated in Fig(4-7b). The bulk current monitors the electron component and the source current measures the hole component. The flatband voltage shifts in the negative direction and may become negative if a sufficiently high negative value of the bias is used. The gate bias is then swept back to the initial positive value. A ramp rate α = 0.2 V/sec

has been used throughout this study.

A set of various current components is illustrated in Fig7-6. This set was obtained at room temperature with V_{SB} = 0 V. For the sake of clarity, I_B and I_S are shown separately in Fig 7-7. The flatband voltage shifts from an initial value of V_{FB1} = 6 V to the new value V_{FB2} = -8.0 V with an application of -20 V on the gate.

Reproducibility of the measurement is illustrated in Fig.7-8 where different values of the Writing voltage are used. The device is prepared by applying +20 V for 1 min before each Writing operation. The initial flatband voltage depends on the Writing voltage as expected. The more the Writing voltage the higher is the shift in the flatband voltage.

The temperature dependence of the charcteristics is shown in Fig.7-9. As expected, the minimum of the quasi-static curve is reduced at the low temperature. In addition, the displacement current at the gate in the charge injection region is reduced below that at room temperature. This can be attributed to the reduced penetration of the centroid of the injected charge due to the increase in capture cross-section of attractive traps at lower temperature as discussed in chapter 6 and illustrated in Fig.6-11(a).

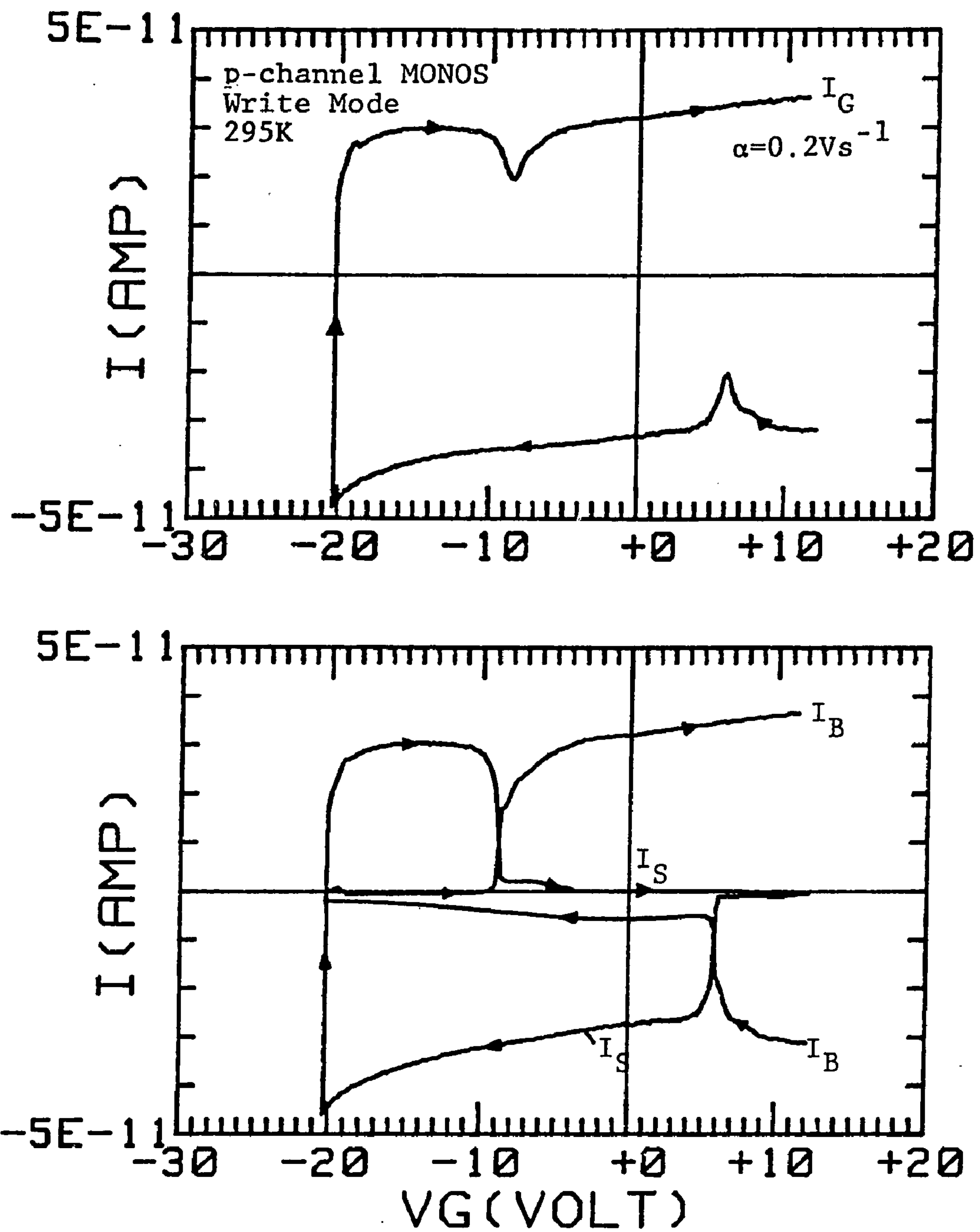


Fig. 7.6 Various current components in the Write mode of the p-channel MONOS structure. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A, Area=1.51x10^{-3}cm^2.

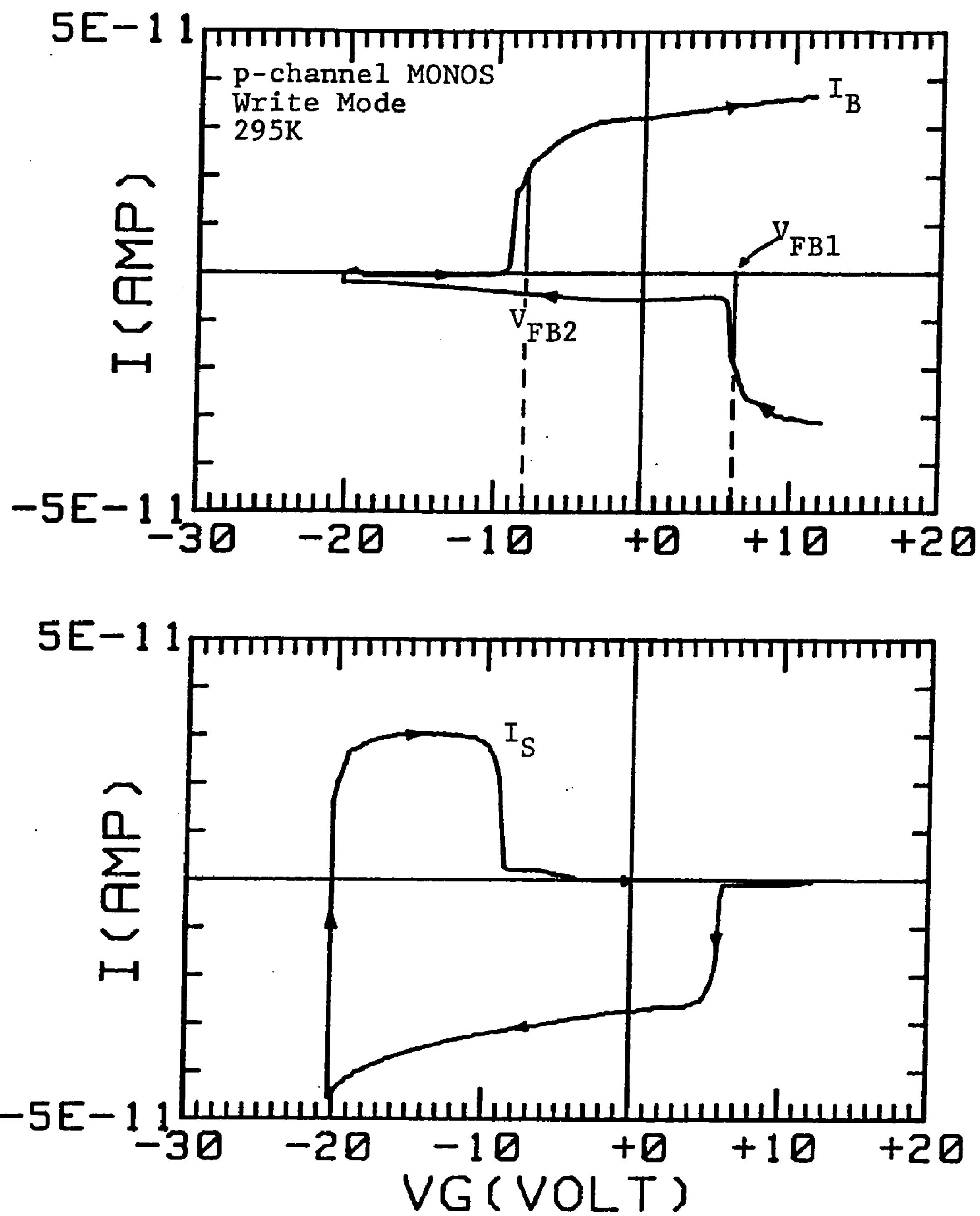


Fig. 7.7 Bulk and source current components of Fig.7.6 shown separately. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A, Area=$1.51\times10^{-3}cm^2$.

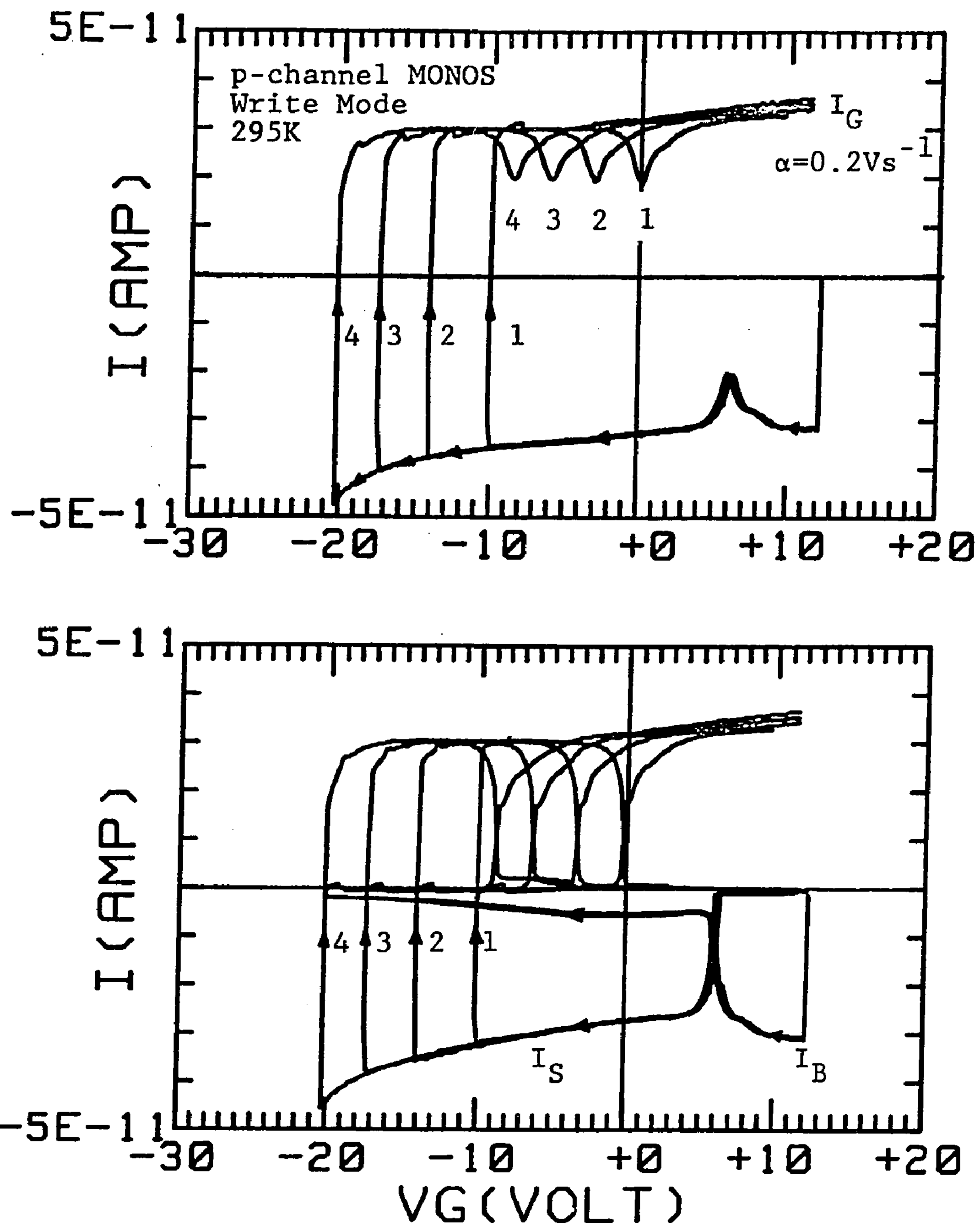


Fig. 7.8 Various current components with different magnitudes of the Writing voltage applied to the p-channel MONOS device. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A, Area=$1.51\times10^{-3}cm^2$.

The effectiveness of the blocking electrode is illustrated by the steady-state current characteristics of the sample for negative gate bias shown in Fig.7-10. In this measurement, -30 V is first applied to the gate electrode relative to bulk for 5 min and then the bias is reduced in steps of 200 mV. The steady-state current in the gate electrode is measured at the end of a small time period (5-10 sec) after the bias is reduced. This current is dominated by the holes injected from the inversion layer and collected by the gate as evidenced by the large source current component. A small amount of electrons are also collected by the semiconductor (monitored by the bulk current I_B). The Poole-Frenkel conduction through the nitride is significantly reduced at the low temperature (100 K). It should be noticed that the steady-state current at the gate bias of -20 V or less is much below the displacement current observed in the ramp measurements. Thus, any injection from the gate electrode can be safely negelected in the analysis of results of the ramp measurements.

The current-voltage curves are integrated according to the Fig.4-10. The resulting ΔQ vs ΔV_{FB} data are shown in Fig.7-11 at room temperature and low temperature. ΔQ_p represents the hole charge per unit device area injected from the semiconductor into the nitride film whereas ΔQ_n represents the electron charge per unit device area released from the nitride film and collected by the semiconductor. It is quite clear that hole injection from the

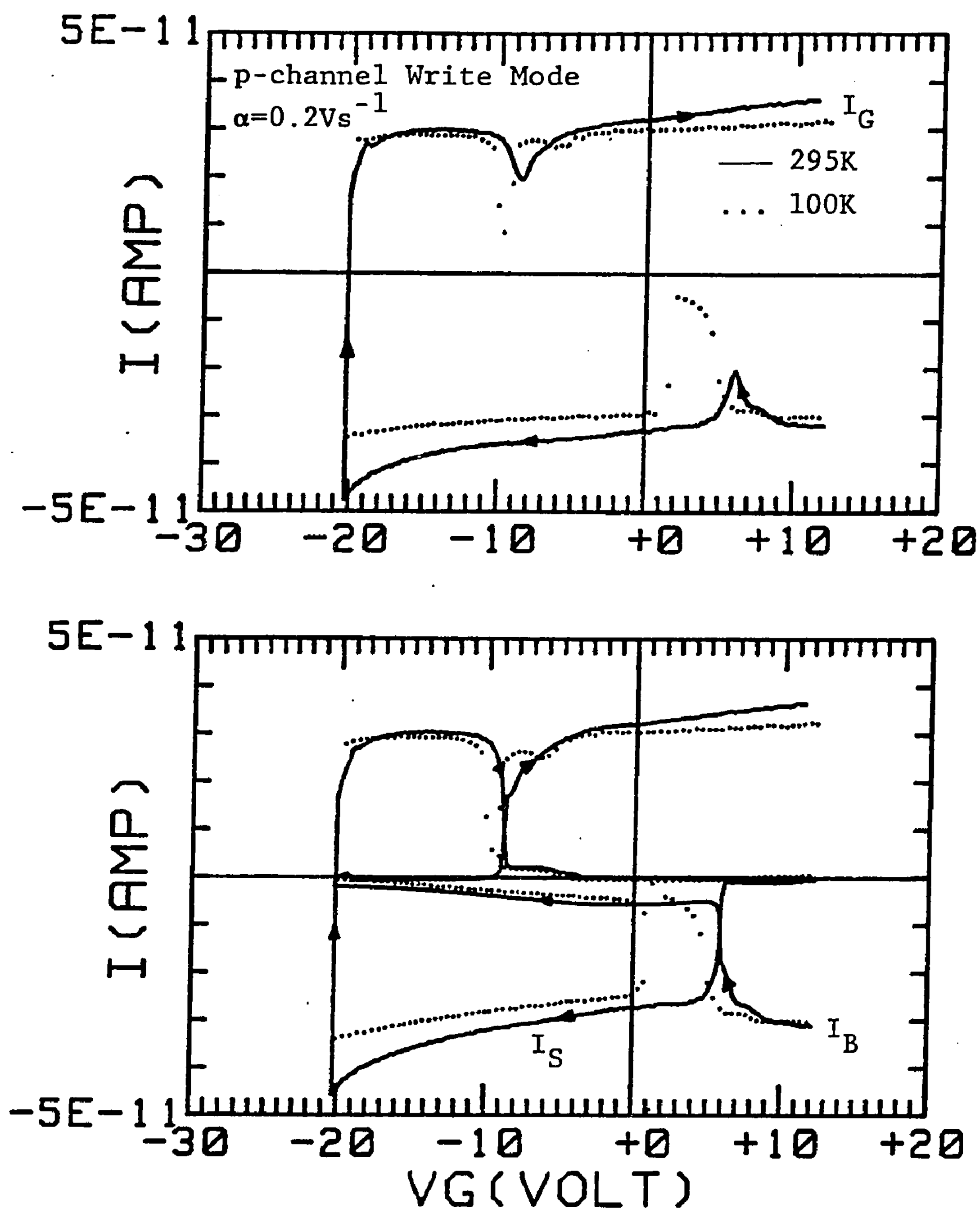


Fig. 7.9 Various current components at two different temperatures for the p-channel MONOS device. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A and area=$1.51\times10^{-3}cm^2$.

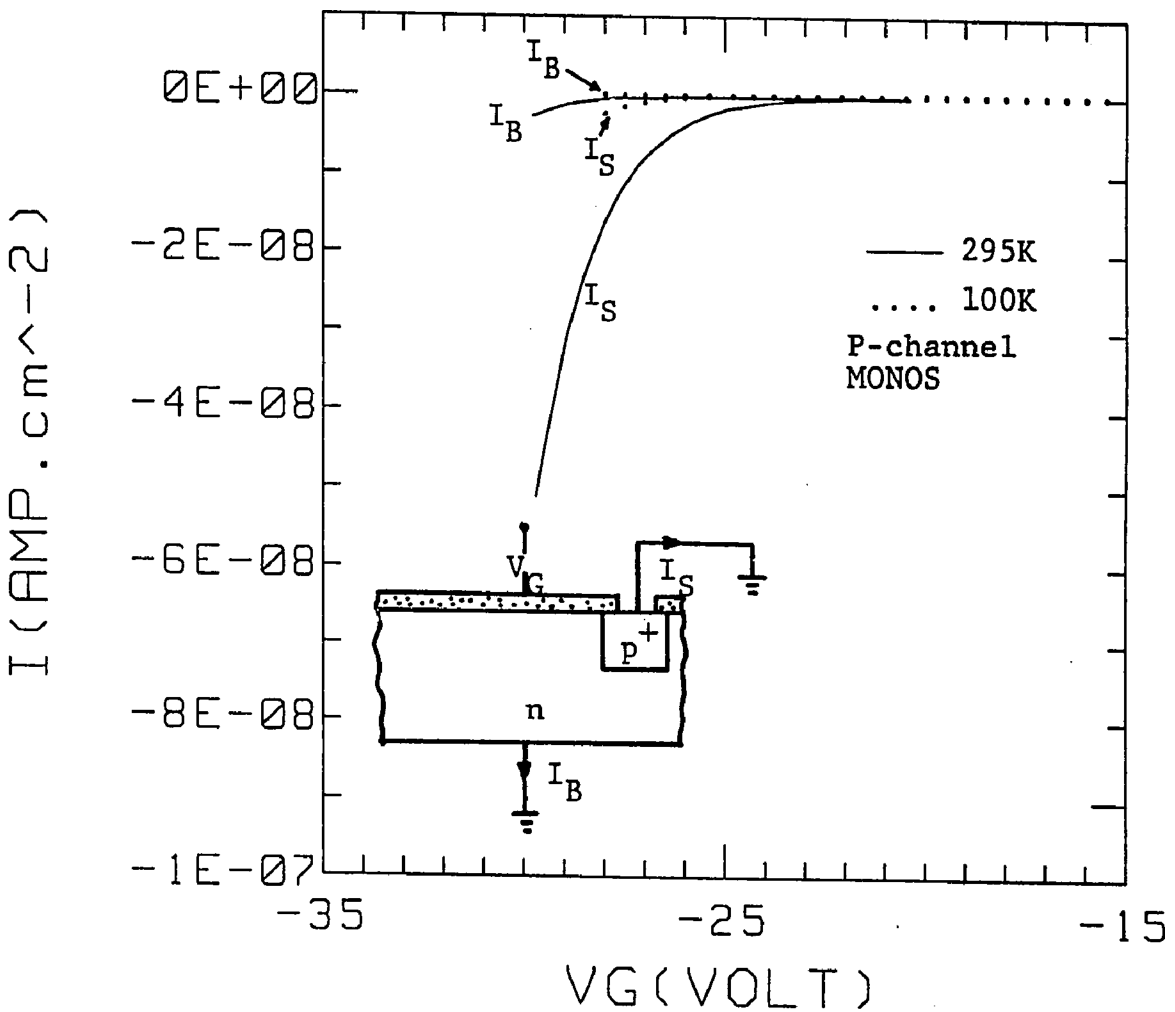


Fig. 7.10 Steady-state conduction in the p-channel MONOS sandwich at two different temperatures.

semiconductor is the dominant mechanism for Writing of the p-channel MONOS device. The temperature dependence is also seen to be in agreement with that predicted in Fig.6-11(c) based on the temperature dependence of the capture cross-section of an attractive center. ΔQ_p vs ΔV_{FB} data at room temperature shows the upward bend characteristic of centroid movement away from the injecting boundary (tunnel oxide-nitride interface).

Erase Mode:

In this mode of operation, the p-channel device is first prepared by applying -20 V on the gate relative to the bulk for 1 min (saturated Write). This results in the transfer of holes from Si to the traps in nitride and electrons from traps to the Si conduction band as shown in Fig.5-7b. The flatband shifts to a negative value. Next, the device is swept from a small negative value of the gate bias (strong inversion) to a large positive value (+10 to +20 V) so that the semiconductor is driven into strong accumulation (Erase Mode). This results in the tunneling of electrons from the semiconductor to the nitride and holes from the traps in nitride to the semiconductor valence band as shown in Fig.5-7a. The bulk current monitors the electron component and the source current measures the hole component. The flatband voltage shifts in the positive direction and may become positive if sufficiently high positive value of the bias is used. The gate bias is then swept back to the initial negative value. Again, the ramp

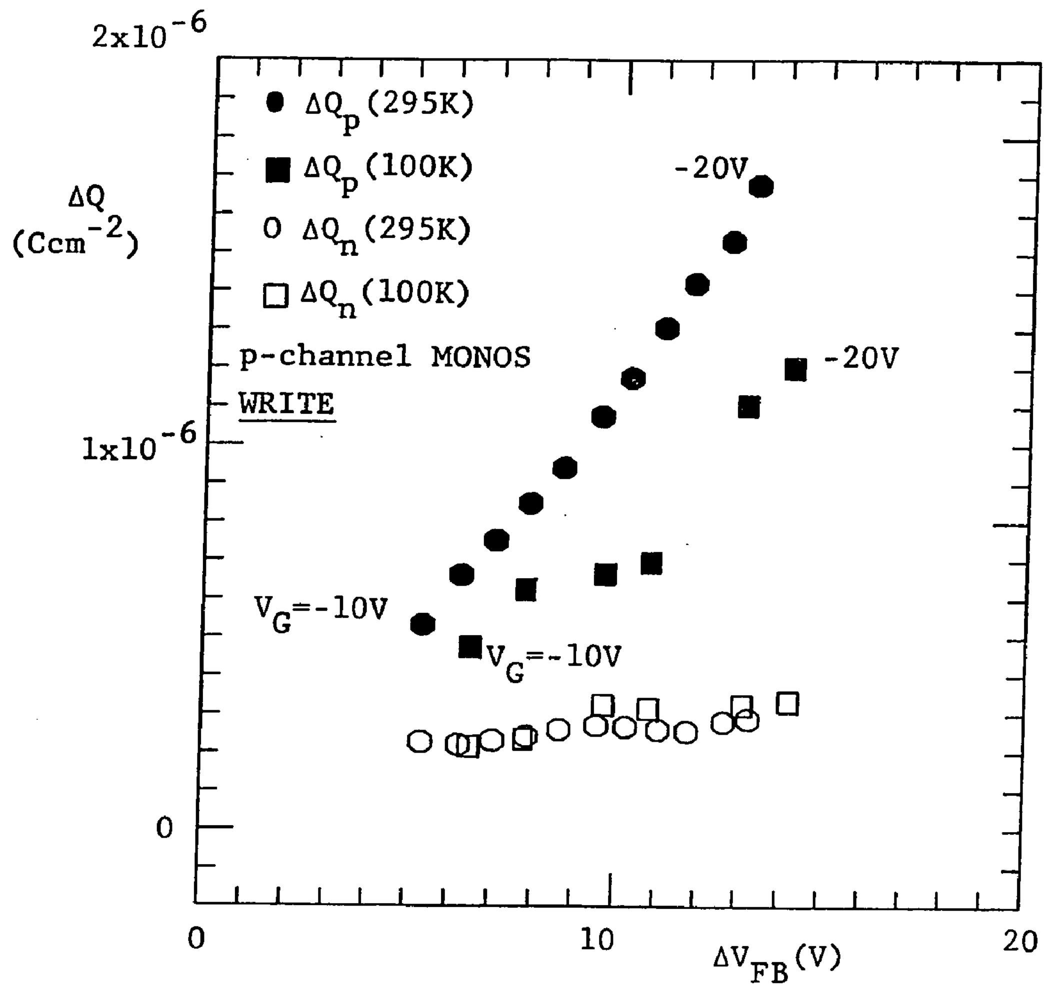


Fig. 7.11 ΔQ vs ΔV_{FB} data extracted from the ramp measurements on the p-channel MONOS device at two different temperatures in the Write mode. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A, and Area=1.51x10^{-3}cm^2.

rate $\alpha = 0.2$ V/sec was used throughout this study.

A set of various components in the Erase Mode are shown in Fig.7-12. It was obtained at room temperature with $V_{SB}=0$ V. For the sake of clarity, I_B and I_S are shown separately in Fig.7-13. The flatband voltage shifts from an initial value of $V_{FB1} = -9$ V to the new value $V_{FB2} = 5$ V with an application of +20 V on the gate.

Reproducibility of the measurement is illustrated in Fig.7-14 where different values of the Erasing voltage are used. The device is prepared by applying -20 V for 1 min on the gate relative to bulk before each sweep. It can be seen that for Erasing voltages of 10V to 14V, flatband voltage shifts in the positive direction but still remains negative. For Erasing voltages in excess of 14V, flatband voltage indeed becomes positive.

The temperature dependence of the Erase characteristic is shown in Fig.7-15. Again, the displacement current at the gate in the charge injection region is reduced below that at room temperature. As before, this is attributed to the increase in capture cross-section of the attractive trap at lower temperature as discussed in chapter 3. The effect of temperature is also manifested in ΔQ vs ΔV_{FB} characteristics in Erase mode [Fig.7-16]. The separation of carriers is not conclusive in the Erase mode because of high recombination probability of the holes released from the nitride

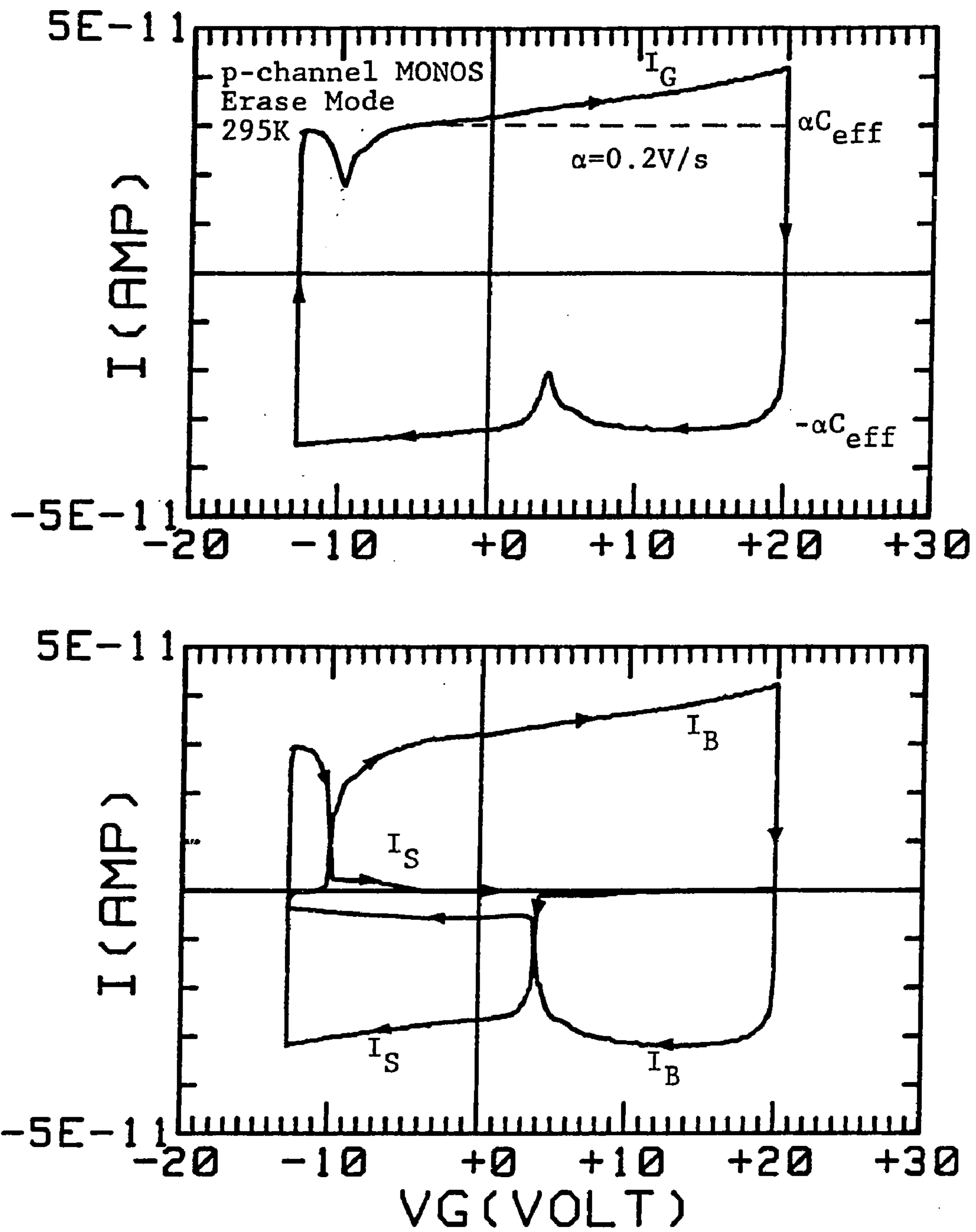


Fig. 7.12 Various current components in the Erase mode of operation of the p-channel MONOS structure. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A and Area=$1.51 \times 10^{-3} cm^2$.

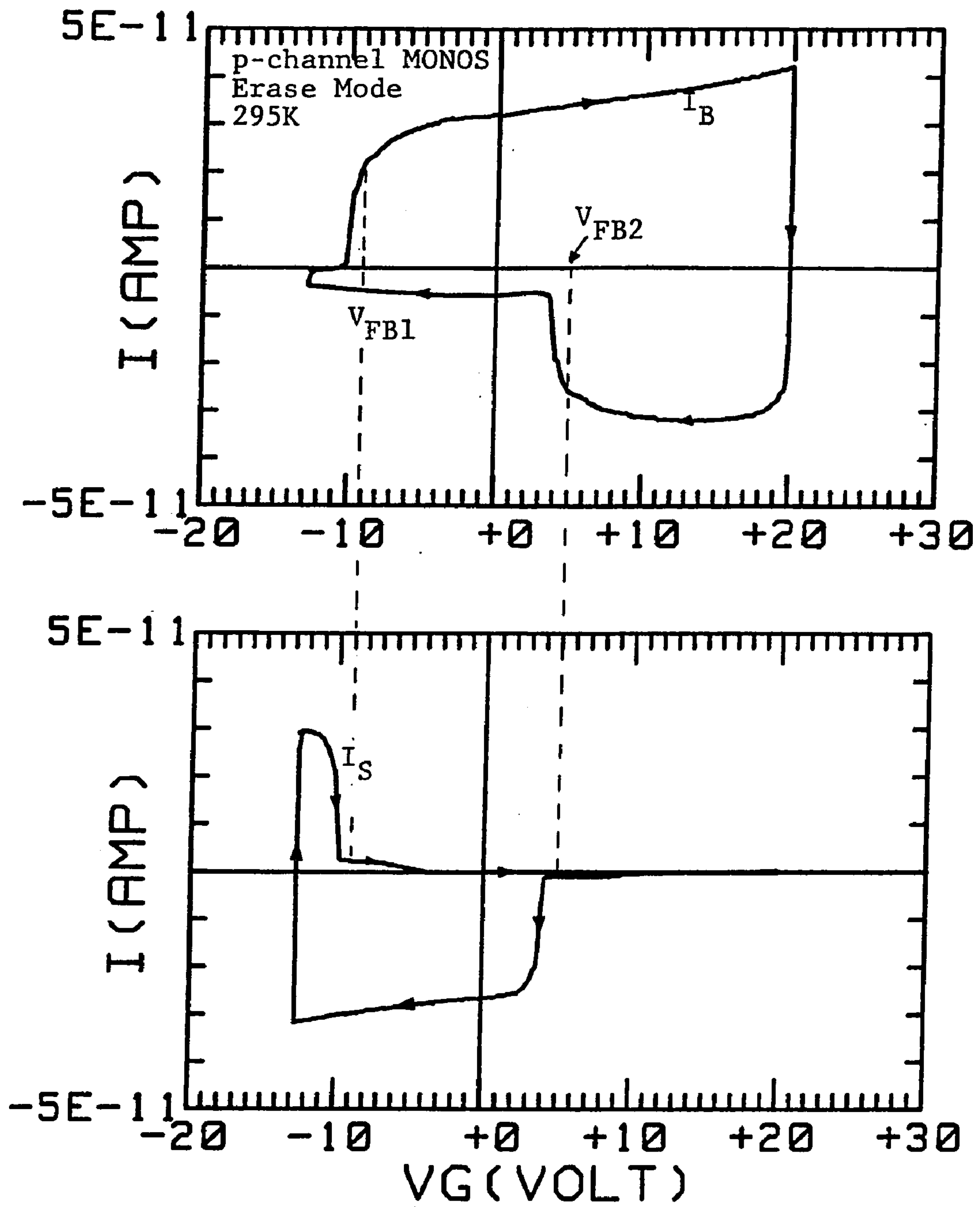


Fig. 7.13 Bulk and Source current components of the Fig.7.12 shown separately. The device dimensions are: X_{oxt}= 10A, X_N=530A, X_{oxB}=75A, and Area=1.51x10^{-3}cm^2.

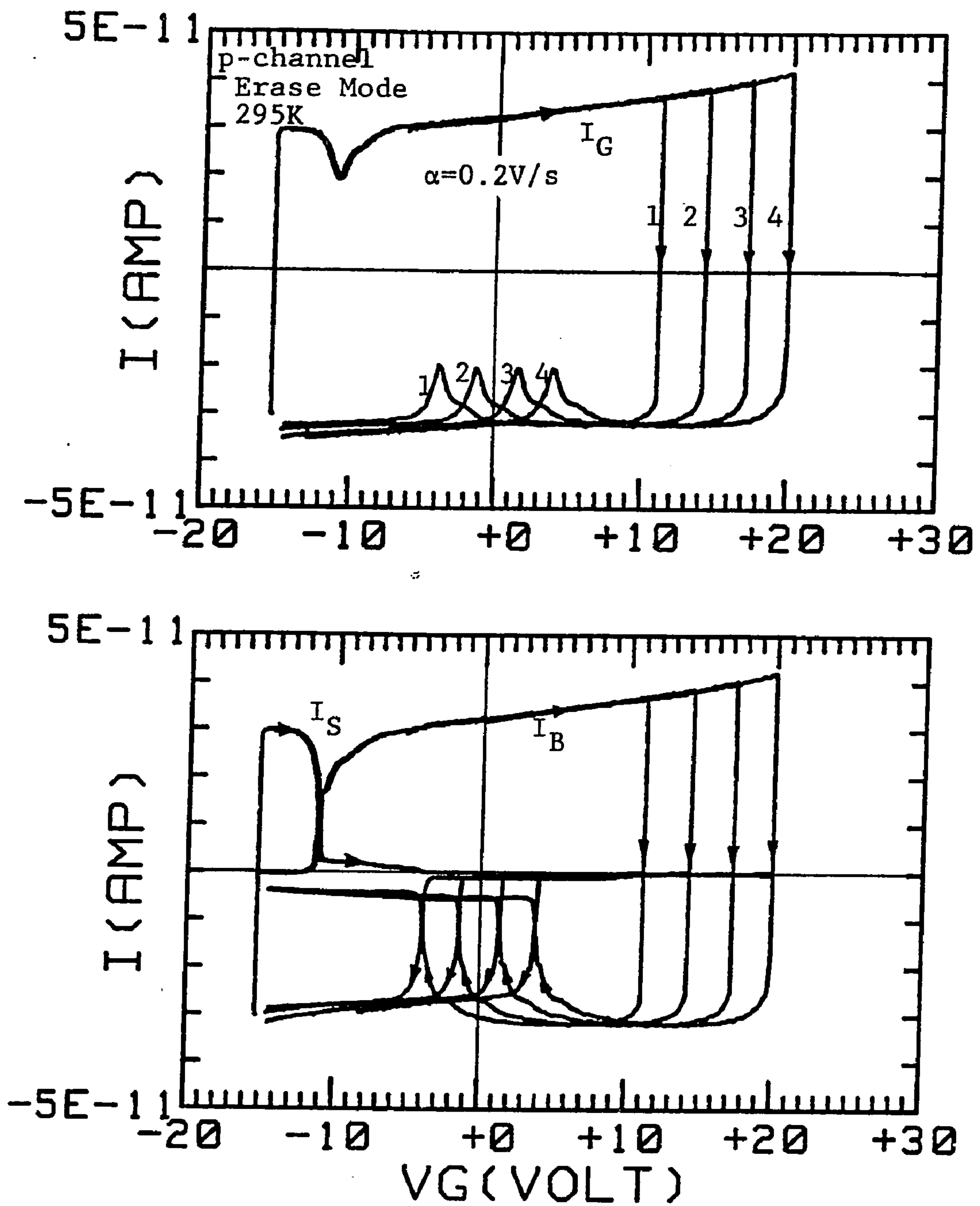


Fig. 7.14 Various current components with different erasing voltages applied to the p-channel MONOS device. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A and Area=1.51×10^{-3} cm^2.

before they can be collected by the source region.

The gated-diode leakage current data for two different values of V_{SB} is shown in Fig.7-17. In this measurement, gate bias is changed in small steps. The source and bulk currents are measured at the end of a small period (5-10 sec) after the step. Since in steady-state, gate current should be zero, it follows that $I_G=I_B+I_S=0$. This is also observed experimentally. The current due to the surface generation is about 30 $nA.cm^{-2}$ corresponding to the surface recombination velocity S_0 ($I_{gen} = qn_iS_0$) of 13 $cm.sec^{-1}$.

7.2.3 Results on the N-Channel MNOS Device

The n-channel device fabricated within the p^- well on the n type substrate with CMOS sequence could not be used because of the floating p-n junction at the back. In this case, relatively small geometry, n-channel devices fabricated on p type substrate were obtained from another source. The photo-micrograph of the device and schematic of its cross-section are shown in Fig.7-18. No blocking layer of oxide was intentionally put down on the nitride. The following parameters were estimated:

Substrate = <100> p type
Substrate resistivity = 1 ohmcm
Tunnel oxide thickness = 20-25 A
nitride thickness = 500A
Device area (memory window) = 0.833×10^{-4} cm^2.

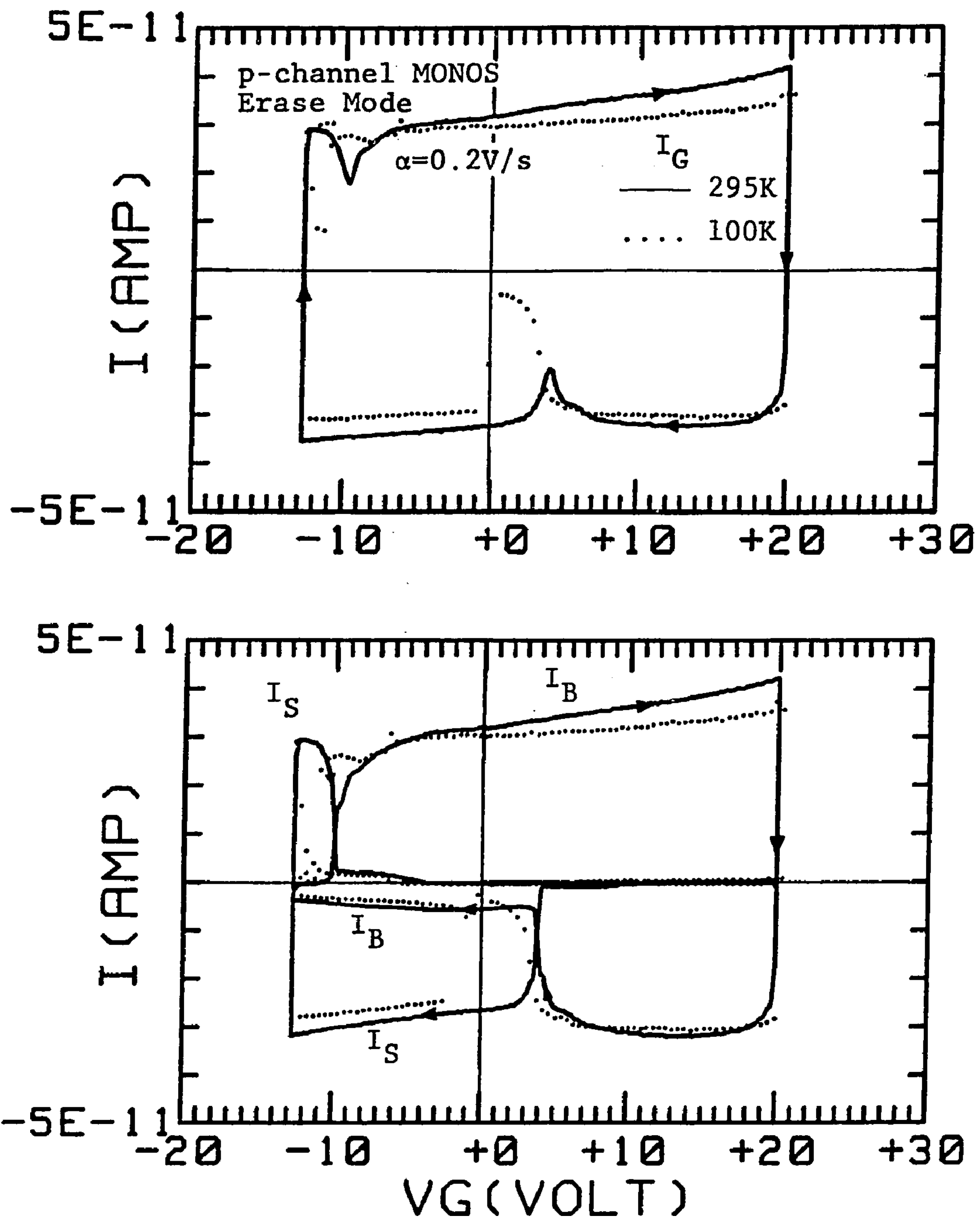


Fig. 7.15 Various current components in the erase mode of the p-channel MONOS device at two different temperatures. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A, and Area=$1.51x10^{-3}cm^2$.

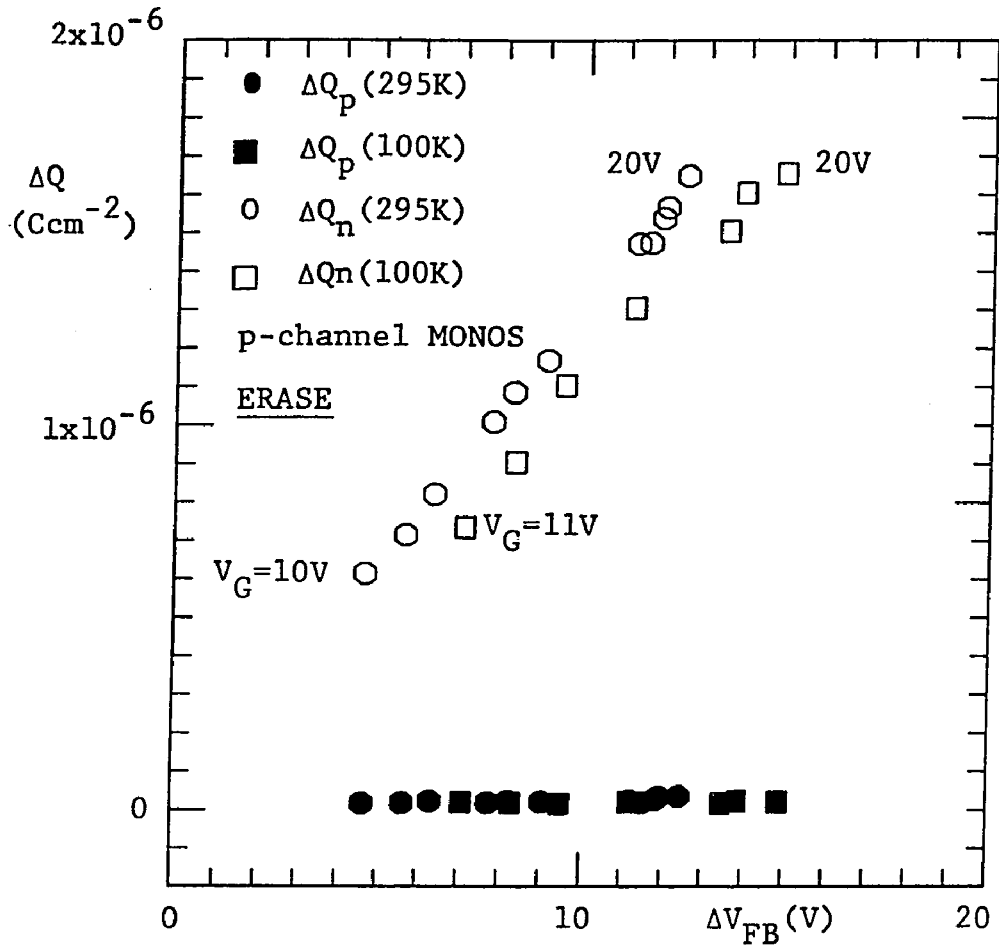


Fig. 7.16 ΔQ vs ΔV_{FB} data extracted from the ramp measurements on the p-channel MONOS device at two different temperatures. The device dimensions are: X_{oxt}=10A, X_N=530A, X_{oxB}=75A, and Area=$1.51\text{x}10^{-3}$ cm^2.

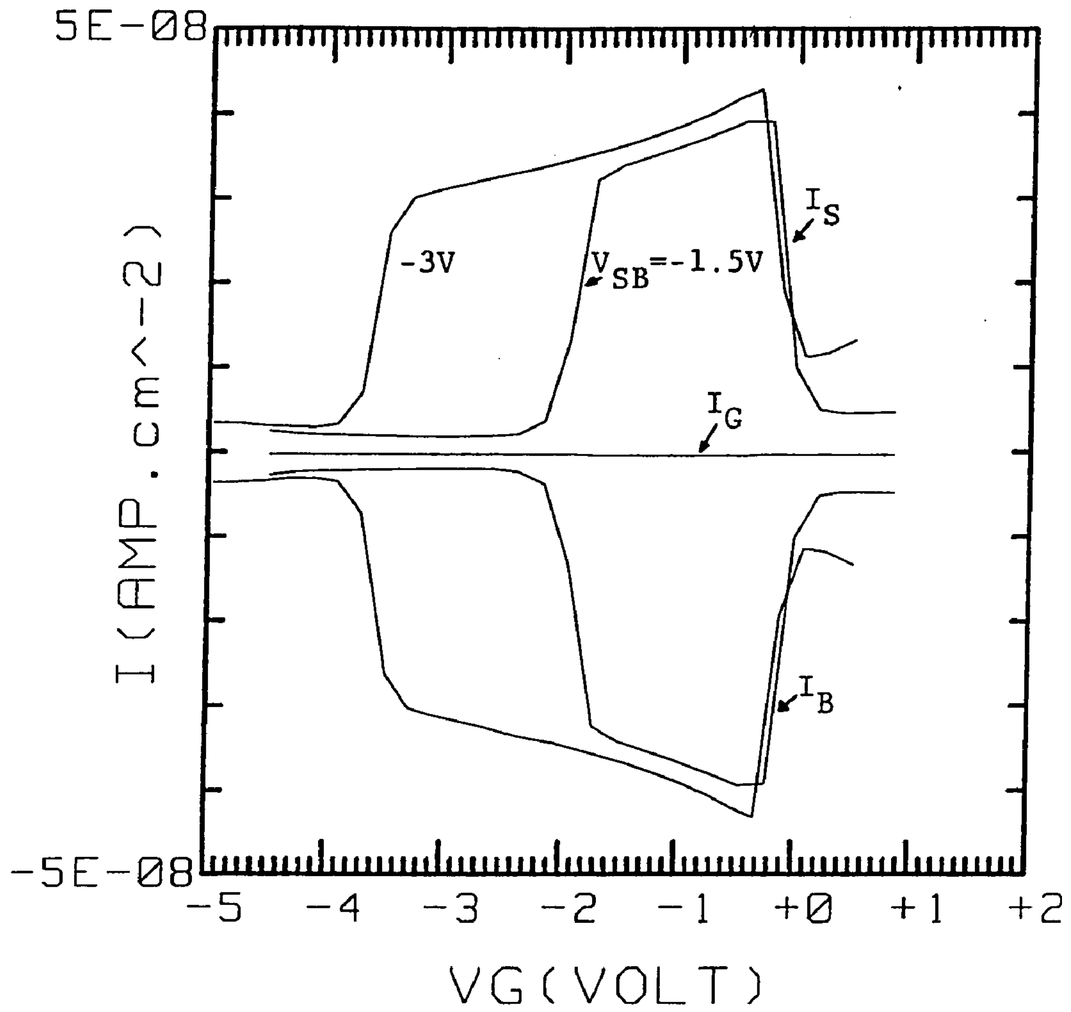


Fig. 7.17 Gated-diode leakage current measurement on the p-channel MONOS device for two different values of the source to bulk bias, V_{SB}. The device area=1.51×10^{-3} cm^2. See inset of Fig.7.10 for the current directions.

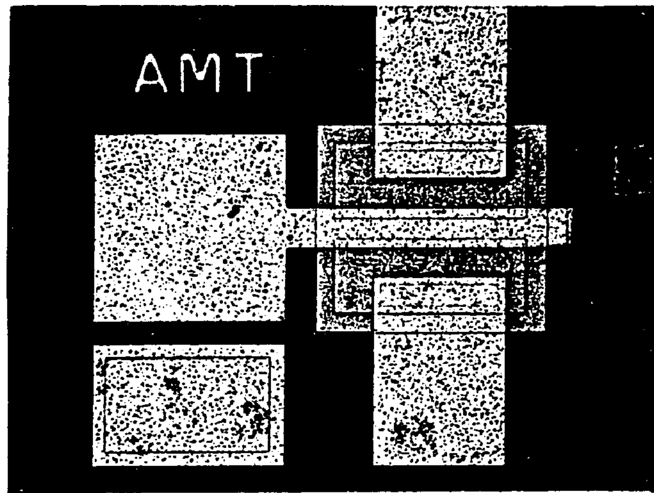


Fig. 7.18 The photomicrograph of the n-channel MNOS device.
By courtesy of Dr. Richard Withers, Lincoln Laboratory.

Write Mode:

Write mode of the n-channel device is similar to the Erase mode of the p-channel device. In this case, the device was prepared by applying -25 V on the gate relative to bulk for 1 min. This shifts the flatband to a negative value. Next, the device is swept from a small nagative value of the gate bias (weak accumulation) to a large positive value and back. The source current monitors the electrons injected from the semiconductor into the nitride layer while the bulk current monitors the electrons released from the nitride layer.

A typical set of curves is shown in Fig.7-19. Non-zero value of the bulk current in the inversion region ($V_G>0$) is due to the pad capacitance, C_{pad} shown in Fig. 5.6. The non-zero value of the source current, $\alpha C_{overlap}$, is due to overlap capacitance, $C_{overlap}$, shown in Fig.5.6. After substracting these parasitics, the various components are redrawn in Fig.7-20. Since the tunneling oxide is relatively thick (20-25 A), a distinct threshold is noticeable in the I-V curves. The bulk and source current components are shown separately in Fig.7-21. Figure7-22 illustrates the Write mode with different magnitudes of the Writing voltage. Figure7-23 compares the various current components at two different temperatures. Again we find that the displacement current at the gate is reduced at lower temperature indicating the reduced penetration of the charge centroid due to an increase in the capture cross-section.

The steady-state conduction in the structure at two different temperatures is illustrated in Fig. 7-24. The conduction is clearly dominated by the source current indicating electron injection from the inversion layer and transport through the insulating film. The steady-state source current is quite negligible up to +30 V and can not be responsible for high non-steady-state value, shown in Fig. 7-23, which can only be due to the trapping of electrons injected from the semiconductor inversion layer into the silicon nitride layer. This is also evident from the ΔQ vs ΔV_{FB} data shown in Fig. 7-25. The temperature dependence of this data also indicates reduced movement of the charge centroid at the lower temperature of 100 K.

Erase Mode:

The Erase mode of the n-channel device is similar to the Write mode of the p-channel device. The device is first prepared by applying +25 V on the gate relative to bulk for 1 min and then the gate bias is swept from strong inversion (+15 V) to large negative bias (accumulation) and back. The flatband voltage shifts in the negative direction. Bulk current monitors the hole flux injected from the semiconductor to the insulator whereas the source current measures the electron flux released from the insulator into the semiconductor. Although the charge separation is not efficient because of the high recombination probability of the electrons and holes in the semiconductor accumulation layer, the gate current

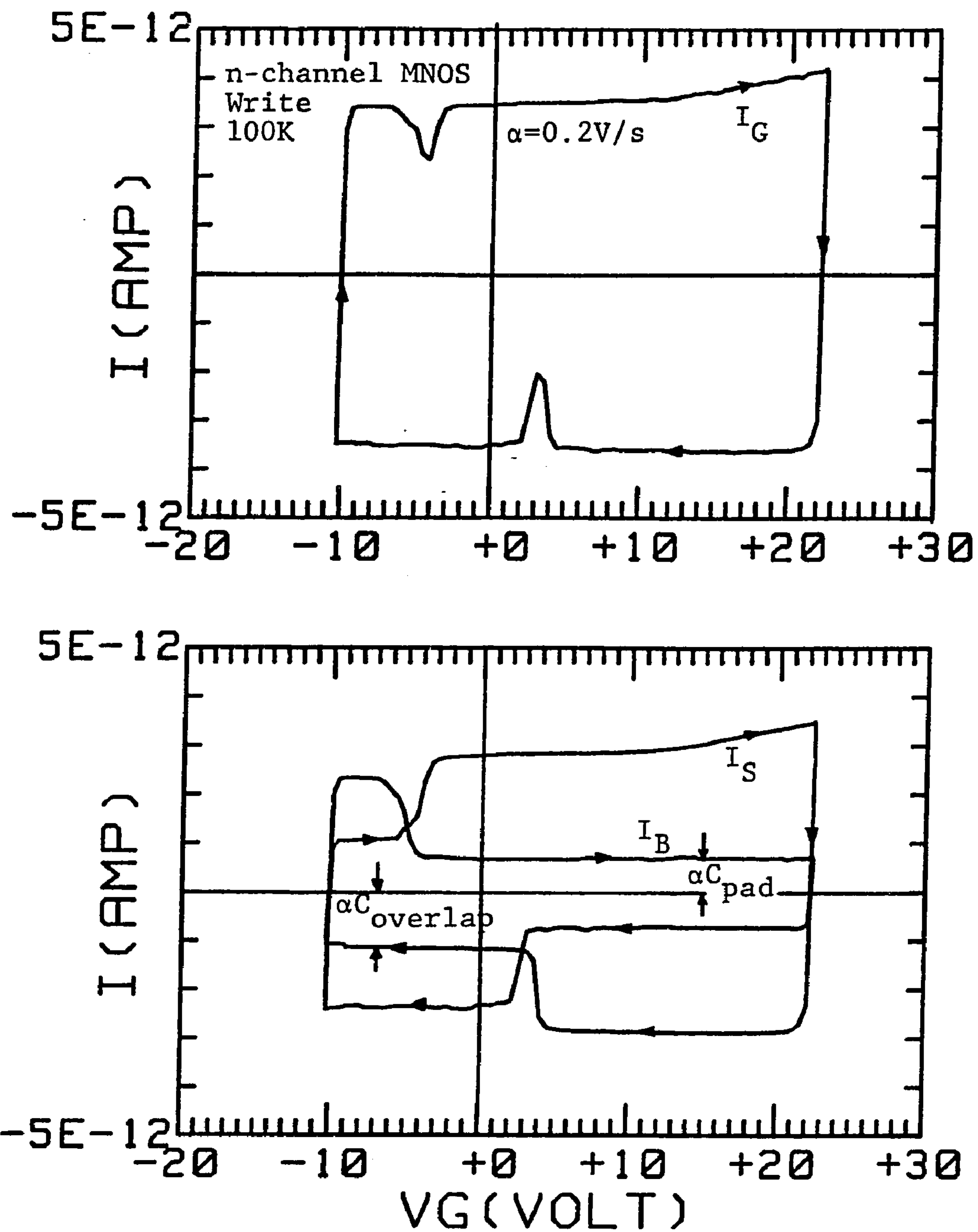


Fig. 7.19 Various current components in the Write mode of operation of the n-channel MNOS structure. The device dimensions are: X_{oxt}=20A, X_N=500A,and Area= $0.833x10^{-4}cm^2$.

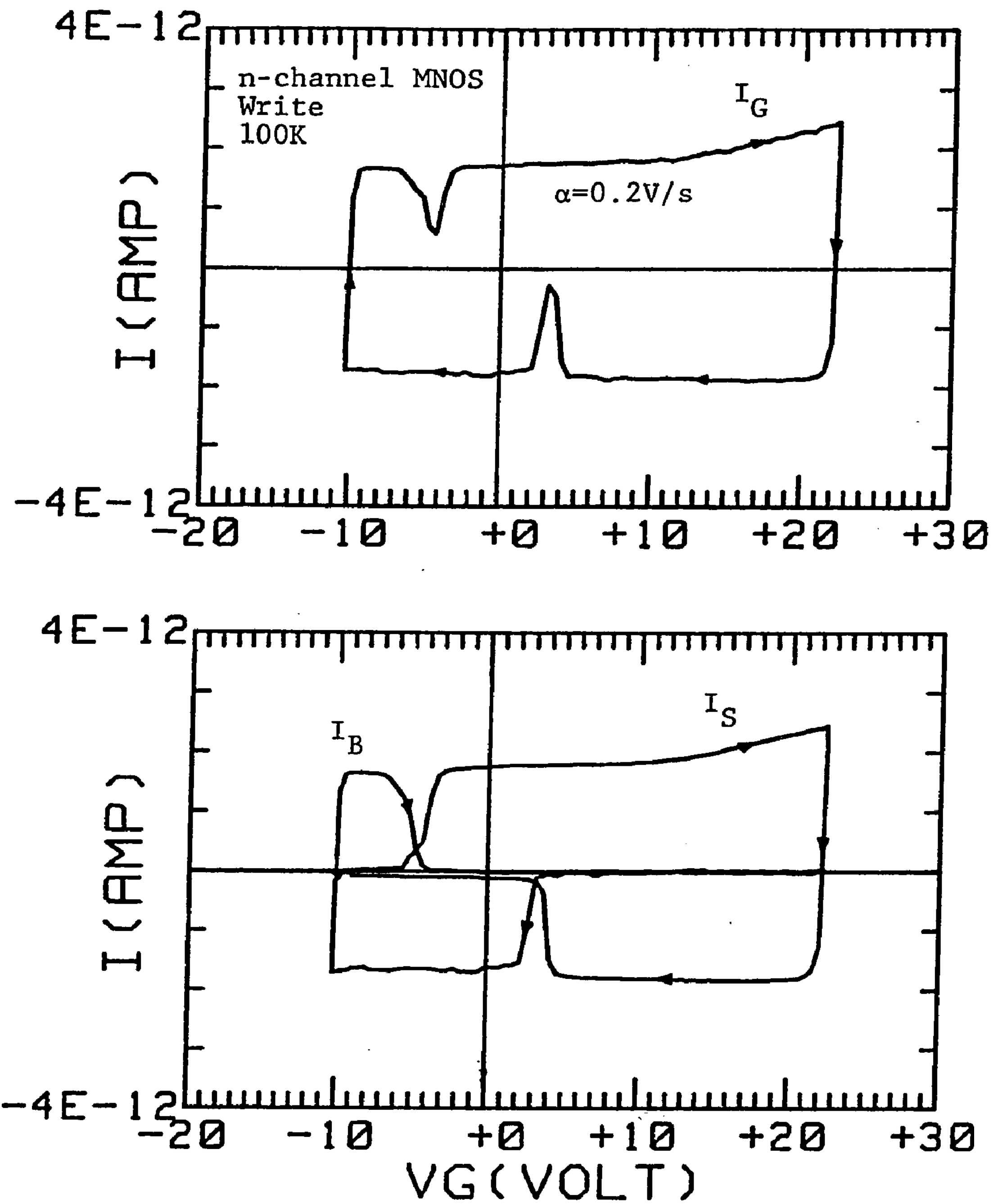


Fig. 7.20 Various current components of Fig. 7.19 after substracting the parasitics. The device dimensions are: X_{oxt}=20A, X_N=500A, and Area=0.833x$10^{-4}$$cm^2$.

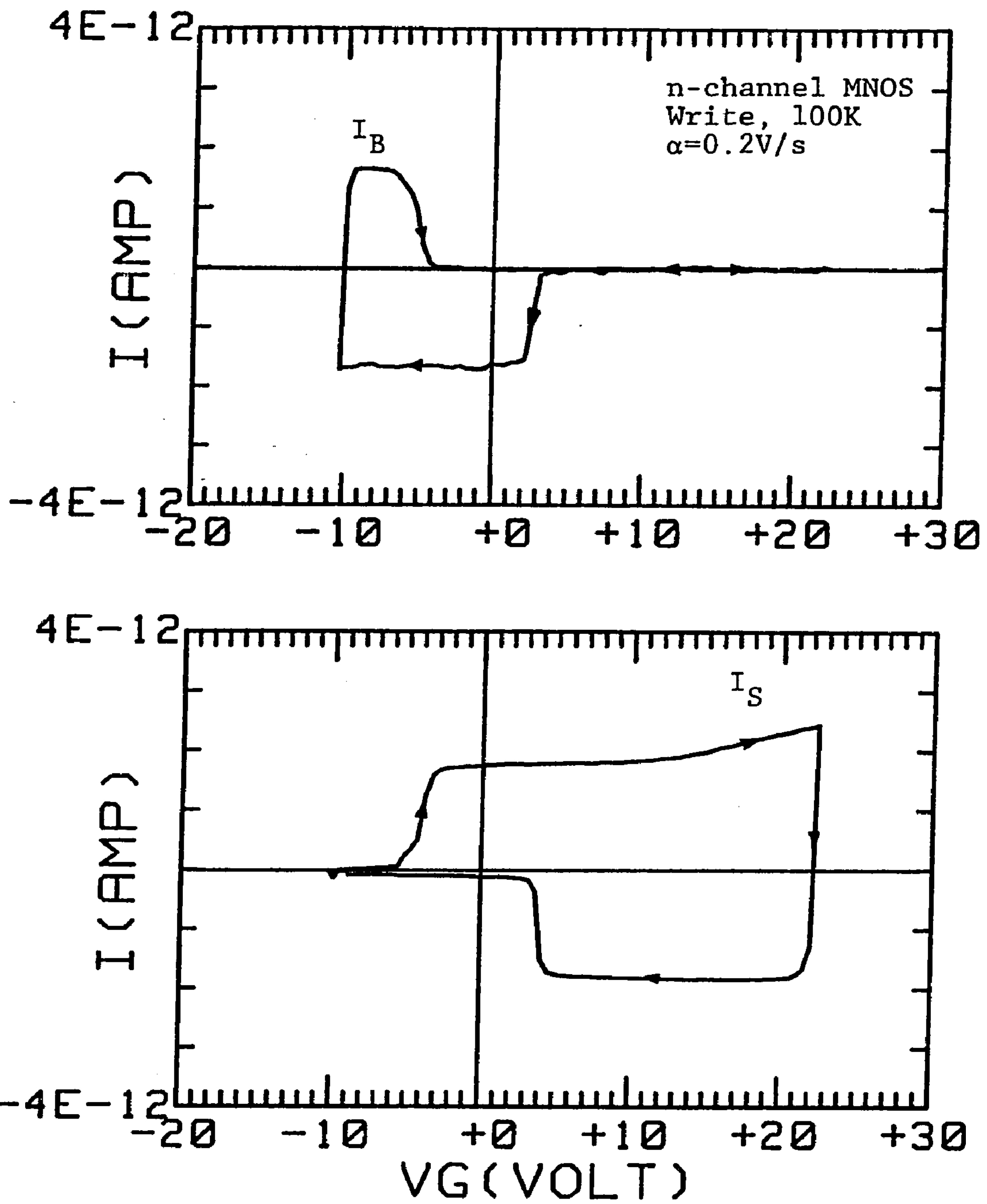


Fig. 7.21 Bulk and Source current components of Fig. 7.20 shown separately. The device dimensions are: X_N= 500A, X_{oxt}=20A, and Area=0.833x10^{-4}cm^2.

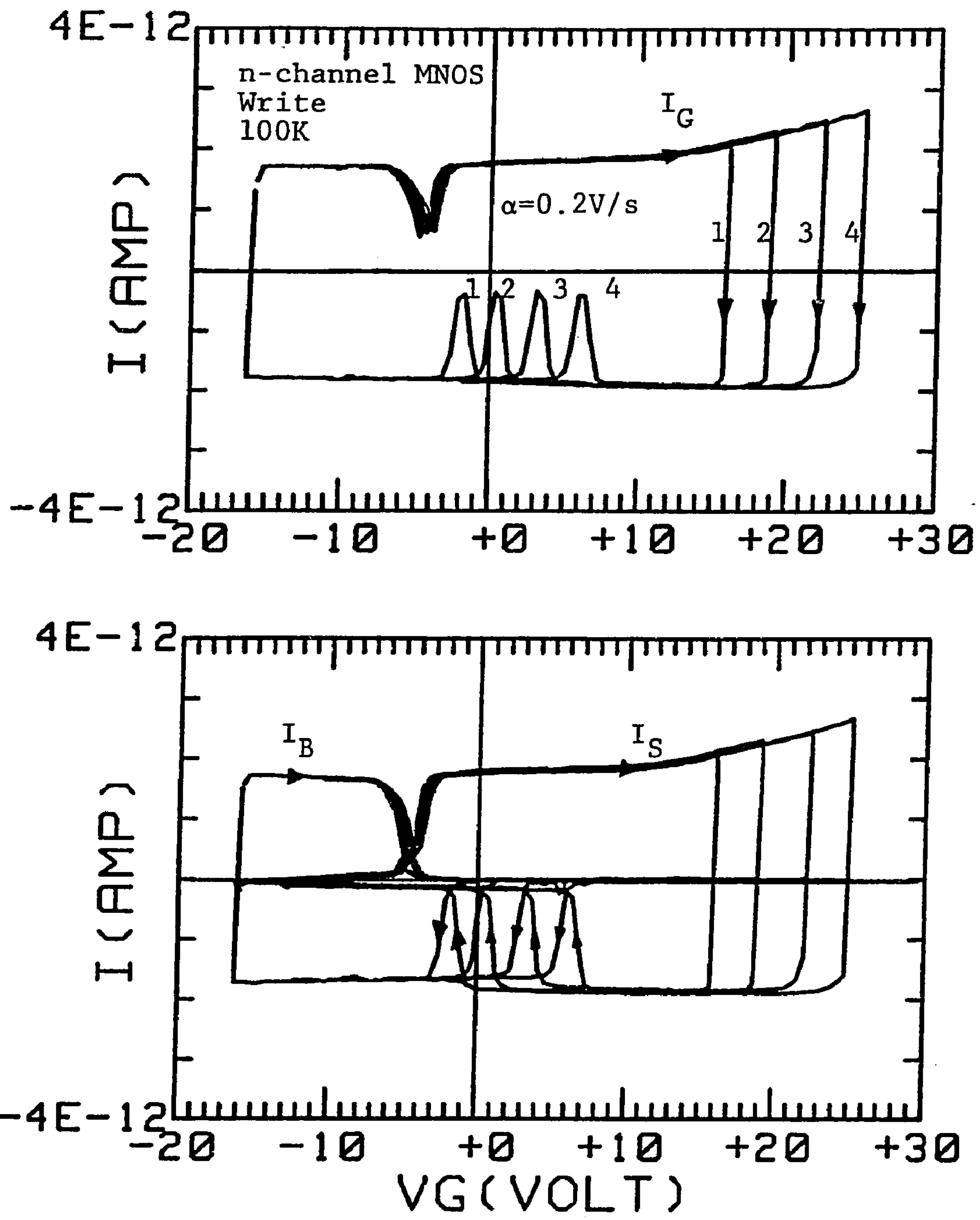


Fig. 7.22 Various current components with different writing voltages applied to the n-channel MNOS device. The device dimensions are: X_{oxt}=20A, X_N=500A, and Area=$0.833 \times 10^{-4} cm^2$.

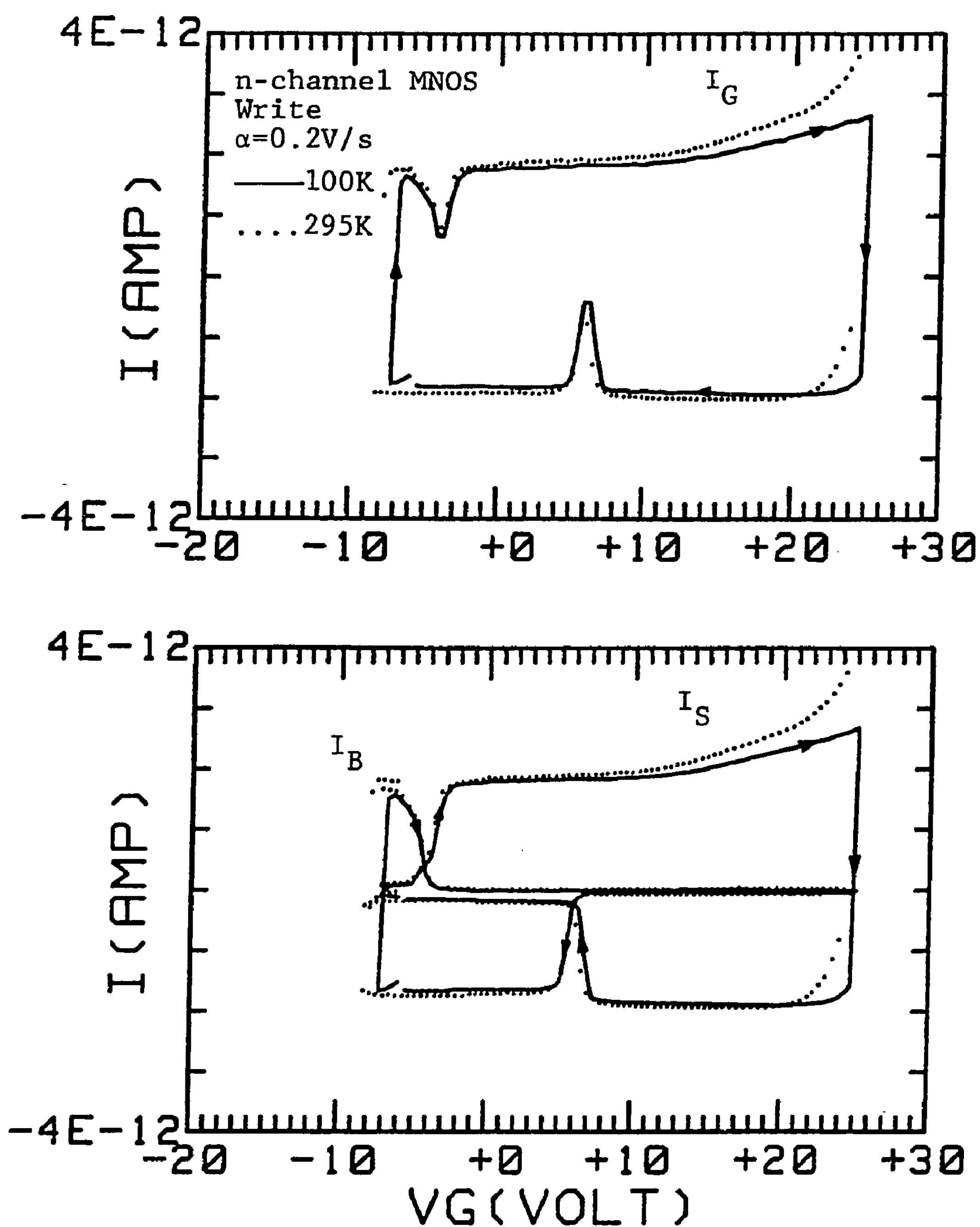


Fig. 7.23 Various current components in the n-channel MNOS device (write mode) at two different temperatures. The device dimensions are: X_{oxt}=20A, X_N=500A, and Area=0.833 $x10^{-4}cm^2$.

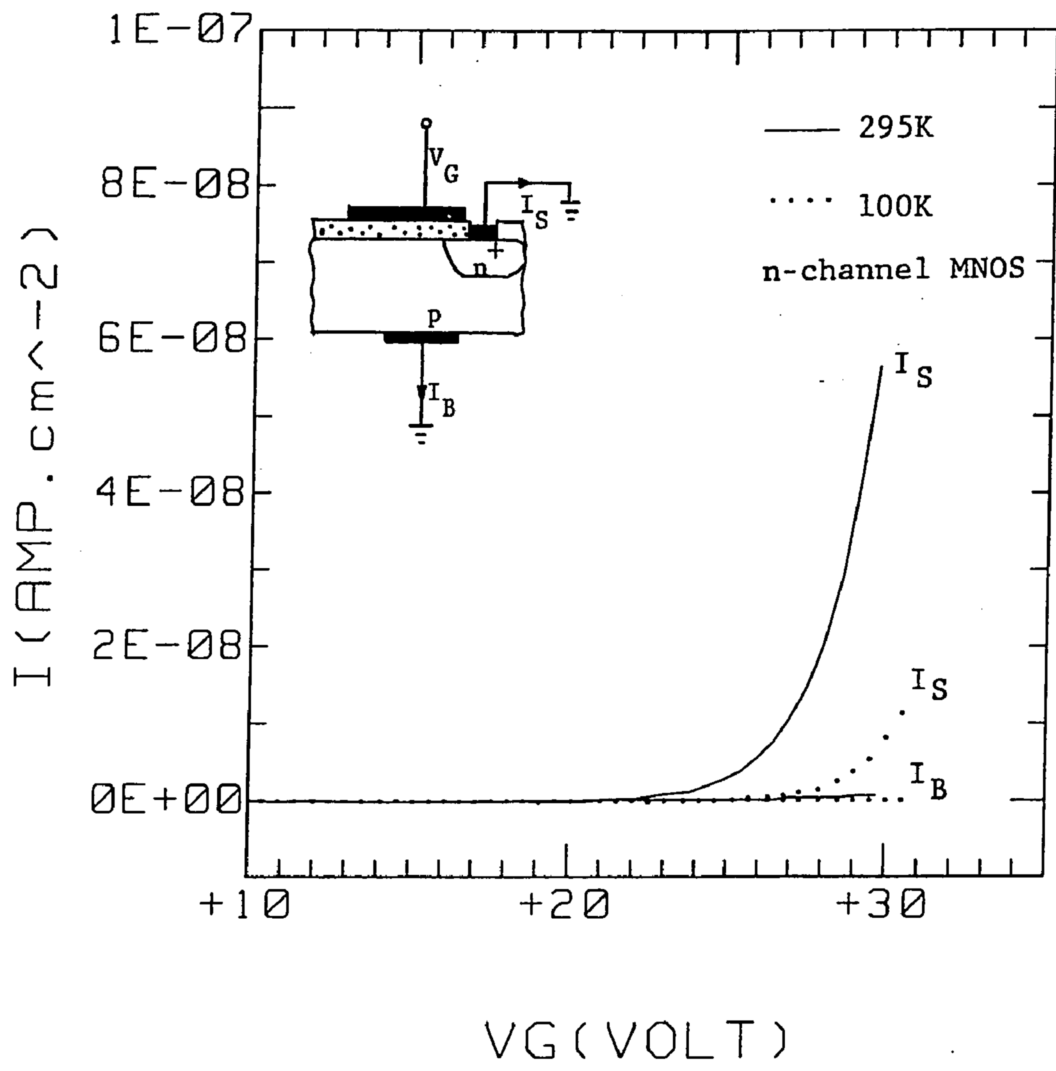


Fig. 7.24 Steady-state conduction in the n-channel MNOS structure at two different temperatures. The device dimensions are: X_{oxt}=20A, X_N=500A, and Area=0.833×10^{-4} cm^2.

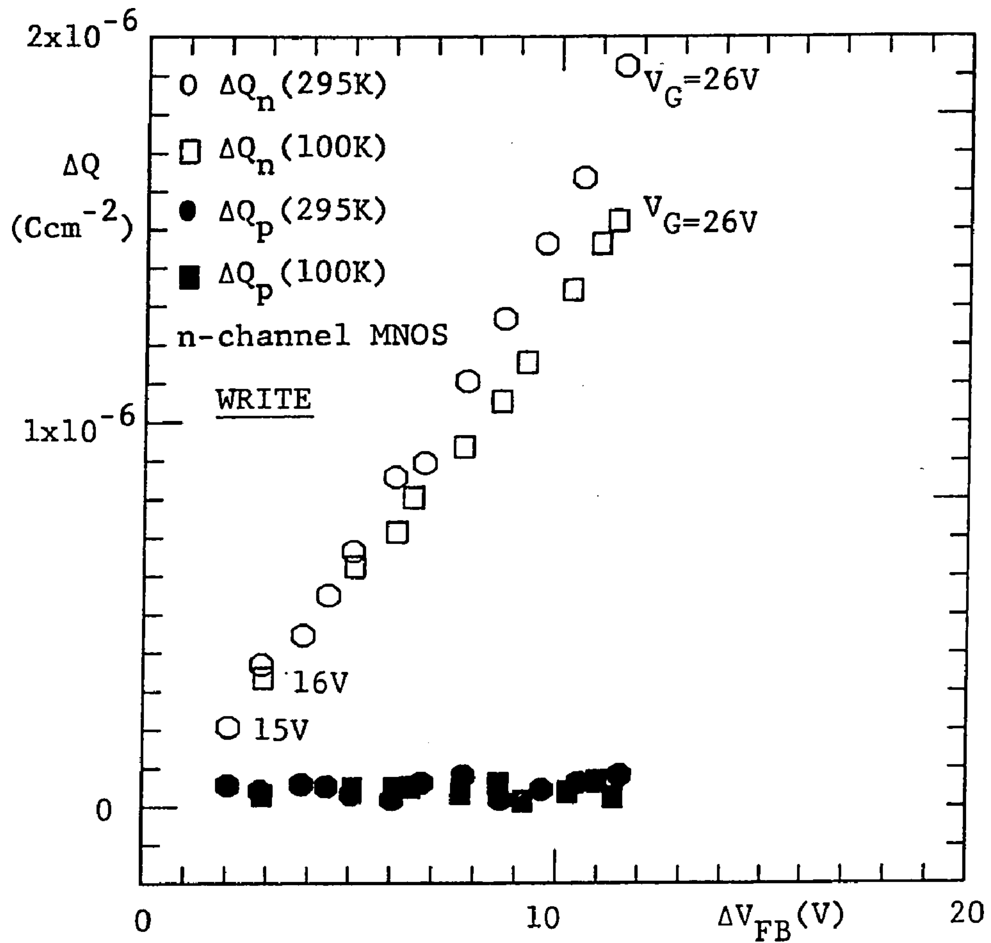


Fig. 7.25 ΔQ vs ΔV_{FB} data extracted from the ramp measurements in the write mode of operation of the n-channel MNOS device at two different temperatures. The device dimensions are: X_{oxt}=20A, X_N=500A, and Area=0.833x10^{-4}cm^2.

reflects the net trapping mechanism in the nitride layer.

A typical Erase data is shown in Fig.7-26. Along with the parasitics αC_{pad} in the bulk component and $\alpha C_{overlap}$ in the source current, an unusual structure in the curves denoted by ABCD and A'B'C'D' is noticeable. This is interpreted as the inversion of the n^+ source diffusion region due to the negative bias on the gate. During the region A'B', holes to the inversion layer of the source region are supplied by the bulk and therefore the bulk current increases negatively whereas the source current (AB) decreases toward zero due to the formation of the inversion layer. During the return sweep, holes from this region are swept out as the bulk current (C'D') and the source current (CD) increases again. Nevertheless, the gate current after substracting parasitics, still, reflects the trapping of carriers in the insulator. ΔQ vs ΔV_{FB} data is still meaningful as long as the charge separation is ignored. Such data at low temperature (100 K) is shown in Fig.7-27.

7.2.4 On the Back-Tunneling

The ramp technique minimizes the back-tunneling of the injected charge before the new state of the device may be determined. This is illustrated in Figure7-28 with a Write operation on an n-channel MNOS device. The device is swept in strong inversion (+25 V) and then swept-back. The new flatband voltage now shifts to a positive value due to electron injection from the semiconductor into the

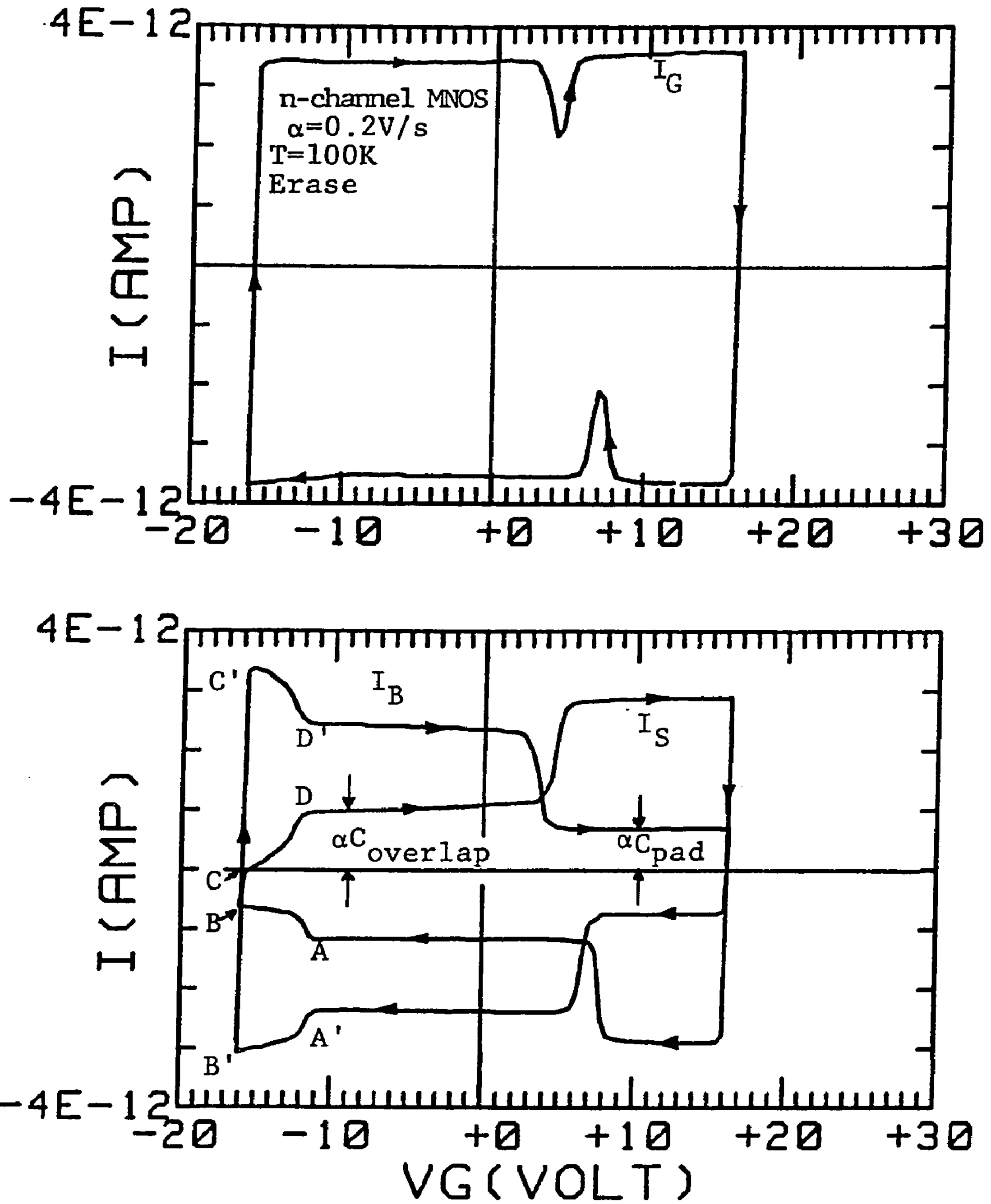


Fig. 7.26 Various current components in the erase mode of operation of the n-channel MNOS device.

nitride film. The position of the portion A on the return sweep determines the new state of the device. It is noted that as long as the device is inverted, the field in the tunneling oxide and the semiconductor is not favorable for the trapped electrons to back-tunnel to the semiconductor. This minimizes the bulk tunneling of the electrons. However, the field becomes favorable for this to happen once the semiconductor is accumulated as shown by the energy band diagram.

Figure 7-29 verifies this argument. Portion 'A' of the Fig.7-28 is redrawn and marked 'A' in Fig.7-29. The same experiment is repeated, but this time the return sweep is stopped at +10 V so the semiconductor is still inverted and the field is not favorable for back-tunneling. After 1 hour, the curve 'B' is traced with a very little change in the flatband voltage of the device indicating minimal back-tunneling. The entire exercise is repeated again but this time the sweep is stopped at 0 V for 1 hour so that the semiconductor is accumulated and the field is favorable for the back-tunneling of electrons to occur. After 1 hour, curve 'C' is traced with significant decrease in flatband voltage indicating appreciable back-tunneling. Next, the sweep is stopped at -5 V for 1 hour and then the curve 'D' is traced indicating larger extent of the back-tunneling.

This experiment proves that the ramp technique minimizes the

back-tunneling of the injected charge before the final state of the device can be read.

7.3 Summary

In this chapter, a charge separation technique, based on the linear voltage ramp, is described to study the charge injection in the MNOS/MONOS multi-dielectric structures. This technique works effectively in the Write mode of operation (semiconductor is driven into strong inversion) and separates the electron and hole components of the current. In addition, it also measures the flatband voltage shift.

Application of this technique to n and p-channel devices reveals that (a) In the Write mode of p-channel device ($V_G<0$), tunneling of holes from the semiconductor into the nitride layer with subsequent capture by deep traps is the dominant mechanism, (b) In the Write mode of n-channel device($V_G>0$), electron tunneling from the semiconductor into the nitride layer with subsequent capture by the traps is the dominant mechanism. Although the interpretation regarding the charge separation is ambiguous in the Erase mode of operation, it is reasonable to assume that Erase mode is equivalent to the Write mode of the conjugate structure. With this assumption, it then follows that for a given device, holes for the semiconductor are injected into the nitride valence band for one gate bias polarity and electrons are injected from the semiconductor into the

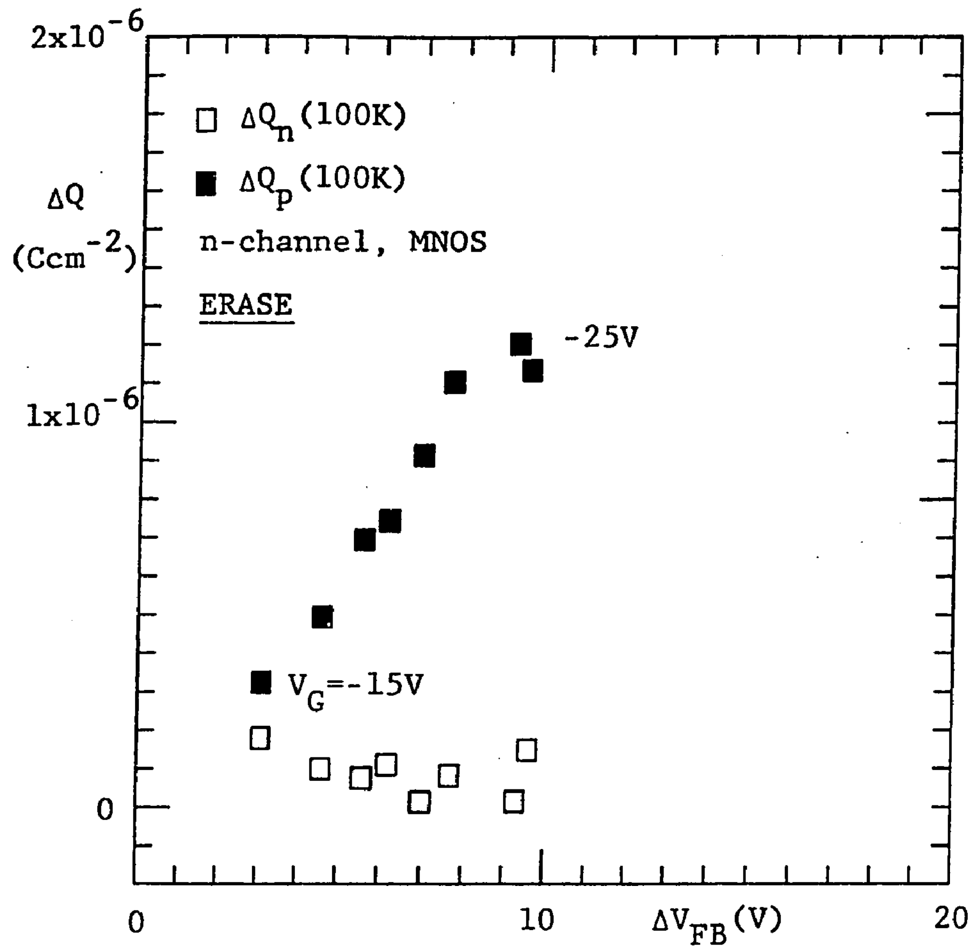


Fig. 7.27 ΔQ vs ΔV_{FB} data extracted from the ramp measurements in the erase mode of operation of the n-channel MNOS device at 100K. The device dimensions are: X_{oxt}=20A, X_N=500A, and Area=$0.833 \times 10^{-4} cm^2$.

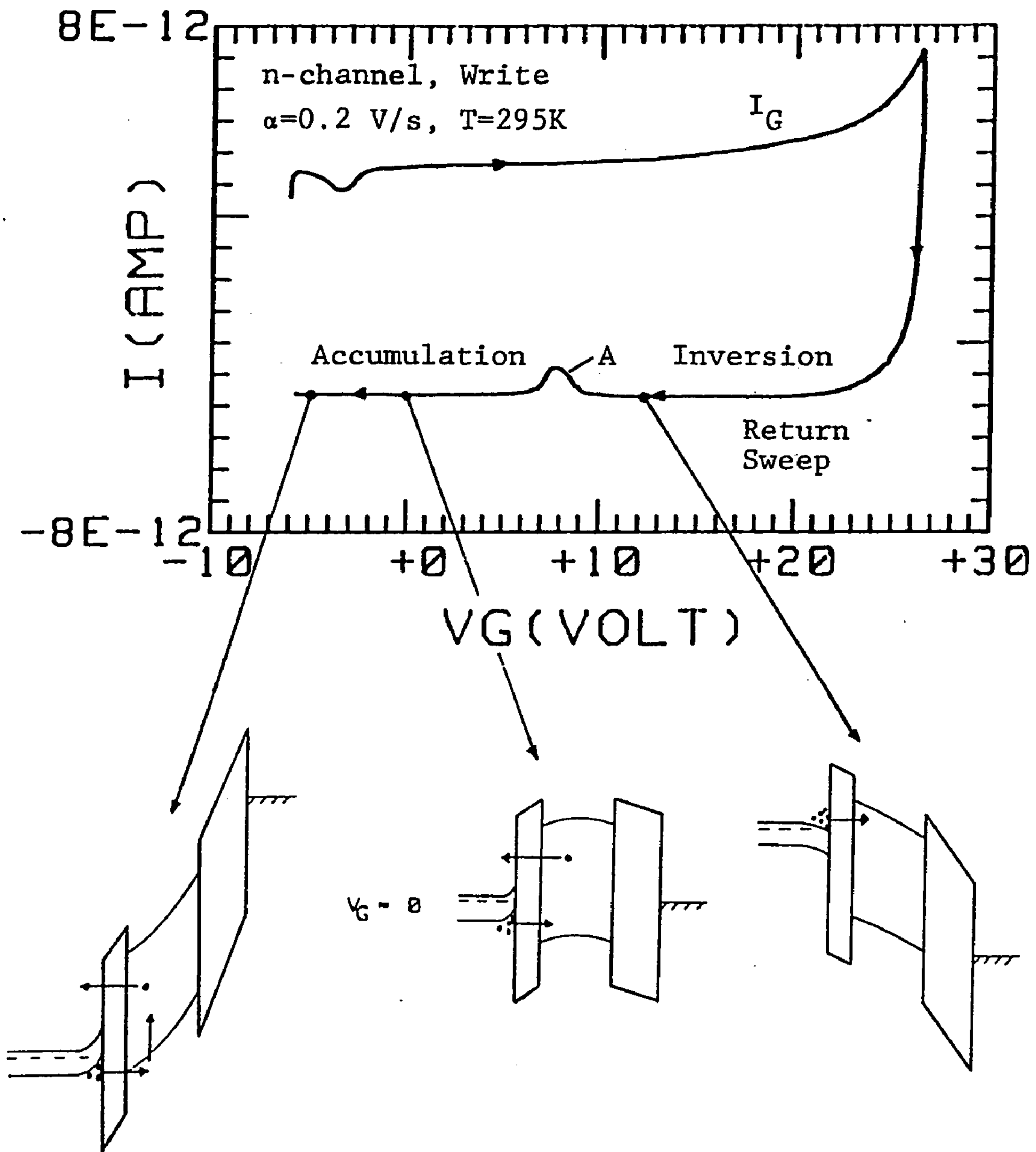


Fig. 7.28 I_G-V_G curve for the n-channel device with energy band-diagrams at various points during the reverse sweep.

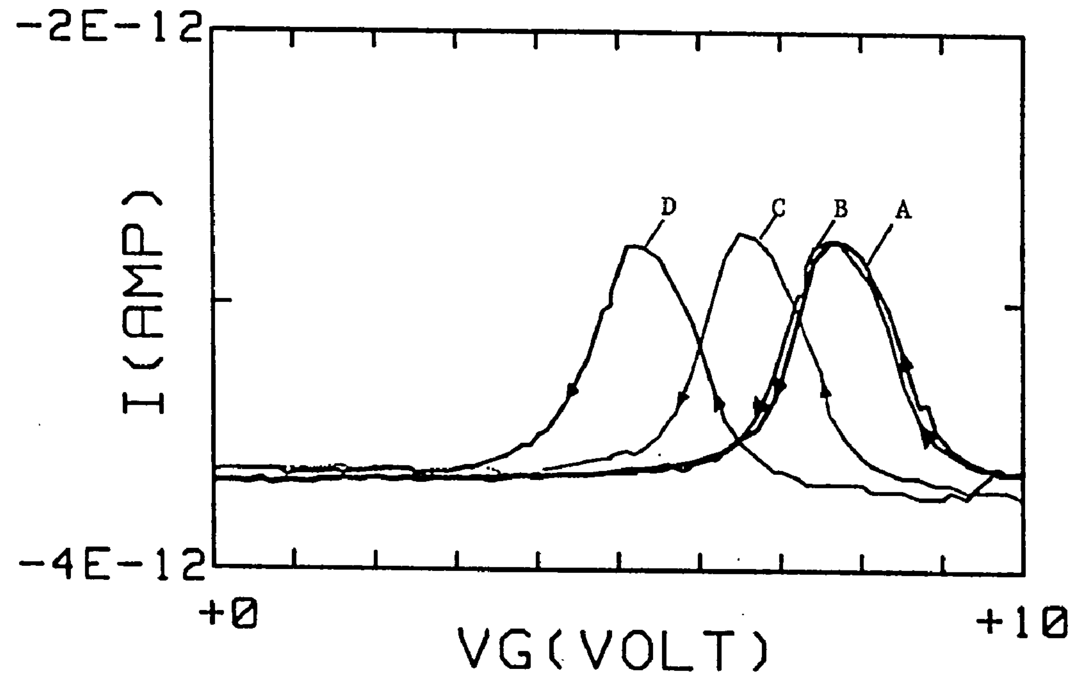


Fig. 7.29 The experiment on the back-tunneling. The curve A represents the portion 'A' of the Fig.7.28. The curves B,C and D represent the reverse sweeps after the Write operation with 1 hr. delay at +10V, 0V, and -5V respectively.

nitride conduction band for the opposite polarity. It is speculated that the recombination occurs in the nitride layer via deep lying amphoteric traps.

Temperature dependence of the observed data may be explained on the basis of increased capture cross-section of an attractive trap in the nitride layer at the reduced temperature. Finally, the ramp technique minimizes the back-tunneling of the injected charge before the final state of the device can be determined.

CHAPTER 8

CONCLUSIONS

This chapter is a summary of the key analytical and experimental conclusions arrived at in the preceding chapters. Also included is a brief description of related work that is presently under way, or envisaged for the future, using the analytical and experimental techniques developed in this thesis.

8.1 Studies on $Si-SiO_2$ Deep-Interface States

The following conclusions may be drawn from the study of $Si-SiO_2$ interface traps presented in chapters 2 and 3:

1. The two-terminal equilibrium technique of small signal admittance has been successfully extended to the three-terminal non-equilibrium case with the ability to extract trap parameters (D_{it}, σ_n, σ_p etc.) near the midgap of the semiconductor. The Shockley-Read-Hall (SRH) statistics, when applied to the $Si-SiO_2$ interface traps in a deep-depleted gated-diode structure, predicts the "pinning" of the steady-state occupancy function. The "pinning" occurs at a trap level which is $(1/2)\ln(\sigma_n/\sigma_p)kT$ eV below (above if $\sigma_n/\sigma_p < 0$) the intrinsic level at the interface. Interface traps below this "pinning" level always remain occupied regardless of

the position of the majority carrier Fermi level in deep-depletion.

2. The theoretical simulations of the small signal admittance predict a rapid decay or "pinning" in the normalized conductance ($G_{it}/\omega C_{ox}$) vs frequency curves at a critical value of the surface potential which places the majority carrier Fermi level at the so called "pinning" trap level.
3. It is found that the traps located at the "pinning" trap level are the most efficient generating centers and the generation rate is proportional to $\sqrt{\sigma_n \sigma_p}$.
4. Experimental measurements of small signal admittance as a function of frequency and surface potential confirm the "pinning" action of the normalized conductance vs frequency curves.
5. The trap parameters obtained by fitting the admittance data on the two samples (n-type, <100>, 5-10ohm-cm substrates) studied are found to be consistent with the surface leakage current measurements.
6. In the two samples studied, high values of $R=\sigma_n/\sigma_p=10^4$ and $\sigma_n=2.5\times10^{-14}cm^2$ corresponding to a low value of $D_{it}=7.0\times10^{10}eV^{-1}cm^{-2}$, and the low values of R=50 and $\sigma_n=10^{-15}cm^2$ with a high value of $D_{it}=6.0\times10^{11}eV^{-1}cm^{-2}$ are obtained for the interface traps located (1/2)ln(R)kT eV

below the intrinsic Fermi level at the interface. The relatively high values of R indicate donor like traps in the lower half bandgap of n-type Si.

8.2 Recommendations for Future Work on Interface-Traps

The technique developed in this thesis should be applied to p-type material, as well. Systematic variations in processing and the crystallographic orientation of n and p type substrates are recommended. As the interface state density is reduced below 10^{10} $cm^{-2}eV^{-1}$, the conventional lock-in technique is inadequate. Use of a Bridge and Lock-in amplifier as a null detector is envisaged for studying low densities of the interface states. In addition, the temperature dependence of R can be studied to determine the coulombic nature of these deep level traps.

8.3 Charge Injection and Trapping in MNOS/MONOS Devices:

The principal aim of this work was to determine the dominant tunneling carrier type and to study the trapping process in MNOS/MONOS devices. To this end, both n and p-channel devices were studied. The following conclusions may be drawn from chapters 4 to 7:

1. A new experimental technique based on the linear voltage ramp has been developed which minimizes back-tunneling and separates charge carriers at the injecting boundry during the Write operation with a simultaneous

measurement of the flatband voltage shift.

2. It is found that in the Write operation (both n and p-channel devices) inversion charge tunnels from silicon to silicon nitride and a small amount of majority carrier charge back-tunnels from silicon nitride to silicon.

3. The interpretation of the charge separation in the Erase mode is ambiguous due to possible recombination of electrons and holes near the Si-SiO_2 interface. However, it is reasonable to assume that the Erase mode is equivalent to the Write mode of the conjugate structure. With this assumption, it follows that for a given device, the semiconductor injects holes into the nitride valence band under a negative gate bias and electrons into the nitride conduction band under a positive bias. This strongly suggests that the carriers drift in the nitride bands and recombine via deep-lying "amphoteric" traps which can exist in negative, neutral and positive charge states.

4. Temperature dependence of the observed data for both Write and Erase modes is explained on the basis of increased capture cross-section of an attractive trap in the nitride layer at a reduced temperature.

5. The theory of the linear voltage ramp measurements based on a single carrier injection via modified Fowler

Nordheim tunneling and subsequent capture via donor or acceptor traps in the nitride layer predicts the observed gate current vs gate voltage characteristics.

8.4 Present and Future Investigation

The technique developed in this thesis is presently being applied to study the charge injection in the scaled-down MONOS devices. Technology/trap relationship leading to the specific tailoring of the traps in Si_3N_4 is also being investigated. The problem of identifying the possible "amphoteric" trap still remains unsolved.

REFERENCES

1. M. M. Atalla, E. Tannebaum and E. J. Scheibner, Bell Syst. Tech. J., 38, 749 (1959).

2. S. R. Hofstein and F. P. Heiman, Proc. IEEE, 51 , 1190 (1963).

3. C. H. Sequin and M. F. Tompsett, Charge Transfer Devices, Academic Press, New York, 1975.

4. J. J. Chang, Proc. IEEE, 64, 1039 (1976).

5. J. F. Verwey, Advances in Electronics and Electron Physics 41, Academic Press, 1976, 249-309.

6. P. Bergveld, IEEE Trans. Biomed. Engg. BME-17 , 70 (1970).

7. S. K. Clark and K. D. Wise, IEEE Trans. Elect.Dev. ED-26, 1887 (1979).

8. F. J. Feigl, Characterization of Dielectric Films, Academic Press, VLSI Electronics Microstructure Science Vol. 6, 1983.

9. B. J. Baliga and E. Sun, IEEE Trans. Elect. Dev., ED-24, 685 (1977).

10. C. W. Pearce, L. E. Katz and T. E. Seidel, Semiconductor Silicon 1981, Electrochem. Soc., Pennington, NJ, 1981, p.705.

11. W. Shockley and W. T. Read, Phys. Rev., 87 387 (1952).

12. R. N. Hall, Phys. Rev. 87, 387 (1952).

13. A. F. Tasch, Jr., and C. T. Sah, Phys. Rev. B1 , 800 (1970).

14. C. T. Sah, L. Forbes, L. I. Rosier, A. F. Tasch, Jr., and A. B. Tole, Appl. Phys. Lett. 15 145 (1969).

15. E. H. Nicollian and J. R. Brews, MOS Physics and Technology, Ch. 5, John Wiley and Sons, New York, 1982.

16. E. H. Nicollian and A. Goetzberger, Bell Syst. Tech. J. 46, 1055 (1967).

17. H. Deuling, E. Klausmann and A. Goetzberger, Solid-St. Electron. 15, 559 (1972).

18. G. Declerck, R. Van Overstraeten and G. Broux, Solid-St. Electron. 16, 1451 (1973).

19. J. A. Cooper, Jr., and R. J. Schwartz, Solid-St. Electron. 17, 641 (1974).

20. K. Lehovec, Appl. Phys. Lett. 8, 48 (1966).

21. T. C. Poon and H. C. Card, J. Appl. Phys. 51, 5880 (1980).

22. S. M. Sze, Physics of Semiconductor Devices, John Wiley and Sons, New York, 1981, p. 37.

23. N. F. Mott and R. W. Gurney, Electronic Processes in Ionic Crystals, Oxford University Press, New York, 1940, p. 172.

24. A. Rose, Phys. Rev. 97, 1538 (1955).

25. R. W. Smith and A. Rose, Phys. Rev. 97, 1531 (1955).

26. M. A. Lampert, Phys. Rev. 103, 1648 (1956).

27. R. I. Frank and J. G. Simmons, J. Appl. Phys. 38, 832 (1967).

28. J. Frenkel, Phys. Rev. 54, 647 (1938).

29. D.L. Pulfrey, A. H. M. Shousha and L. Young, J. Appl. Phys. 41, 2838 (1970).

30. A. Many and G. Rakavy, Phys. Rev. 126, 1980 (1962).

31. A. Many, S. Z. Weisz and M. Simhony, Phys. Rev. 126, 1989 (1962).

32. P. C. Arnett, J. Appl. Phys. 46, 5236 (1976).

33. J. J. O'Dwyer, J. Appl. Phys. 37, 599 (1966).

34. S. M. Sze, J. Appl. Phys. 38, 2951 (1967).

35. S. M. Hu, J. Electrochem. Soc. 113, 693 (1966).

36. A. S. Ginovker, V. A. Gritsenko and S. P. Sinatsa, Phys. Status Solidi 26a, 489 (1974).

37. Z. A. Weinberg, W. C. Johnson and M. A. Lampert, Appl. Phys. Lett. 25, 42 (1974).

38. Z. A. Weinberg and R. A. Pollack, Appl. Phys. Lett. 27, 254 (1975).

39. Z. A. Weinberg, Appl. Phys. Lett. 29, 617 (1976).

40. D. K. Schroder and M. H. White, IEEE Trans. Elect. Dev. ED-26, 899 (1979).

41. P. C. Arnett and Z. A.Weinberg, IEEE Trans. Elect. Dev. ED-25, 1014 (1978).

42. J. G. Simmons and G. W. Taylor, Phys. Rev. B4, 502 (1971).

43. W. E. Dahlke and D. W. Greve, Solid-St. Electron. 22, 893 (1979).

44. A. S. Grove, Physics and Technology of Semiconductor Devices, John Wiley and Sons, New York, 1967, pp. 299-302.

45. T. Sugano, Proc. of the International Conf. INFOS 83, 1983, p. 1.

46. E. Suzuki, H. Hiraishi, K. Ishii and Y. Hayashi, IEEE Trans. Elect. Dev. ED-30, 122 (1983).

47. F. L. Hampton and J. R. Cricchi, IEDM Tech. Digest, 1979, p. 374.

48. Y. Yatsuda, T. Hagiwara, S. Minami, R. Kondo, K. Uchida and K. Uchiumi, 13th Conf. Solid-St. Dev. (Tokyo, Japan), 1981, p. 27.

49. E. Suzuki, Y. Hayashi and H. Yanai, Appl. Phys. Lett. 35, 790 (1979).

50. E. Suzuki, Y. Hayashi and H. Yanai, J. Appl. Phys. 52, 6377 (1981).

51. V. Y. Doo, D. R. Nichols and G. A. Silvey, "Vapor Deposition of Silicon Nitride," Fall Meeting of Electrochemical Society, Buffalo, NY, 1965.

52. V. Y. Doo, "Silicon Nitride, A New Diffusion Mask, IEDM Tech. Digest, Washington, DC, 1965.

53. V. Y. Doo, IEEE Trans. Elect. Dev. ED-13, 561 (1966).

54. S. M. Hu, D. R. Kerr and L. V. Gregor, Appl. Phys. Lett. 10, 97 (1967).

55. T. L. Chu, J. R. Szedon and C. H. Lee, Solid-St. Electron. 10, 897 (1967).

56. C. L. Hutchins and R. W. Lade, Proc. IEEE 55, 1494 (1967).

57. B. E. Deal, P. J. Fleming and P. L. Castry, J. Electrochem. Soc. 115, 300(1968).

58. T. L. Chu, J. R. Szedon and C. H. Lee, J. Electrochem. Soc. 115, 318 (1968).

59. H. A. R. Wegener, A. J. Lincoln, H. C. Pao, M. R. O'Connel and R. E. Oleksiak, IEDM Tech. Digest, Washington, DC, 1967.

60. H. C. Pao and M. O'Connel, Appl. Phys. Lett. 12, 260 (1968).

61. D. Frohman-Bentchkowsky and M. Lenzlinger, J. Applied Phys. 40, 3307 (1969).

62. C. Svensson and I. Lundstrom, Electron. Lett. 6, 645 (1970).

63. J. T. Wallmark and J.H. Scott, Jr., RCA Rev. 30, 335 (1969).

64. E. C. Ross and J. T. Wallmark, RCA Rev. 30, 367 (1969).

65. G. Dorda and M. Pulver, Phys. Status Solidi (a) 1, 71 (1970).

66. A. V. Ferris-Prabhu, Appl. Phys. Lett. 20, 149 (1972).

67. A. V. Ferris-Prabhu, Phys. Status Solidi (a) 11, 81 (1972).

68. A. V. Ferris-Prabhu, Solid-St. Electron. 16, 1086 (1973).

69. M. H. White and J. R. Cricchi, IEEE Trans. Elect. Dev. ED-19, 1280 (1972).

70. K. I. Lundstrom and C. M. Svensson, IEEE Trans. Elect. Dev. ED-19, 826 (1972).

71. J. J. Chang, IEEE Trans. Elect. Dev., ED-24, 511 (1977).

72. B. H. Yun, Appl. Phys. Lett. 23, 152 (1973).

73. B. H. Yun, "Non-destructive Measurements of Dielectric Properties," U.S. Patent 3840809, Oct. 8, 1974.

74. B. H. Yun, Appl Phys. Lett. 25, 340 (1974).

75. G. W. Taylor and J. G. Simmons, Solid-St. Electron. 17, 1 (1974).

76. L. S. Wei and J. G. Simmons, Solid-St. Electron. 17, 591 (1974).

77. H. Maes and R. Overstraeten, IEDM Tech. Digest, 1974, p. 119.

78. K. Lehovec, C. H. Chen and A. Fedotowsky, IEEE Trans. Elect. Dev. ED-25, 1030 (1978).

79. L. Lundkvist, I. Lundstrom and C. Svensson, Solid-St. Electron. 16, 811 (1973).

80. C. A. Neugebauer and J. F. Burgess, J. Appl. Phys. 47, 3182 (1976).

81. M. H. White, J. W. Dzimianski and M. C. Peckerar, IEEE Trans. Elect. Dev. ED-24 , 577 (1977).

82. M. Kuhn, Solid-St. Electron. 13, 873 (1970).

83. K. Board and J. G. Simmons, Solid-St. Electron 20, 859 (1977).

84. U. Kelberlau and R. Kassing, Solid-St. Electron. 24, 321 (1981).

85. L. Faraone, A. G. Nassibian and J. G. Simmons, J. Appl. Phys. 50, 5865 (1979).

86. L. Faraone, J. G. Simmons, A. K. Agarwal and P. D. Tonner, Solid-St. Electron. 24, 709 (1981).

87. P. D. Tonner and J. G. Simmons, Solid-St. Electron. 25, 733 (1982).

88. R. C. Hughes, Phys. Rev. Lett. 30, 1333 (1973).

89. P. C. Arnett and B. H. Yun, Appl Phys. Lett. 26, 94 (1975).

90. K. Lehovec and A. Fedotowsky, J. Appl. Phys. 48, 2955 (1977).

91. F. B. Hildebrand, Introduction to Numerical Analysis, McGraw Hill, NY, 1974.

92. A. G. Milnes, Deep Impurities in Semiconductors, John Wiley and Son, New York, 1973, Ch. 5.

93. V. L. Bonch-Bruevich and E. G. Landsberg, Phys. Status Solidi, 29a, 9 (1968).

94. M. Lax, J. Phys. Chem. Solids 8, 66 (1959).

95. M. Lax, Phys. Rev. 119, 1502 (1960).

96. G. A. Dussel and K. W. Boer, Phys. Status Solidi 39, 375 (1970).

97. J. R. Cricchi, F. C. Blaha and M. D. Fitzpatrick, IEDM Tech. Digest, 1973, p. 126.

VITA

Anant Kumar Agarwal was born in Rochester, N.Y., on June 5, 1957. His family moved to Allahabad, India the same year. He attended elementary schools in Allahabad and received a Bachelor of Engineering Degree in Electrical Engineering from the University of Allahabad in July, 1978. In the Fall of 1978, he joined the University of Tennessee Space Institute. After finishing the requirements for a Master of Science degree in Electrical Engineering, he joined Lehigh University in the Spring of 1980 to work towards a Ph.D. degree in the area of Microelectronics. The requirements for the Ph.D. degree were finished in June 1984.

LIST OF PUBLICATIONS:

1. L. Faraone, J. G. Simmons and A. K. Agarwal, "Interpretation of Non-equilibrium Measurements on MOS Devices Using the Linear Voltage Ramp Technique," Solid-St. Electron. 24, 709 (1981).

2. L. Faraone, A. K. Agarwal and J. G. Simmons, "The Non-equilibrium Linear Voltage Ramp Technique as a Diagnostic Tool for the MOS Structure," presented at the International Conference on Insulating Films on Semiconductors (INFOS), 1981.

3. A. K. Agarwal, F. M. Rhodes and M. H. White, "Characterization of $Si\text{-}SiO_2$ Interface States with Admittance Measurements on a Surface-Potential-Controlled (SPC)-MOS Capacitor," presented at the 1982 Semiconductor Interface Specialist Conference.

4. A. K. Agarwal, F. M. Rhodes amd M. H. White, "Experimental Verification of Electron-Hole Recombination Theory for $Si\text{-}SiO_2$ Interface Traps with Non-equilibrium Steady-State, Admittance Measurements," Appl. Phys.

Lett., Sept. 1 (1983).

5. A. K. Agarwal and M. H. White, "On the Non-equilibrium Statistics and Small Signal Admittance of Si-SiO_2 Interface Traps in the Deep- Depleted Gated-Diode Structure," J. Appl. Phys., May 15 (1984).

6. A. K. Agarwal, C. Chao, R. H. Vogel and M. H. White, "On the Transient and Steady-State Transport of Electrons and Holes in MNOS and MONOS Devices," presented at the 1983 International Electron Device Meeting (IEDM), Washington DC.