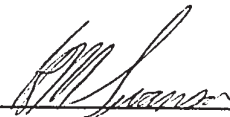


**SIPOS HETEROJUNCTION CONTACTS
TO SILICON**

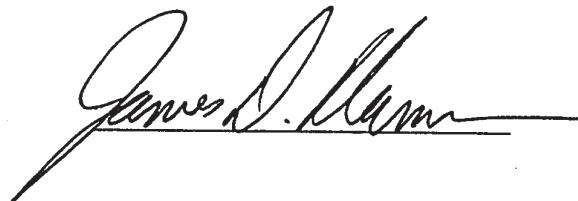
**A DISSERTATION
SUBMITTED TO THE
DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

**By
Young Hoon Kwark
June 6, 1984**

I certify that I have read this thesis and that in my opinion
it is fully adequate, in scope and quality, as a dissertation for
the degree of Doctor of Philosophy


(Principal Adviser)

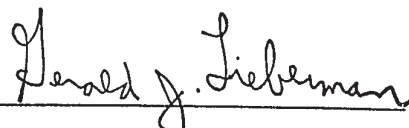
I certify that I have read this thesis and that in my opinion
it is fully adequate, in scope and quality, as a dissertation for
the degree of Doctor of Philosophy



I certify that I have read this thesis and that in my opinion
it is fully adequate, in scope and quality, as a dissertation for
the degree of Doctor of Philosophy



Approved for the University Committee
on Graduate Studies


Dean of Graduate Studies & Research

ACKNOWLEDGEMENT

The author would like to thank his advisors Professors R. M. Swanson and J. D. Plummer for their guidance, and their friendship during what has been a rewarding experience. The students in the Thermophotovoltaics group and the Integrated Circuits and Solid State Labs are to be thanked for their many friendships. Special thanks go to my wife and to my family for their patience during this period. This thesis is dedicated to my grandparents.

CONTENTS

1	Introduction	1
1.1.	Introduction	1
1.2.	Approach and Outline	3
2	Device Applications of Selective Contacts	6
2.1.	Introduction	6
2.2.	Ideal Contacts	7
2.3.	Practical Contacts	8
2.4.	The Solar Cell	10
2.5.	The Point Contact Cell	11
2.6.	Bipolar Transistors	17
2.7.	Homojunction BJT	17
2.8.	Idealized BJT	19
2.9.	Shallow Emitter Devices	22
2.10.	Selective Contacts	24
2.11.	Current Crowding	25
2.12.	A Circuit Application	27
2.13.	Summary	29
3	Low Frequency Behavior of Emitter Structures	31
3.1.	Introduction	31
3.2.	Lumped Parameter Equivalent Circuit	32
3.2.1.	Transport Equations	32
3.2.2.	Diffused or Bulk Regions	34
3.2.3.	Interfaces	38
3.2.4.	Tunneling Barriers	39
3.2.5.	Integration into a Network	41
3.3.	Practical Emitter Structures	42

3.3.1. Homojunction Emitter	43
3.3.2. Heterojunctions	50
3.3.3. SIS contacts	57
3.3.4. Majority Carrier Series Resistance	62
3.4. Summary	64
4 Dynamic Behavior of Emitter Structures	66
4.1. Introduction	66
4.2. Time Dependent Transport Equations	67
4.3. Charge Control	67
4.3.1. Constant Profile Emitter	68
4.3.2. Application of Charge Control	70
4.3.3. Validity of Charge Control	71
4.4. Exact Calculations for the Constant Profile	72
4.5. Comparison with Charge Control	77
4.6. Nonconstant Profiles	82
4.7. Validity of Assuming Constant S_c	84
4.8. SUMMARY	87
4.9. Bipolar Junction Transistors	88
4.10. Current Gain	88
4.11. Charge Control	89
4.12. Phasor Viewpoint	92
4.13. Impact of Emitter and Base Structures on Device Performance . . .	94
4.14. Maximum Frequency of Oscillation	108
4.15. SUMMARY	109
5 SIPOS	110
5.1. Introduction	110
5.2. SiO_2 Films	110
5.3. CVD SIPOS	111

5.4. Device Applications for SIPOS Films	114
5.5. Application to Selective Contacts	117
5.6. Minority Carriers	119
5.7. SIPOS Deposition	122
5.8. Summary	129
6 Experimental Determination of J_{oc}	130
6.1. Introduction	130
6.2. The Base Current	131
6.3. Series Resistance	136
6.4. Temperature Measurement	137
6.5. Measurement Procedure	138
6.6. Design Considerations for the Test Device	139
6.7. Summary	140
7 Experimental Values of J_{oc} for POLY and SIPOS Emitters	141
7.1. Introduction	141
7.2. Format of Presentation	143
7.3. Conventional Emitters	144
7.3.1. Influence of Aluminum Anneal Temperature	147
7.3.2. Influence of Final High Temperature Anneal	148
7.3.3. Modeling the Conventional Emitter	149
7.4. SIPOS Emitters	151
7.4.1. Role of Oxygen in Determining J_{oc} :	152
7.4.2. The $\gamma = 0$ Film	155
7.4.2.1. J_{oc} vs. Junction Depth	155
7.4.2.2. J_{oc} vs. Anneal Temperature	155
7.4.3. Influence of Anneal Temperature and Interfacial Oxide	156
7.4.4. Devices with No Chem Ox	157
7.4.5. Modeling SIPOS Contacts Without a Chem Ox	158

7.4.6. SIPOS Devices with Chem Ox	163
7.4.7. Modeling of a Chem Ox Device	166
7.4.8. 800°C SIPOS Devices With Chem Ox	167
7.5. POLY Contacted Emitters	172
7.5.1. Devices without a Chem Ox	175
7.5.2. Devices with Chem Ox	176
7.6. Comparison of SIPOS and POLY Structures	177
7.7. Proposed Mechanism for SIPOS Emitter Operation	178
7.8. Determining the Bandgap of SIPOS Using a C-V Technique	182
7.9. summary	184
8 Conclusions	185
A Mechanical Analogy for the Solar Cell	188
B Bipolar Process Using Selective Emitter Contacts	193
C Proof of Ebers-Moll Reciprocity	195
C.1. Plausibility Argument	195
C.2. Formal Proof	196
D Model Parameters Used in Simulations	200
E Refinement of the Charge Control Result	204
F Preheater Calculations	206
G Conductivity Decay Technique for J_{oe}	211
H Perimeter Streamline Patterns	214
I Base Current Crowding	218
J Effect of BWM on the Collector Temperature	221

K	Source of the Base Perimeter Current	224
L	Experiments with Chemically Grown Oxide	228
M	Bibliography	231

TABLES

Table 4.1. parameters and results for BJT simulations	97
Table 7.1. measured J_{oe} vs. V_{be} for BJT with conventional emitter	146
Table 7.2. J_{oe}, G_b vs. FGA temperature ($T_A = 850^\circ C$) EM-B device	147
Table 7.3. J_{oe}, G_b vs. T_A , final high temperature anneal	148
Table 7.4. lifetime parameters used in J_{oe} simulations	151
Table 7.5. process data for J_{oe} vs. γ device set	153
Table 7.6. J_{oe}, G_b vs. T_A for SIPOS devices	156
Table 7.7. J_{oe} vs. T_A for POLY devices types B and C	173
Table L.1. chemox growth sequence	229

FIGURES

Figure 2.1. ideal N -type contact	7
Figure 2.2. N -type contact where $\phi_M > \phi_S$	9
Figure 2.3. cross section of FSC and BSC point contact cells	12
Figure 2.4. band diagrams of a P - I - N solar cell	14
Figure 2.5. solar cell energy conversion efficiency vs. contact J_o	16
Figure 2.6. Band diagrams for BJT	18
Figure 2.7. Ideal one-dimensional BJT	20
Figure 2.8. cross-section of emitter sidewall region	25
Figure 2.9. generalized differential amplifier stage	28
Figure 3.1. General bulk region	35
Figure 3.2. equiv. circuit for bulk region	36
Figure 3.3. equiv. circuit for uniform region	37
Figure 3.4. equiv. circuitfor interfaces	39
Figure 3.5. equiv. circuit for tunnelable insulator	40
Figure 3.6. boundary conditions for matching adjacent regions	42
Figure 3.7. generalized contact structure	43
Figure 3.8. band structure and equiv. circuit for homojunction emitter	44
Figure 3.9. $\overline{v_{diff}}$ for a uniform homojunction emitter	46
Figure 3.10. $\overline{p_o}$ for a uniform homojunction emitter	47
Figure 3.11. J_{oe} for a uniform homojunction emitter	48
Figure 3.12. $\overline{v_{diff}}, \overline{p_o}, J_{oe}$ for Gaussian emitter	49
Figure 3.13. J_{oe} vs. emitter depth for uniform emitter	50
Figure 3.14. J_{oe} vs. emitter depth for Gaussian emitter	51
Figure 3.15. J_{oe} vs. emitter depth for ERFC emitter	52
Figure 3.16. heterojunction emitter+contact structure	53
Figure 3.17. J_{oe} vs. POLY thickness for $N+$ POLY emitter	55
Figure 3.18. J_{oe} for mono+polycrystalline emitter	56

Figure 3.19. SIS contact structure	58
Figure 3.20. J_{oe} vs. emitter depth, S_c for uniform emitter	59
Figure 3.21. J_{oe} vs. N, S_c for uniform emitter	62
Figure 3.22. J_{oe} vs. N_{surf}, S_c for ERFC emitter	63
Figure 4.1. symbol definition for uniform emitter	69
Figure 4.2. exact $J_{oe}(\omega)$ solutions for uniform emitter	74
Figure 4.3. effect of W and τ_p on $J_{oe}(\omega)$	76
Figure 4.4. comparison of charge control and exact solution, case 1	78
Figure 4.5. comparison of charge control and exact solution, case 2	79
Figure 4.6. comparison of charge control and exact solution, case 3	80
Figure 4.7. $J_{oe}(\omega)$ for ERFC emitter	83
Figure 4.8. influence of J_{ox} on $J_{oe}(\omega)$ for POLY SIS contact	85
Figure 4.9. $J_{oe}(\omega)$ for a 5000 Å POLY + outdiffused emitter	86
Figure 4.10. simplified model for charge control analysis of BJT	89
Figure 4.11. vector diagrams for phasor interpretation of β rolloff	92
Figure 4.12. vector diagrams for BJT at $f = f_T$	94
Figure 4.13. emitter dopant profile for Device A simulation	98
Figure 4.14. simulation of β vs. freq for Device A	99
Figure 4.15. simulation of J_{oe}, J_{ob} vs. freq. for Device A	100
Figure 4.16. simulation of β vs. freq. for Device B	101
Figure 4.17. simulation of J_{oe}, J_{ob} vs. freq. for Device B	102
Figure 4.18. simulation of β vs. freq for Device C	104
Figure 4.19. simulation of J_{oe}, J_{ob} vs. freq for Device C	105
Figure 4.20. simulation of β vs. freq. for Device D	106
Figure 4.21. simulation of J_{oe}, J_{ob} vs. freq. for Device D	107
Figure 5.1. shell model of SIPOS	114
Figure 5.2. SIPOS passivation of high voltage devices	117
Figure 5.3. SIPOS resistivity vs. N_2O flow rate, T_{anneal}	118

Figure 5.4. first generation SIPOS deposition apparatus	125
Figure 5.5. SIPOS deposition tube — overall view	126
Figure 5.6. SIPOS deposition tube — closeup of mixing chamber	127
Figure 5.7. SIPOS film resistivity vs. T_A and γ	129
Figure 6.1. components of the base current	131
Figure 6.2. determination of τ_p in the base from BWM	135
Figure 6.3. effect of series resistance on BJT characteristics	136
Figure 6.4. cross-section of BJT test structure	140
Figure 7.1. selective contact test structures	142
Figure 7.2. effect of series resistance on BJT characteristics	145
Figure 7.3. dopant profiles used in simulations of EM-A,B	149
Figure 7.4. J_{oe} vs. S_c for EM-A (A) and EM-B (B)	150
Figure 7.5. influence of γ on J_{oe} of SIPOS contacts	154
Figure 7.6. experimental J_{oe} vs. T_A for SIPOS contact without chemox . . .	157
Figure 7.7. simulated J_{oe} vs. S_c for SIPOS without chemox	160
Figure 7.8. dopant profiles used to simulate SIPOS contacted devices	161
Figure 7.9. simulated J_{oe} for $T_A = 1000^\circ C$ SIPOS without chemox	162
Figure 7.10. experimental J_{oe} for SIPOS contacts with chemox	164
Figure 7.11. effect of chemox on series resistance of SIPOS contacts	165
Figure 7.12. simulated J_{oe} vs. S_c for SIPOS contact with chemox	167
Figure 7.13. corrected Gummel plot for SIPOS devices w/ and w/o chemox .	168
Figure 7.14. β vs. I_c for SIPOS devices w/ and w/o chemox	169
Figure 7.15. corrected Gummel plot for small SIPOS device w/ chemox . . .	170
Figure 7.16. β vs. I_c for small SIPOS device w/ chemox	171
Figure 7.17. band diagrams for SIS structure with thick insulator	172
Figure 7.18. experimental J_{oe} vs. T_A for POLY contacted emitters	174
Figure 7.19. influence of annealing on SIPOS microstructure	179
Figure 7.20. cross-section of SIPOS layers for contact resistance	181

Figure 7.21. $C - V$ technique to determine SIPOS workfunction	183
Figure A.1. mechanical analogy for a solar cell	188
Figure B.1. bipolar process using selective emitter contacts	194
Figure C.1. lumped parameter equivalent circuit for transport equations . .	196
Figure C.2. control volume used for reciprocity proof	197
Figure C.3. equivalent circuit for control volume of Fig. C.2	197
Figure D.1. model used for hole diff. coefficient vs. dopant density	200
Figure D.2. band gap narrowing model used in simulations	201
Figure D.3. hole lifetime model used in simulations	202
Figure D.4. J_{oe} vs. N for semi-infinite uniform region using model parameters	202
Figure F.1. control volume for heat transfer equation	206
Figure F.2. gas temperature profile for N_2 preheater	209
Figure G.1. conductivity decay lifetime measurement apparatus	211
Figure G.2. band diagram of lifetime test wafer	212
Figure H.1. fringing component of electrons injected into the base	214
Figure H.2. idealized geometry for fringe components	216
Figure I.1. quasi-two-dimensional analysis of base current crowding	218
Figure I.2. J_c and V distributions for current crowding	220
Figure J.1. coordinate definition for BWM analysis	222
Figure J.2. graphical representation of effective slope m	222
Figure J.3. collector ideality factor, m , vs. V_{be}	223
Figure K.1. cross-section of BJT with field plate	225
Figure K.2. Gummel plots of field plate equipped BJT	226
Figure L.1. ellipsometric chemox thickness for several process conditions . .	230

Chapter 1

Introduction

1.1. Introduction

Contacts are sometimes thought of solely as a means to gain access to a semiconductor device; yet, they can play an important role in determining device behavior. For unipolar devices where only one type of carrier is involved, it is sufficient to characterize a contact in terms of the effective series impedance it presents to majority carrier flow. For a large class of applications, this description can be reduced to a simple contact resistance. For bipolar devices, a class which includes solar cells and transistors, two types of carriers are present and if the device geometry is such that minority carriers are present in the vicinity of a contact, then contacts can influence device behavior in a fundamental manner.

In distinguishing between unipolar and bipolar applications, it is important to remember that minority carriers are a metastable species when present in excess of their thermal equilibrium values. It is the distribution of this minority carrier recombination that determines the currents in a bipolar device. Therefore, if minority carriers are present in significant amounts near contact regions, this pattern of recombination may be altered depending upon the nature of the interaction between the carriers and the contact.

The trend in modern bipolar structures is toward shallow diffusions. There are several driving forces behind this reduction of vertical dimensions. For high

speed transistors, the base width must be kept small to minimize the transit time. Since the active base is usually formed as the result of two subsequent diffusions (or implants), process controllability requires that the emitter depth be comparable to the base width. Therefore, as the emitter depth decreases to values less than that of the diffusion length of injected minority carriers, the emitter becomes transparent, and minority carriers may survive to reach the emitter contact on the surface instead of recombining in the bulk of the emitter.

Traditional methods used to achieve low series resistance ohmic contacts to silicon inevitably result in a high surface recombination velocity for minority carriers. As emitters become shallower, the presence of this recombination sink in close proximity to the emitter-base junction results in an increase in the reverse injected current. This reduction in emitter injection efficiency translates into a lower current gain for the device.

For solar cells, the minority carriers which result from photogeneration of hole-electron pairs by the incident solar radiation must be collected before they recombine, since the potential energy of a recombining hole-electron pair is mostly dissipated to the lattice in the form of heat and is therefore unavailable to do useful electrical work. In conventional solar cells, the surface layer is heavily doped to reduce the lateral series resistance. Since the heavy doping results in a short diffusion length for photogenerated pairs, this layer must be kept thin if good blue response is to be achieved. Yet, a thin topside collecting junction (which is the analog of the emitter in a bipolar transistor) results in enhanced minority carrier recombination for the same aforementioned reasons. This translates into a lower energy conversion for the cell.

For both cases, it is seen that the presence of an active recombination site for minority carriers in close proximity to the injecting junction results in degradation of device performance. This degradation can be circumvented by the use of selective contacts which act as barriers to minority carrier flow, yet do not impede majority carrier transport by exhibiting a large series resistance. Such contacts, often called

minority carrier mirrors, may be composed of a semiconducting material different from that of the bulk device. The use of such heterojunctions is a common strategy in bipolar transistors made from compound semiconductors, but only recently have practical selective structures been developed for silicon.

The degree of interaction of the contact with the minority carriers is reflected in J_o , the saturation current due to the injected carriers. Through the use of a selective contact in the emitter structure, not only can J_{oe} , the emitter saturation current, be prevented from reaching excessively high values for thin emitters, but also it is possible to reduce it far below what is achievable with current homojunction structures. The possibility of trading-off the reduction in J_{oe} , or equivalently for bipolar transistors the high current gains that result, for a greater degree of freedom in device design is an intriguing prospect.

1.2. Approach and Outline

The application of selective contact structures to solar cells and to bipolar junction transistors is discussed in Chapter 2. There it will be shown how the placement of energy barriers in the minority carrier bands can be used to channel minority carrier flow in a more efficient manner than is possible than by relying solely on the built-in potentials of the diffusions. However, the idealized barriers assumed are not necessarily readily obtained in practice. Therefore, in Chapter 3, practical candidate structures, the conventional homojunction contact, a heterojunction contact and an SIS (Semiconductor-Insulator-Semiconductor) contact are analyzed to see how static J_o depends on structural parameters. The key sources of recombination in these candidate structures will be identified which will permit general processing guidelines to be established. In Chapter 4, this analysis is extended to finite frequencies; it is shown that although the dynamic behavior as reflected in the frequency dependent junction admittance differs for the various contact structures, the dynamic performance of the BJT in particular is dominated by other factors.

The use of contacts with low static values of emitter saturation current, J_{oe} , is seen to improve the low frequency performance, namely the static current gain, β_o , but that high frequency performance is still controlled mostly by the dynamics of the base region.

SIPOS (Semi-Insulating Polycrystalline Silicon) displayed promising results when used in selective contacts fabricated at SONY [1]. SIPOS is a silicon oxide (SiO_x , $x < 2$) formed by CVD (Chemical Vapor Deposition) using a reaction between silane and nitrous oxide. The film properties, which are sensitive to the oxygen content, can be controlled by varying the N_2O/SiH_4 ratio in the deposition ambient. Chapter 5 will provide background information on this family of oxides and will review the techniques and apparatus developed at Stanford to achieve uniform deposition of these films.

As was previously mentioned, J_o the saturation current, is a figure of merit for selective contacts indicating the degree to which minority carrier transport is suppressed, a lower J_o indicating greater blocking action. By measuring the base current in bipolar transistor test structures, one can experimentally determine the J_{oe} for various emitter + contact structures. The test structures and measurement procedures used to extract parasitics from the base current to yield J_{oe} is detailed in Chapter 6 and associated appendices. Chapter 7 presents the values of J_{oe} obtained for SIPOS and POLY (POLYcrystalline silicon) contact structures. For the POLY contacted emitters, annealing temperature and the presence of a chemically grown interfacial layer are shown to be the key factors that influence the values of J_{oe} that can be obtained. SIPOS contacts are shown to be equally vulnerable to J_{oe} degradation if excessively high annealing temperatures are used, but they apparently do not rely on the deliberate introduction of an interfacial oxide. These observations are given semi-quantitative corroboration using the modeling concepts introduced in Chapter 3.

The remainder of the discussion, focusing on the possible transport mechanisms in SIPOS films, is more speculative in nature since the microstructure of the film

is not yet well understood. Determination of the effective bandgap of SIPOS using C-V flatband voltage measurements are proposed. Chapter 8 provides a summary of this investigation and suggests guidelines for future work.

Chapter 2

Device Applications of Selective Contacts

2.1. Introduction

In order to be able to control device potentials and currents from external terminals, it is usually desired that the potential drop across the contact be smaller than internal voltage drops. If the contact also possesses no photovoltaic effect and a linear I - V characteristic, it is referred to as an “ohmic” contact [1]. Bipolar devices are sensitive to the influence of contacts on the pattern of minority carrier recombination. Ohmic contacts that are optimized for low majority carrier series resistance may have deleterious effects on overall device performance due to the potential of interaction of the contact with the minority carrier species. The relative impact of this interaction depends to a great extent on device geometry, operating point and frequency. Control over this interaction represents an additional design consideration in bipolar devices which is not present in majority carrier devices.

In this chapter, low frequency application of selective contacts to two bipolar structures — the solar cell and the bipolar junction transistor (BJT) are presented. These examples will serve to illustrate the performance enhancements that accrue from the use of such contacts. A quantitative analysis of the low frequency behavior of these selective contacts is presented in the next chapter. Before proceeding to the device applications, a brief discussion of contact structures is presented.

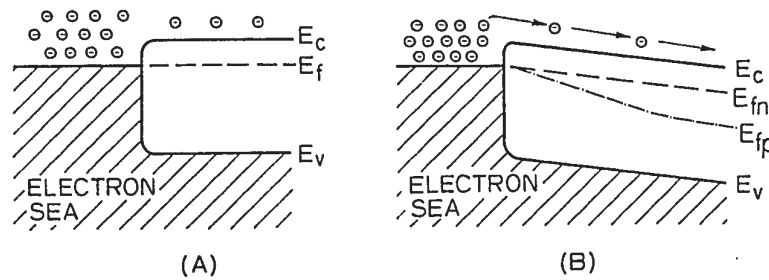


Figure 2.1: “ideal” N -type contact in thermal equilibrium (A) and under bias with hole injection (B)

2.2. Ideal Contacts

Metals, because of their characteristic high conductivity and ease of fabrication, are most commonly used to contact semiconductor devices, thereby forming a metal-semiconductor junction. Since the contacts to P -type regions behave in an analogous manner, only N -type contacts will be discussed. In an “ideal contact”, the work function of the metal matches that of the semiconductor so that no surface barriers are formed, and flat band conditions prevail in the bulk. The metal does not alloy with or modify the semiconductor through doping or by creating defect states. Furthermore, no interfacial layers are present, and the current density is assumed low enough that space charge effects are not important. Since the Fermi energy in the metal is at least several electron volts, it will be assumed that the conduction band of the metal will form a continuous “sea of electrons” with the silicon valence band (although the use of semimetals suggests some intriguing possibilities, they will not be discussed here). The band diagram for the ideal N -type contact would resemble Fig. 2.1, where some rounding of the band edge discontinuities may arise from image force lowering and carrier “spillover” from the metal. For the majority carriers (in this case electrons), the transport is limited by the bulk semiconductor resistivity.

Minority carriers (in this case holes) flow in an analogous manner, but since

the thermal equilibrium density is many orders of magnitude below that of the majority carriers, they do not play a significant role in unipolar contacts. If excess minority carriers are present however, due to the presence of a nearby injecting contact or due to photogeneration, the excess hole density will tend to relax to the equilibrium value. The relevant time constant for this process in the metal is the thermal relaxation time for hot holes. The corresponding parameter in the semiconductor is the hole recombination lifetime which is usually many orders of magnitude larger than the relaxation time in the metal. The metal therefore tends to act as a recombination sink for minority carriers due to their short lifetime in the metal. The flux of holes at the contact interface can be phenomenologically characterized by S_c , the contact recombination velocity for minority carriers, i.e.

$$J_p = q \cdot p'(0) \cdot S_c \quad (2.1)$$

where J_p is the hole current density, $p'(0)$ is the excess hole density at the contact interface and q is the electronic charge. Although an accurate value for S_c is difficult to extract from device measurements, the values tend to be quite large. A value of $3 \times 10^5 \text{ cm/sec}$ has been quoted by deGraaf for a sintered $Al-Si$ contact [2]; others have used values from several 10^6 cm/sec to $1 \times 10^7 \text{ cm/sec}$ [3], presumably because this is the kinetic limit imposed by the thermal velocity of the carriers. As will become clear from the analysis in Chapter 3, once S_c becomes large compared to the diffusion velocity of minority carriers, its exact value is of little consequence. Since low resistivity metallic contacts act as minority carrier sinks, approximating S_c by the thermal velocity usually suffices for the analysis of conventionally contacted structures.

2.3. Practical Contacts

Since most metals used for contacts have a work function greater than that of N^+ silicon, they tend to deplete the surface region in the silicon. The band diagram then resembles Fig. 2.2. The barrier height does not follow simply from

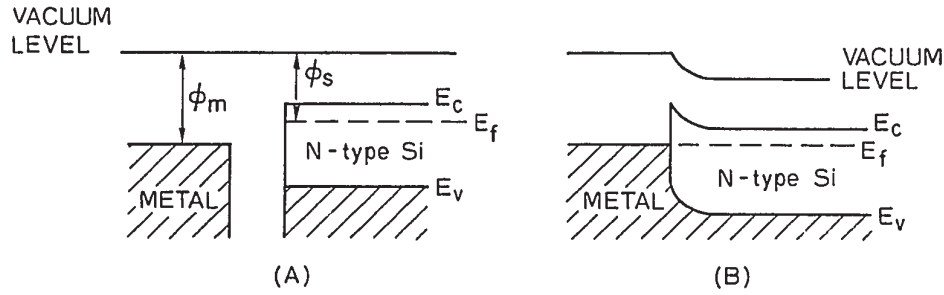


Figure 2.2: *N*-type contact where $\phi_M > \phi_S$

the difference in work functions as the influence of surface states can not usually be neglected, at least for covalent semiconductors such as silicon [4]. The presence of the depleted region inhibits electron transport; the effective contact resistance that results depends on several variables, but is most sensitive to the doping level, sintering treatment, and the choice of contact metal. In some cases, the presence of a thin interfacial layer can alter contact properties depending upon the method used to deposit the metal. For practical low resistance contacts to *N*-type silicon, heavy doping concentrations are required ($> 1 \times 10^{20} \text{ cm}^{-3}$). The resulting reduction in the barrier strength (height and width) has the important consequence that tunneling currents become appreciable and contact resistance falls sharply.

Since most contacts receive some type of sinter either to improve contact properties through alloying (e.g. silicide formation) or to anneal surface states in MOS devices, the planar interface is somewhat idealized. However, a linear boundary condition where the flux is proportional to the excess hole density via S_c , the contact recombination velocity, remains a useful concept. Selective contact structures which block minority carrier transport while maintaining low series resistance can possess values of S_c that are orders of magnitude smaller than for metallic contacts, hence the equivalent terminology “low surface (contact) recombination velocity contacts.”

2.4. The Solar Cell

Solar cells are energy converters transforming incident radiant energy from the sun into electrical potential energy. One convenient mechanism that can be used for this conversion process is to invest the energy of the incident photon into the potential energy of a hole–electron pair in a semiconductor. Semiconductors are used because the energy of pair formation (essentially the bandgap energy) can be chosen to match the energy spectrum that is available from incident sunlight. The photogenerated carriers can be highly mobile, and though metastable, are sufficiently long-lived to permit their separation before they recombine.

In order for a semiconductor solar cell to be an efficient energy converter, there are three universal characteristics it must possess [5]:

(1) Efficient photogeneration: Incident photons with energies above the bandgap must result in the photoproduction of hole–electron pairs. The use of anti-reflection coatings and surface texturizing techniques [1] are common techniques employed to ensure that the incident light is tightly coupled to the bulk of the cell where generation takes place. The bandgap of the semiconductor should also be chosen to balance the energy loss due to thermalization of hot carriers which result from the more energetic photons against the loss of energy in photons with sub-bandgap energies (which can not contribute to photoproduction). Direct gap materials are also preferred since carrier pair generation does not require the participation of phonons and is hence a more efficient process in the sense that the absorption coefficient is higher than in indirect gap materials.

(2) Efficient collection of carriers: Holes and electrons must be separated and transported to the external circuit before they recombine. The potential energy a hole–electron pair that recombine is lost to the lattice as heat. Even radiative recombination is essentially an irreversible loss of energy due to the weak absorption near the bandgap and due to geometrical factors. Efficient collection is promoted by diffusion lengths longer than the collection length and by physical segregation

of carriers away from regions of high recombination.

(3) Maximization of the terminal voltage: The maximum energy available to do useful work outside of the cell is given by the separation of the quasi-fermi levels (imrefs). Unfortunately, this is also the driving force behind various parasitic mechanisms which result in energy loss through carrier recombination. Some of the available work is also spent in transporting the carriers out of the generation region to the collecting contacts. This arises from the need to establish carrier gradients (or more generally, a gradient in the electrochemical potential) to motivate carrier flow as well as the ohmic drops that are associated with current flow in various parts of the cell. To reduce these energy losses, it is important to minimize competing recombination mechanisms by ensuring high lifetimes in regions of large minority carrier concentrations or by excluding them from regions such as the contacts and contact diffusions where their lifetime is short. The other losses can be minimized by optimization of the cell geometry to eliminate the bottlenecks that restrict carrier flow and by reducing internal cell, contact and metallization grid resistances.

2.5. The Point Contact Cell

The development of selective contact structures for a high concentration high conversion efficiency solar cell has been the primary objective of this investigation. The cell under consideration is a *P-I-N* structure fabricated from high lifetime float zone silicon. The diffused contacts, which act as collection regions, form a grid of regularly spaced point contacts. One version of the cell places the *P*-type and *N*-type diffusions on opposite sides of the cell (Front Surface Cell or FSC); the other, in which both types of contacts are on the backside (BackSide Cell or BSC) (see Fig. 2.3) is currently under development. Although crystalline silicon possesses some fundamental disadvantages for photovoltaic applications when compared to some of the compound semiconductors, its technology is the most highly refined. The greater flexibility in cell design that results allows many of the potential dis-

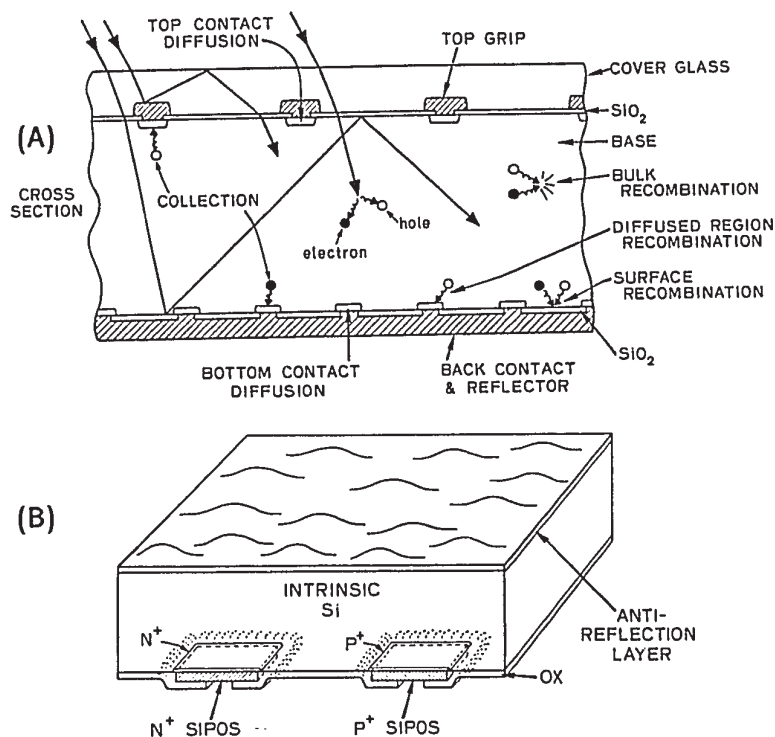


Figure 2.3: cross section of FSC (A) and BSC (B) point contact cells

advantages to be circumvented. Initial experimental efficiencies of 20% have been obtained for the BSC configuration, and conversion efficiencies in excess of 26% are believed to be feasible through refinements of currently available technology. This section will discuss how the design philosophy used in the cell design addresses the aforementioned requirements for high conversion efficiency solar cells and how selective contacts in conjunction with other cell refinements can be used to further increase the cell efficiency.

Since silicon is an indirect gap semiconductor, it is only weakly absorbing for energies near the bandgap. Unfortunately, this represents a significant portion of the available energy spectrum. To efficiently utilize these photons, the optical path length must be extended to compensate for the weaker photogeneration. To

accomplish this, the backside of the cell is made highly reflective at these energies. The frontside of the cell can also be texturized so that the angle of incidence of the light reflected from the back side will be greater than that required for total internal reflection. The net effect is to trap the light inside the cell thereby increasing the probability of photogeneration [5]. To obtain this backside reflectivity, the contact coverage fraction is reduced to a value much less than unity, since a sintered low resistivity contact tends to parasitically absorb incident photons, resulting in poor reflectivity. The use of an array of point contacts meets this requirement.

The cell is processed in an ultraclean facility; the measured bulk lifetime after high temperature processing has been measured at several milliseconds. The resulting diffusion length greatly exceeds the cell thickness so that carrier losses due to bulk defect recombination are minimal. However, a long diffusion length alone is not sufficient to ensure good collection efficiency. In a conventional cell, the full area backside ohmic contact acts as a recombination sink for carriers generated in its vicinity, thereby resulting in a loss in collection efficiency regardless of the bulk properties. In the BSC configuration where the region between the small point contacts is passivated by a high quality thermal oxide, recombination losses due to this mechanism are minimized. Collection efficiencies close to unity have been achieved.

The maximization of the terminal voltage for a given terminal current represents the final requirement. This involves a complex interplay of many factors; some of the tradeoffs which are involved can be appreciated by referring to Appendix A in which a mechanical analogy of the solar cell is presented. A simplified version which ignores this interdependence can be explained by referring to the band diagram of a $P-I-N$ cell in thermal equilibrium shown in Fig. 2.4a. The potential barriers resulting from the presence of the $P+$ and $N+$ diffusions serve to prevent the majority carriers from diffusing into regions of lower concentration. Under illumination, the barrier heights are reduced somewhat (because the diode is effectively forward biased), but are still effective in separating the photogenerated carriers. The electrons and holes are collected into the $N+$ and $P+$ regions respectively as shown in Fig. 2.4b.

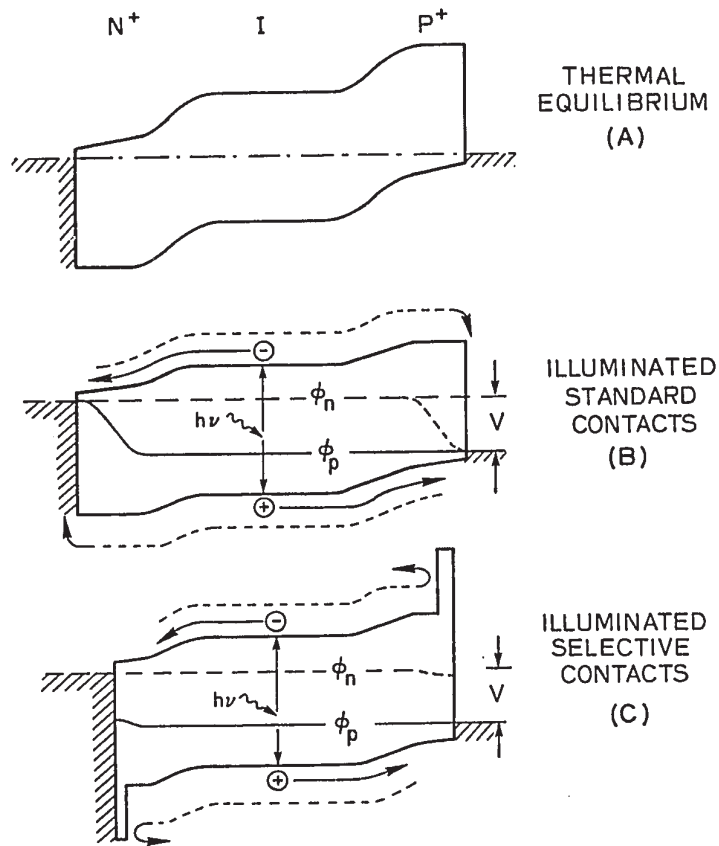


Figure 2.4: band diagrams of a *P-I-N* solar cell

This direction of carrier flow allows useful work to be done by the electron as it cascades down the potential V in the external circuit eventually to recombine with the hole. That the quasi-Fermi levels are shown separated by the amount V follows from the fact that photogenerated holes and electrons are present in excess of their thermal equilibrium values in the generation region. Although the illustration of flat Fermi levels is a simplification, the qualitative impact of contact recombination is not altered. The band diagram for the illuminated device is seen to be equivalent to a very short base forward biased *P-I-N* diode. Referring to Fig. 2.4b, it can

be seen that the reduced barrier heights that are a consequence of forward bias result in a greatly enhanced presence of electrons and holes in regions where they are minority carriers. The enhancement is given by the Boltzman factor. The presence of these carriers in such high densities is not in itself detrimental (their density is still far less than what it is in the active generation regions). However, the lifetime in these heavily doped regions tends to be quite small and the ohmic contact interface presents a very high effective contact recombination velocity. The presence of large numbers of carriers in regions of high recombination results in large currents. This direction of carrier flow also opposes the photocurrent and the energy in these recombining carriers is lost to the silicon lattice or to the contact mostly as heat. If the height of the barriers could be increased, this loss could be dramatically reduced (approximately one decade per $60mV$), but there is a limit to what can be achieved with homojunctions. Doping levels above $1 \times 10^{18} cm^{-3}$ don't seem to increase the minority carrier barrier heights; instead, the primary effect on the minority carriers is to reduce their recombination lifetime, ostensibly through Auger recombination.

If it were possible to place strategic barriers in the valence and conduction bands as shown in Fig. 2.4c, then the possibility arises of excluding minority carriers from these defective regions. The loss of photocurrent to contact recombination could be greatly reduced. Another way of seeing this is to view the additional barriers as “checkvalves” that redirect errant carriers. In order for this stratagem to prove effective, both types of barriers are needed, i.e. for both the conduction and valence bands. It is also important to realize this barrier in the minority carrier band without affecting majority carrier transport. High concentration cells with their attendant high current densities are especially vulnerable to contact resistance, and for the point contact cell, where the contact coverage fraction is only several percent, it is imperative that the minority carrier blocking action be obtained without introducing excessive contact resistance.

J_o , the saturation current associated with the injection of minority carriers into

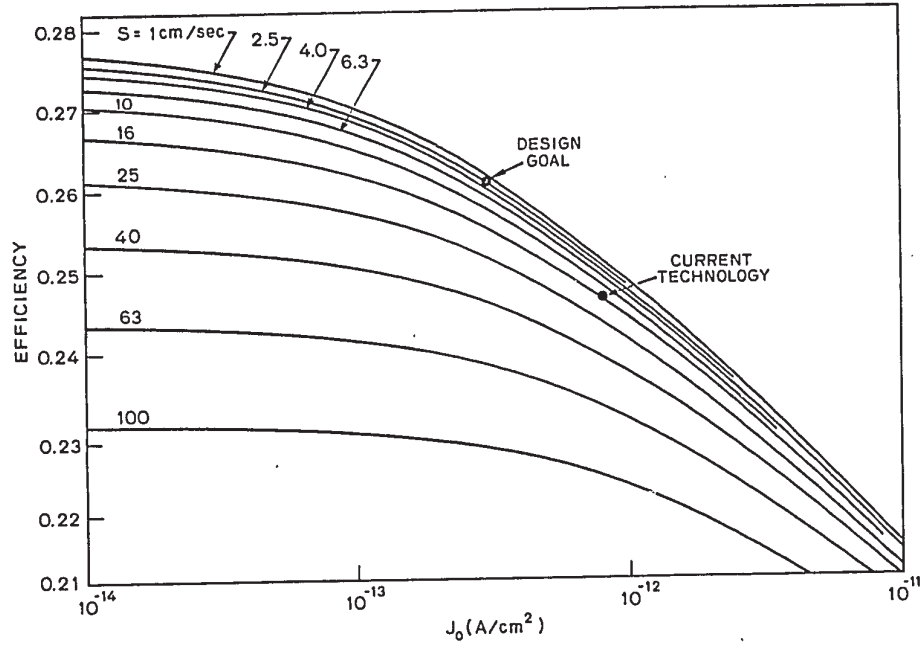


Figure 2.5: energy conversion efficiency for point contact cell vs. contact region J_o , parameter is recomb. velocity at passivating oxide-Si interface, from Swanson [5]

the contact diffusions, reflects the degree to which the minority carrier flow is obstructed. For conventional homojunction diffusions, the combined effects of Auger recombination and band-gap narrowing limit J_o to values above $1 \times 10^{-13} \text{ A/cm}^2$, lower values requiring very deep diffusions. Present cell technology can only offer values of $8 \times 10^{-13} \text{ A/cm}^2$, and because of this large value, contact recombination represents the dominant loss mechanism in present cells. The improvement in conversion efficiency that can result from reductions in J_o are shown in Fig. 2.5 which is reproduced from Swanson [5]. The parameter on the plot is the surface recombination velocity, S_{ox} , in the intercontact region which is passivated by a high quality thermal oxide. Reducing J_o by two orders of magnitude, down to $1 \times 10^{-14} \text{ A/cm}^2$, can significantly improve efficiency, although the most impressive gains obtain if the current geometry is revised to take maximum advantage of the low J_o structures

and if a synergistic reduction of S_{ox} can also be achieved. In Chapter 6, data will be presented for N -type selective contacts which have met this goal.

2.6. Bipolar Transistors

The basic principles of bipolar junction transistor (BJT) action are well known [6,7]. It will be assumed that the reader is already familiar with standard terminology relating to the device. Although often thought of as a current driven device, the BJT is, like its cousin the MOSFET, a voltage controlled barrier device. Under static conditions, like the MOSFET, an “ideal” BJT would require no base current. Any current that does flow at 0 Hz arises as a result of various parasitic mechanisms that can in principle be made negligibly small. Minimizing the base current drive has the desirable effect of increasing the power gain and input impedance of the device. Conversely, degradation of beta has negative impact on these properties.

To identify the sources of base current, it is convenient to divide them into static and dynamic components. The use of the term “static” in this case implies that these components are still present at 0 Hz (DC); in general they have a frequency dependence. Both components are present for finite frequencies. The dynamic terms arise because of the internal rearrangement of charges associated with, for example the space charge layer capacitances and neutralization of injected minority carriers for time varying signals. The “static” terms which arise primarily from carrier recombination will be the subject of analysis in this section. The dynamic terms will be discussed in Chapter 4 where it will be shown that optimization of high frequency response relies on a different set of criteria.

2.7. Homojunction BJT

The band diagram for a homojunction N - P - N BJT is shown in Fig. 2.6. When the emitter–base junction is forward biased, a net injection of electrons from the emitter into the base results. These carriers are eventually collected by the electric

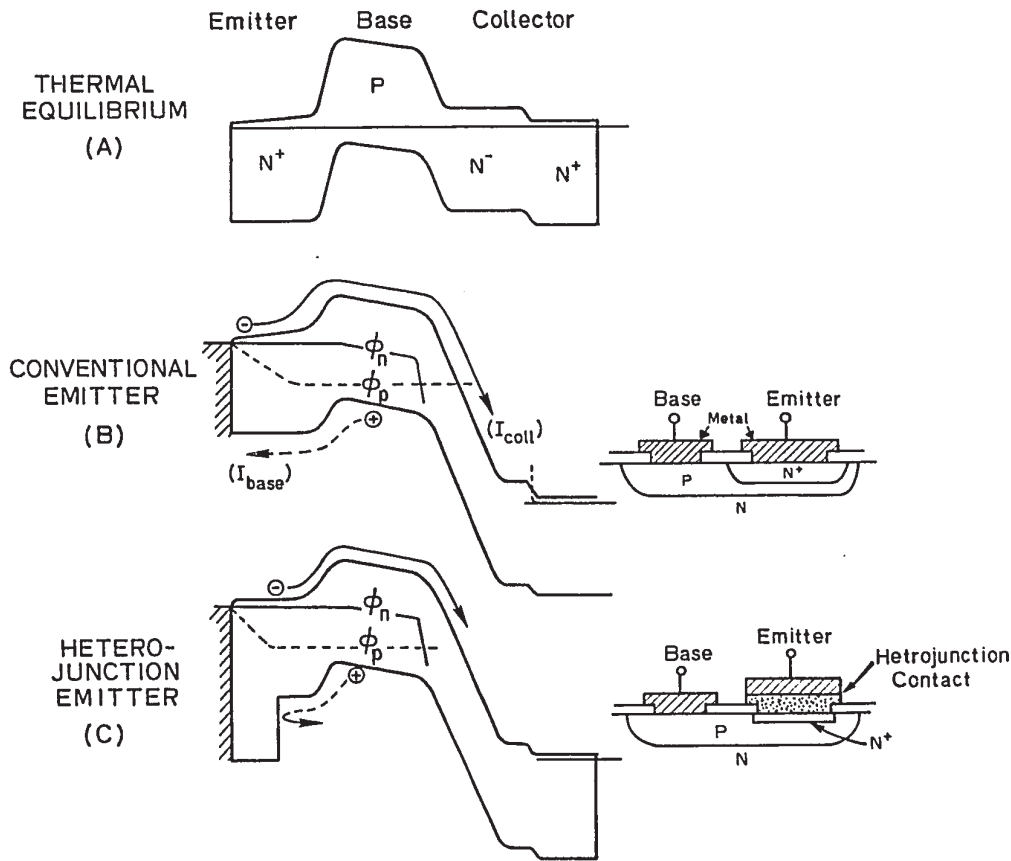


Figure 2.6: Band diagrams for BJT

field at the collector-base junction unless they recombine somewhere along the transport path. This flow of electrons (shown by the solid path) which constitutes the collector current, is then available to do work outside the device. It is this flow that is controlled by modulating the barrier height in the base via the base-emitter voltage. The BJT is in essence, a voltage controlled barrier device.

This picture is not yet complete, for an analogous situation occurs for holes. In thermal equilibrium, the holes rest in a potential well formed by the built-in potentials arising from the diffusions. Reverse biasing the base-collector junction

only increases the barrier height toward the collector. With the application of a forward bias to the base-emitter junction, the barrier is lowered toward the emitter resulting in an enhancement of hole density in this region. A flux of holes from the base to the emitter is required to replenish those that are lost to recombination to maintain steady state. If the recombination rate in this region can be minimized, this reverse injected current can be reduced.

From this description, it can be seen that under static conditions, base current is required to support recombination in various regions of the device. Which component dominates is dependent upon device geometry and operating point. For high voltage power devices with their deep emitters and wide bases, bulk recombination in the base may be the most significant. For modern high speed designs where the base is made narrow to minimize the transit time, reverse injection into the emitter tends to dominate. For both types of devices, operation at low current levels may lead to space charge recombination in the bulk or at the surface becoming the principle source of base current. It is only for the class of devices in which reverse injection dominates that selective contacts can offer significant performance enhancements.

2.8. Idealized BJT

A qualitative understanding of the influence of the emitter contact on BJT characteristics can be obtained by analyzing an idealized one-dimensional transistor with uniformly doped base and emitter regions. In practice, doping gradients can have a significant impact on device parameters, but the qualitative trends are still preserved in the simpler idealized model. Results of quantitative modeling of realistic device profiles will be presented in Chapter 4 where the 0 Hz results can be viewed in context of the overall frequency performance of the device.

Consider the transistor depicted in Fig. 2.7 where W_e and W_b denote the undepleted widths for the emitter and base. N_e and N_b are the corresponding doping

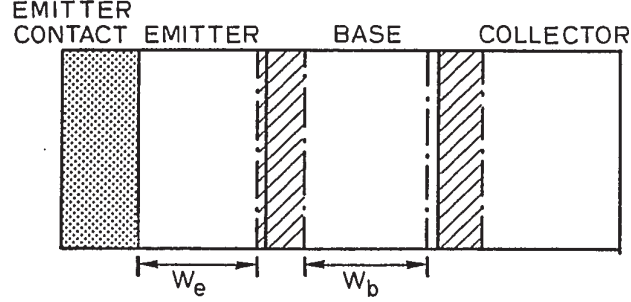


Figure 2.7: Ideal one-dimensional BJT

densities. The collector current can be written as

$$I_c = A \cdot J_{oc} \cdot (e^{V_{be}/V_T} - 1); \quad V_T = kT/q \quad (2.2)$$

where A is the active area, usually of the emitter. J_{oc} , the collector saturation current, is given by

$$J_{oc} = \frac{qn_{ie,b}^2 D_{n,b}}{N_b L_{n,b} \tanh(W_b/L_{n,b})} \quad (2.3)$$

where $D_{n,b}$ and $L_{n,b}$ represent the diffusion coefficient and length for electrons in the base and $n_{ie,b}^2$ represents the effective thermal equilibrium pn product in the base due to heavy doping effects. For $L_{n,b} \gg W_b$ and lightly doped base regions ($N_b < 10^{17} \text{ cm}^{-3}$), this reduces to

$$J_{oc} = \frac{qn_{io}^2}{N_b} \frac{D_{n,b}}{W_b} \quad (2.4)$$

or in an alternate form by

$$J_{oc} = \frac{qn_{io}^2}{G_b}; \quad G_b = \frac{W_b N_b}{D_{n,b}} \quad (2.5)$$

where G_b is defined as the base Gummel number [8]. The Gummel number is a measure of the effective majority carrier charge per unit area divided by an effective minority carrier mobility. Because of the inverse relationship, a large Gummel number is associated with a small saturation current; equivalently, a region with a large Gummel number is more effective in impeding minority carrier transport.

The component of base current, due to neutral base recombination is given by:

$$I_{b1} = A \cdot J_{oe} \cdot \left[1 - \frac{1}{\cosh(W_b/L_{n,b})} \right] (e^{V_{be}/V_T} - 1) \quad (2.6)$$

$$\approx A \cdot J_{oe} \cdot \frac{W_b^2}{2L_{n,b}^2} (e^{V_{be}/V_T} - 1) \quad (2.7)$$

for $L_{n,b} > W_b$ The reverse injected component is given by:

$$I_{b2} = A \cdot J_{oe} \cdot (e^{V_{be}/V_T} - 1) \quad (2.8)$$

where V_{be} is the base-emitter bias. J_{oe} , the emitter saturation current, can be expressed as:

$$J_{oe} = \frac{qn_{ie}^2 D_{p,e}}{N_e L_{p,e} \tanh(W_e/L_{p,e})} \quad (2.9)$$

In this expression, $L_{p,e}$ is the effective hole diffusion length due to SRH and Auger recombination; n_{ie}^2 is the effective pn product taking bandgap narrowing (BGN) effects into account. Actually BGN is not the only mechanism responsible for the enhanced pn product. Changes in the effective mass due to alterations in the density of states also play a key role; however, all the effects associated with heavy doping will be assumed to be accounted for by the n_{ie}^2 term and will be collectively referred to as BGN. An equivalent method for accounting for BGN effects is to keep n_{io}^2 , the pn product for intrinsic silicon, and redefine a new quantity, the effective doping density

$$N_{e,eff} \equiv N_e \cdot \left(\frac{n_{io}^2}{n_{ie}^2} \right) \quad (2.10)$$

This is the doping level that in the absence of BGN would produce the same equilibrium hole density [9]. As $N_{e,eff} < N_e$, it can be seen that the effect of BGN is to decrease the effective barrier to reverse injection as seen by the holes from the base. This definition allows the reverse injected current to be expressed as:

$$J_{oe} = \frac{qn_{io}^2 D_{p,e}}{N_{e,eff} L_{p,e} \tanh(W_e/L_{p,e})} \quad (2.11)$$

This can be cast into a form similar to that for the collector current:

$$J_{oe} = \frac{qn_{io}^2}{G_e} \quad (2.12)$$

where G_e , the emitter Gummel number is given by:

$$G_e = \frac{N_{e,eff} L_{p,e} \tanh(W_e/L_{p,e})}{D_{p,e}} \quad (2.13)$$

Large values for G_e imply more effective barriers to reverse injection. Since other components of the base current are assumed to be insignificant, the static current gain, the ratio of I_c to I_b , is simply

$$\beta_o \equiv \frac{I_c}{I_b} = \frac{I_c}{I_{b1} + I_{b2}} \quad (2.14)$$

$$= \left(\frac{W_b^2}{2L_{n,b}^2} + \frac{J_{oe}}{J_{oc}} \right)^{-1} = \left(\frac{W_b^2}{2L_{n,b}^2} + \frac{G_b}{G_e} \right)^{-1} \quad (2.15)$$

Equivalently, the overall transport factor (common base current gain) is

$$\alpha_o \equiv \frac{\beta_o}{\beta_o + 1} = 1 - \frac{W_b^2}{2L_{n,b}^2} - \frac{J_{oe}}{J_{oc}} = 1 - \frac{W_b^2}{2L_{n,b}^2} - \frac{G_b}{G_e} \quad (2.16)$$

It can be seen that transport losses arise from the loss of base transparency and the non-unity injection efficiency due to the reverse injection of holes.

2.9. Shallow Emitter Devices

The speed at which the collector current can be modulated by a change in base current is limited by the delays associated with the charging of the parasitic and junction capacitances as well as the transit times required for the carriers to traverse the neutral base and collector depletion layers [6]. Since the transit time for a uniformly doped base under low level injection is given by $W_b^2/2D_b$ in the diffusive transport regime (i.e. non-ballistic), a potential exists for improving device speed by reducing W_b . As the base is made narrower, neutral base recombination becomes less of a factor, and beta will approach the reverse injection limited value given by

$$\beta_{o,max} = \frac{G_e}{G_b} \quad (2.17)$$

or simply the ratio of emitter to base Gummel numbers. This represents an upper limit to the achievable current gain (discounting collector multiplication).

Since the neutral base is usually formed as the result of sequential base and emitter diffusions or implants, process controllability dictates that the emitter depths be comparably scaled as base widths are made narrower. Otherwise, process induced variations in the base-emitter junction depth will represent an increasing percentage of the neutral base width and yields will suffer as a result. An example of this problem is illustrated by the reduced yields of linear circuits that employ narrow base superbeta transistors.

Reducing the emitter depth, however, adversely impacts the value of the emitter Gummel number

$$G_e = \frac{N_{e,eff} L_{p,e} \tanh(W_e / L_{p,e})}{D_{p,e}} \quad (2.18)$$

If $W_e < L_{p,e}$ then G_e approaches the value

$$G_e = \frac{N_{e,eff} W_e}{D_{p,e}} \quad (2.19)$$

The maximum current gain is then

$$\beta_{o,max} \leq \frac{1}{G_b} \frac{N_{e,eff} W_e}{D_{p,e}} \quad (2.20)$$

If as a rough guide, W_e is scaled proportionately to W_b , then

$$\beta_{o,max} \leq \frac{D_{n,b}}{D_{p,e}} \left(\frac{N_{e,eff}}{N_{b,eff}} \right) K \quad (2.21)$$

where K is a constant of order unity and $N_{b,eff}$ is the effective doping density in the base if BGN effects become significant. Since the base doping N_b must increase to prevent punchthrough and to keep the base sheet resistivity from rising excessively, $N_{b,eff}$ increases as W_b is decreased until BGN effects dominate. Since $D_{n,b}$ also falls with rising N_b and since $N_{e,eff}$ remains relatively constant because of BGN effects which are more pronounced in the heavily doped emitter, the current gain of the device falls as base widths are decreased; whether it falls precipitously depends upon the scaling rule followed, but clearly the inevitable reduction of static current gain will significantly impact the tradeoffs involved in device optimization.

2.10. Selective Contacts

For thin emitters, the large increase in reverse injected current arises because of the presence of the ohmic contact in close proximity to the injecting base-emitter junction. The metal-silicon interface presents a large effective surface recombination velocity for holes (e.g. $> 3 \times 10^5 \text{ cm/sec}$ for *Al* on $N^+ \text{ Si}$ [2]). In the limit that the emitter is so shallow that transport is limited by S_c , the contact recombination velocity,

$$J_{oe} = \frac{qn_{io}^2}{N_{e,eff}} S_c > 10^{-11} \text{ A/cm}^2 \quad (2.22)$$

where the deGraaf value for S_c and a value of $1 \times 10^{18} \text{ cm}^{-3}$ is chosen as an upper limit for the effective doping density in the emitter due to BGN. This results in a value for emitter Gummel number of:

$$G_e = \frac{qn_{io}^2}{J_{oe}} < 3 \times 10^{12} \text{ sec/cm}^4 \quad (2.23)$$

if a value for S_c of $3 \times 10^5 \text{ cm/sec}$ is assumed. If the thermal velocity were used for S_c , G_e would decrease by roughly another order of magnitude. Since typical values for G_e are of the order of several 10^{13} cm/sec^4 , this can represent an order of magnitude reduction in the static current gain of the device if the base Gummel number were kept constant. If it is increased the degradation would be more severe.

Since the problem lies with the large recombination of holes at the contact interface, if a barrier could be placed in the valence band as shown in Fig. 2.6c to repel holes from this defective region, recombination could be greatly reduced. The emitter diffusion is made shallow from process control constraints, so bulk recombination in the emitter is also reduced. The analysis of thin transparent emitters presented in the next chapter will quantify some of these observations.

The use of selective contacts also presents new possibilities for device optimization. Through their use, it is possible not only to counter the degradation associated with shallow emitters, but also to substantially enhance the current gain even compared to devices with deeper "standard" emitters contacted with conventional ohmic

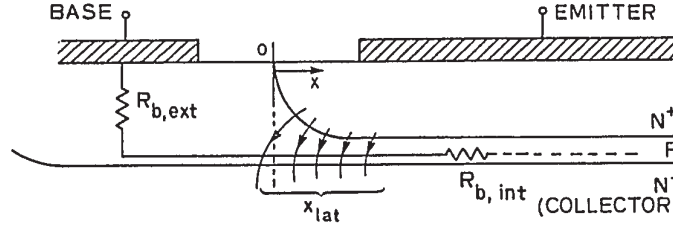


Figure 2.8: cross-section of emitter sidewall region

contacts. Such high values of static current gain may not directly impact the high frequency beta which is dominated by the dynamic components of base current, but if attained will allow other tradeoffs to be made that in turn improve device speed.

2.11. Current Crowding

As lateral dimensions are scaled, devices are required to operate at greater current densities. Parasitic capacitances tend to scale more slowly than device area due to fringing components, differing scaling requirements for interconnects, and higher doping levels that lead to increased specific junction capacitances. Therefore, to avoid speed penalties, the current drive capability per unit area of active device must increase. Other devices that operate in the high current density regime include power devices which are called upon to handle large absolute values of terminal currents.

The three dimensional nature of the BJT becomes apparent when potential differences due to ohmic drops in the device result in non-uniformities in the current flow. Fig. 2.8 shows a cross section of the emitter sidewall region of a typical BJT. For low voltage devices, the base is sufficiently narrow that recombination in the neutral base can be neglected. Base current that supports reverse injection into the emitter encounters ohmic drops as it flows through the extrinsic ($R_{b,ext}$) and intrinsic ($R_{b,int}$) base resistances. $R_{b,int}$ which represents the pinched base resistance

is usually the dominant term due to the narrow active base and the lower doping densities that are encountered there. If these ohmic drops in the active base become comparable to the thermal voltage, the central region of the emitter furthest from the base contact will lose its forward bias and cease to contribute to the collector current. As the base current increases, more and more of the emitter shuts down. When this “current crowding” effect is severe, most of the collector current flows through the perimeter of the emitter as shown in Fig. 2.8. The lateral crowding distance over which the emitter is still effectively injecting is given by [10] :

$$x_{lat} = \sqrt{\frac{2V_T\beta_o}{\rho_{BS}J_c(0)}} \quad (2.24)$$

where $J_c(0)$ is the collector current density at the perimeter of the emitter, V_T is the thermal voltage kT/q , and ρ_{BS} is the pinched base sheet resistivity, including conductivity modulation effects if significant. Since this peripheral injection makes use of only a small fraction of the available injecting area, it represents a poor utilization of emitter area. For a given terminal collector current, $J_c(0)$ is much larger when crowded into the perimeter than when uniform injection applies. Therefore high injection effects such as conductivity modulation and base push out (Kirk effect) which degrade current gain and frequency performance occur at terminal currents far below what would be expected on the basis of uniform injection. The increase in path length for the minority carriers due to the fringing nature of the current streamlines in the perimeter region (see Appendices H,I) increases the effective base width which degrades device performance in a similar manner.

Current crowding becomes a limitation when $x_{lat} \approx L$ where L is a characteristic lateral dimension of the emitter. For high speed bipolar devices, the emitter current can not be scaled with area to keep $J_c(0)$ constant because of a severe speed penalty. Therefore $J_c(0)$ scales roughly as L^{-2} . Therefore

$$\frac{x_{lat}}{L} \sim \sqrt{\frac{\beta_o}{\rho_{BS}}} \quad (2.25)$$

If β_o is reverse injection limited, then scaling N_b to keep the base sheet resistance

constant will prevent introducing current crowding as a limiting factor, at least within the limits of validity of this one dimensional analysis.

Selective contacts, with their potentially large beta enhancements, relax this base sheet resistance requirement. This freedom can be used to find a more optimal choice for the base doping level, or for a given base sheet resistivity, to increase the current level at which current crowding occurs. For power devices, this means that the emitter area is more efficiently utilized, permitting the same size die to handle much larger terminal currents. For output devices in power linear integrated circuits such as voltage regulators, it may be possible to design them to occupy less of the chip area without any penalty in the current handling capability.

2.12. A Circuit Application

It was pointed out in a previous section that current gains could be maintained at reasonable values despite the use of a larger Gummel number in the base. Yet another possibility is to keep the base parameters unchanged and utilize emitters with selective contacts to obtain devices with large current gains.

The emitter coupled circuit shown in Fig. 2.9 is commonly used as the input stage in differential op amps. There are several figures of merit associated with the op amp, many of which are determined by the input stage. Quiescent current must be supplied to the base terminals of the input transistors $Q1$ and $Q2$. If the circuit is to be used with high impedance sources, this current should be minimized. Small input bias currents translate into smaller equivalent input offset voltages due to input impedance mismatches or due to input offset currents. Since $I_b = I_c/\beta_o$, base currents are minimized by using small collector bias currents and by using devices with very large current gains. The standard fabrication method is to use a separate emitter diffusion cycle into the same base tub as the standard devices which possess much lower β_o 's. This emitter is driven in deeper resulting in a very narrow neutral

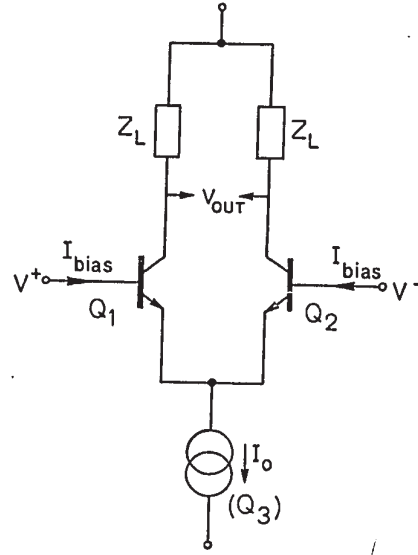


Figure 2.9: generalized differential amplifier stage

base region, hence a smaller base Gummel number. Since

$$\beta_{o,max} \leq \frac{G_e}{G_b}$$

and G_e for the deeper emitter is similar to the standard emitter, this translates into very high gain devices. Unfortunately, this narrow base limits the base-collector voltage to the order of a volt and its larger sheet resistance can compromise noise performance. The voltage restriction can be circumvented by using a bootstrap circuit that forces the emitter and collector to track common mode variations in the input signal [11] at the cost of added complexity. A more fundamental disadvantage of the thin base device is the high output conductance, or equivalently the low Early voltage, due to base width modulation. The small signal gain of a differential stage is given by

$$A_v = g_m (r_L || r_o)$$

where r_L is the effective load resistance and r_o is the output resistance of the active devices ($Q1, Q2$). A low value for r_o limits the gain available in the input stage.

This may require the use of more stages to meet the gain requirement which tends to complicate the task of compensation. This drawback can be sidestepped by using circuit topologies such as the cascode configuration, again at the cost of additional circuit complexity. By using selective emitter contacts in devices $Q1$ and $Q2$, the superbetas can be achieved without a reduction in r_o or sacrificing the collector-base breakdown voltage. If low input bias currents are not the primary objective, some of the beta gain can be traded off for a higher quiescent collector current. The increase in g_m not only improves the gain and common mode rejection, but also improves the frequency response of the input stage by reducing r_x .

Another figure of merit for the differential stage is the common mode rejection ratio, defined as the ratio of differential to common mode gain [12]. For the circuit shown, it is simply

$$CMRR = 1 + g_m r_{o3}$$

where r_{o3} is the output impedance of the current source. For “standard” devices where the current gain must be limited to maintain breakdown voltages (BV_{CEO}), a different base profile with a larger base Gummel number could be used to reduce beta to an acceptable value. This device would possess a much greater output resistance which would make it suitable for use in the current source application. The reduced base sheet resistance would also make it a likely candidate for the output power devices where current crowding is a consideration. A process sequence which incorporates SIPOS or POLY selective contacts suited to linear integrated circuit fabrication is outlined in Appendix B. The details of this process will become clearer after these contact structures are introduced in Chapter 7.

2.13. Summary

The use of selective contacts with their low values of J_{oe} can circumvent many of the disadvantages associated with the use of shallow emitter structures. This can translate into higher energy conversion efficiencies for solar cells. For BJT's,

the relaxation of the constraints on the base Gummel number afforded by the large emitter Gummel numbers attainable with selective contacts allows greater freedom to tailor the base profile to suit various applications.

Chapter 3

Low Frequency Behavior of Emitter Structures

3.1. Introduction

Chapter 2 has illustrated some of the low frequency device performance enhancements that accrue through the use of selective contact structures. It remains to be seen how these structures can be realized in practice. Since barriers to minority carrier transport are necessary, these contacts generally call for the use of materials in addition to silicon in the contact structure. In the sense that heavy doping alters the material properties of silicon, a heavily doped emitter already comprises a heterojunction; hence all the structures discussed in this chapter are strictly speaking heterojunctions, the various cases reflecting additional levels of complexity when compared to the baseline homojunction structure. However, for the sake of convenience, the term “heterojunction” will be reserved for a structure incorporating two dissimilar semiconductors where the interfacial layers, if present, do not affect minority carrier transport.

Using the linearity of the transport equations, an equivalent circuit that models the various recombination currents in contact structures is developed. This model points out the interactions that exist between the various components and allows the development of general guidelines relevant to low J_{oe} structures. A simplified emitter modeling technique developed by del Alamo and Swanson [1] is introduced to permit quantitative evaluation of the low frequency performance of these contact

structures. Extension of these results to finite frequencies is presented in Chapter 4.

3.2. Lumped Parameter Equivalent Circuit

The linearity of the transport equations allows a representation of the minority carrier transport properties for various regions of the bipolar device in terms of an equivalent circuit. The discussion here will use terminology relevant to the $N+$ emitter since it is the calculation of J_{oe} , the emitter saturation current, that is of primary interest, but the concepts are applicable to transport in all regions of the device. The transport equations will be numerically solved later in this chapter, but the equivalent circuit provides insight not readily available from calculations alone.

In order to construct this equivalent circuit, we need elements to represent several types of structures encountered in the emitter and emitter contact. These will be integrated to form a representation for the general heterojunction structure. Various types of available contact structures will be seen to be special cases of this generalized structure.

3.2.1. Transport Equations

For N -type regions and contacts, we are interested in hole transport. In general, the hole current density is given by

$$\vec{J}_p = q\mu_p p \nabla \phi_p \quad (3.1)$$

$$\nabla \cdot \vec{J}_p = q(G - R) \quad (3.2)$$

where ϕ_p the hole imref (quasi-Fermi level) is referenced to the electron imref. For non-degenerate hole concentrations, we can write for low level injection

$$p = p_o e^{\phi_p/V_T} \quad (3.3)$$

where p_o is the local thermal equilibrium hole density. Using the effective doping

concept introduced in Chapter 2, we can write

$$p_o = \frac{n_{io}^2}{N_{eff}} \quad (3.4)$$

and therefore

$$p = \frac{n_{io}^2}{N_{eff}} e^{\phi_p/V_T} \quad (3.5)$$

Rearranging and subtracting unity allows us to define a new quantity

$$u \equiv \frac{pN_{eff}}{n_{io}^2} - 1 = e^{\phi_p/V_T} - 1 \quad (3.6)$$

Taking the gradient of both sides and solving for $\nabla\phi_p$ gives

$$\nabla\phi_p = \frac{V_T \nabla u}{(u + 1)} \quad (3.7)$$

Substituting back into eqn. (3.1) yields

$$\vec{J}_p = qp\mu_p V_T \left(\frac{\nabla u}{(u + 1)} \right) \quad (3.8)$$

If we define an effective diffusion coefficient for holes as

$$D_p \equiv V_T \mu_p \quad (3.9)$$

$$\vec{J}_p = \left(\frac{qn_{io}^2}{N_{eff}/D_p} \right) \nabla u \quad (3.10)$$

As is pointed out in [1], the driving force behind hole transport is u , the imref separation, or equivalently proportional to the pn product. The continuity equation

$$\nabla \cdot \vec{J}_p = q(G - R)$$

for dark characteristics reduces to

$$\nabla \cdot \vec{J}_p = -\frac{q(p - p_o)}{\tau_p} \quad (3.11)$$

and since

$$u = \frac{(p - p_o)N_{eff}}{n_{io}^2}$$

as can be seen by rearranging eqn. (3.6) we can write

$$\nabla \cdot \vec{J}_p = -\frac{qn_{io}^2}{N_{eff}\tau_p} u \quad (3.12)$$

where τ_p is the effective hole lifetime. To summarize,

$$\vec{J}_p = \frac{qn_{io}^2}{N_{eff}/D_p} \nabla u \quad (3.13)$$

$$\nabla \cdot \vec{J}_p = -\frac{qn_{io}^2}{N_{eff}\tau_p} u \quad (3.14)$$

These two equations suggest a representation in the variables $\vec{J}_p$ and u where u is the driving term and $\vec{J}_p$ is the response term.

3.2.2. Diffused or Bulk Regions

The analysis in this section applies to all bulk regions be they deposited layers such as POLY or whether they are formed by implantation or diffusion. These will be referred to collectively as “diffused” layers. To obtain an equivalent circuit for a diffused region, we can adopt a strategy that is used in the analysis of the base region of a BJT. Consider the diffused region shown in Fig. 3.1a which could represent the diffused emitter. Using the notation in [1], J_{pw} and J_{po} represent the hole currents entering and exiting the diffused region. A two port equivalent circuit using the variables u, J is shown in Fig. 3.1b, where it will be understood that J refers to hole currents.

Since the u, J representation is linear, the two port can be characterized entirely in terms of its two port parameters, in this case, the Y parameters.

$$\begin{bmatrix} J_p(W) \\ -J_p(0) \end{bmatrix} = \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (3.15)$$

By setting u_2 to 0 and measuring the current responses at ports 1 and 2 we can determine

$$Y_{11} = \left. \frac{J_1}{u_1} \right|_{u_2=0}$$

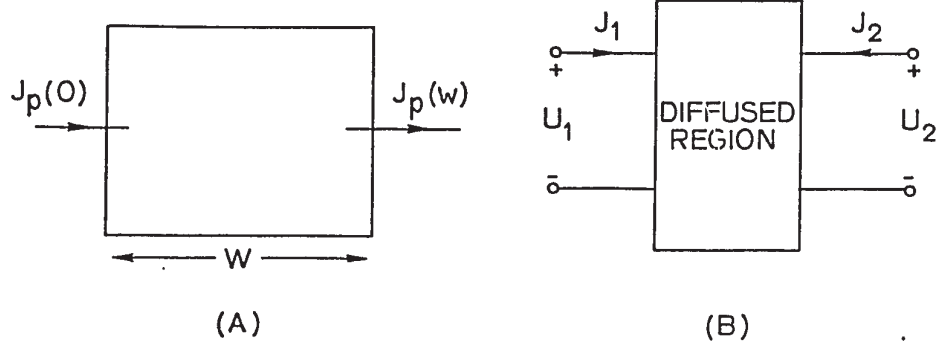


Figure 3.1: General bulk region showing hole currents(A), two port representation(B)

$$Y_{21} = \left. \frac{J_2}{u_1} \right|_{u_2=0} \quad (3.16)$$

Repeating with $u_1 = 0$

$$Y_{12} = \left. \frac{J_1}{u_2} \right|_{u_1=0}$$

$$Y_{22} = \left. \frac{J_2}{u_2} \right|_{u_1=0} \quad (3.17)$$

Using the terminology of del Alamo and Swanson [1], we get

$$Y_{11} = J_{of}$$

$$Y_{22} = J_{or} \quad (3.18)$$

if we define

$$\left. \frac{J_2}{J_1} \right|_{u_2=0} \equiv -\alpha_f \quad ; \quad \left. \frac{J_1}{J_2} \right|_{u_1=0} \equiv -\alpha_r \quad (3.19)$$

then

$$Y_{12} = -\alpha_r J_{or}$$

$$Y_{21} = -\alpha_f J_{of} \quad (3.20)$$

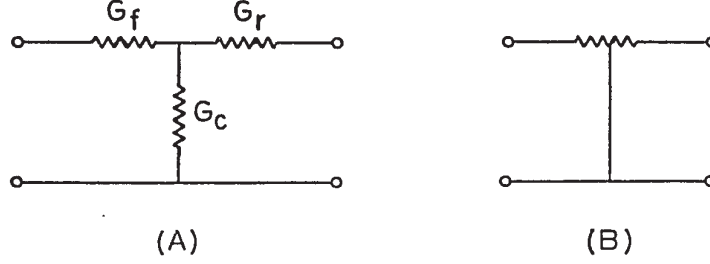


Figure 3.2: T-network equivalent circuit for bulk region(A), shorthand(B)

and we get

$$\begin{bmatrix} J_1 \\ J_2 \end{bmatrix} = \begin{bmatrix} J_{of} & -\alpha_r J_{or} \\ -\alpha_f J_{of} & J_{or} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (3.21)$$

In fact, this sequence of setting u_1 and u_2 successively to zero and measuring the resulting currents forms the basis of the numerical procedure outlined in [1]. This same algorithm was used for the calculations presented in the latter part of this chapter.

In Appendix C, it is demonstrated that reciprocity holds for arbitrary profiles and frequencies. This means that

$$\alpha_r J_{or} = \alpha_f J_{of} \quad (3.22)$$

Therefore three independent parameters are needed to characterize the diffusion. This suggests that a “T” network representation should suffice. For the network shown in Fig. 3.2a , the values are

$$\begin{aligned} G_f &= J_{of} \frac{1 - \alpha_r \alpha_f}{1 - \alpha_r} \\ G_r &= J_{or} \frac{1 - \alpha_r \alpha_f}{1 - \alpha_f} \\ G_c &= J_{of} \frac{1 - \alpha_r \alpha_f}{\alpha_r} \end{aligned} \quad (3.23)$$

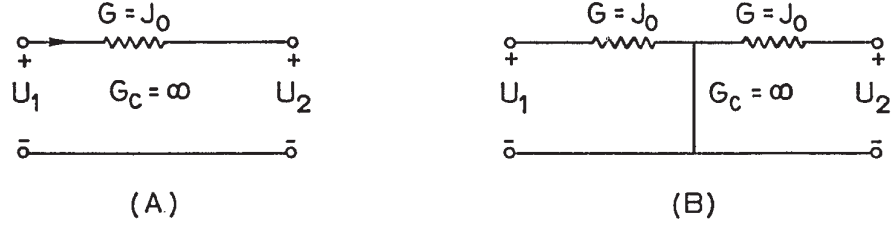


Figure 3.3: equivalent circuit for uniform region; transparent case (A), opaque case (B)

For convenience, a shorthand notation shown in Fig. 3.2b will be utilized to represent the diffused regions. To obtain an intuitive grasp for this representation calls for a more heuristic approach. Consider the special case when the dopant profile is constant. Then $\alpha_r = \alpha_f = \alpha$ and $J_{of} = J_{or} = J_o$ from symmetry, so

$$G_f = G_r = J_o(1 + \alpha)$$

$$G_c = J_o \frac{1 - \alpha^2}{\alpha}$$

Now since α is the ratio of the current that exits from one side of the diffusion to that entering the opposite side, if there is no recombination, $\alpha = 1$ i.e. the region is said to be transparent. In the transparent case,

$$G_f = G_r = 2J_o$$

$$G_c \rightarrow 0$$

so the “T” equivalent circuit reduces simply to a single resistor with the value

$$G = J_o$$

as represented in Fig. 3.3a. The equivalent circuit then says

$$J = J_o(u_1 - u_2)$$

and since for a uniform transparent region

$$J_o = \frac{qn_{i_o}^2}{N_{eff}} \frac{D_p}{W}$$

$$J = \frac{qn_{i_o}^2 D_p}{N_{eff} W} (e^{V_1/V_T} - e^{V_2/V_T})$$

which is the familiar result for a transparent uniformly doped region. It can be seen that the direction of the net current flow is determined by the biases V_1 and V_2 . Now consider the opposite extreme when $\alpha = 0$. Then the diffusion is said to be opaque because no minority carriers survive the transit. Then

$$G_f = G_r = J_o \quad G_c \rightarrow \infty$$

The equivalent circuit is shown in Fig. 3.3b. Now the ports are decoupled by the short that is G_c and the potential u_2 does not affect the value of J (at port 1). This is the result expected from an opaque emitter where the nature of the “surface” boundary condition (i.e. u_2) does not change the J_o simply because there are no holes in appreciable quantities at the surface anyway.

The physical interpretation to G_f , G_r , and G_c are evident. G_c represents a shunting of holes due to recombination mechanisms. G_f and G_r which are proportional to J_{of} and J_{or} respectively represent transport “conductances”. For large J_o , these conductances are large, corresponding to an increased minority carrier current flow for a given bias.

3.2.3. Interfaces

The disruption of crystalline order at the surface and interfaces can lead to the presence of a sheet density of active recombination centers. The exact recombination mechanism is not of concern, so long as the recombination rate is linear in the hole density. Then we can write

$$J_{int} = q \cdot S_{int} \cdot p_{int} \tag{3.24}$$

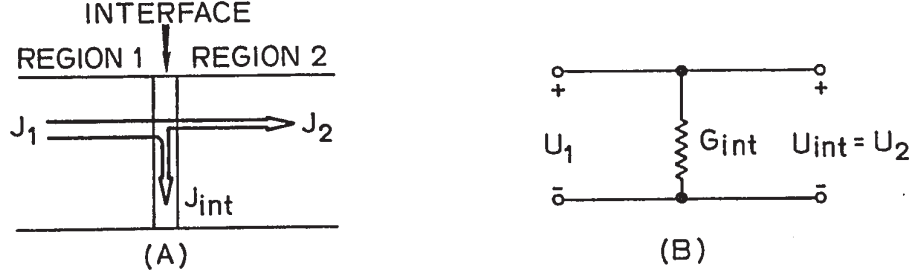


Figure 3.4: Interfacial region: band diagram (A), equivalent circuit (B)

where p_{int} is the local hole density at the interface and S_{int} is a phenomenological interfacial recombination velocity. Referring to Fig. 3.4a,

$$\begin{aligned}
 J_1 &= J_{int} + J_2 \\
 &= q \cdot S_{int} \cdot p_{int} + J_2 \\
 &= S_{int} \frac{qn_{io}^2}{N_{eff}} u_{int} + J_2
 \end{aligned} \tag{3.25}$$

This leads to the equivalent circuit of Fig 3.4b where

$$G_{int} = \frac{qn_{io}^2}{N_{eff}} S_{int} \tag{3.26}$$

If the interfacial recombination velocity is small, G_{int} is small and minority carrier transport is not materially affected. If S_{int} is large as is the case for highly defective or poorly passivated interfaces, then G_{int} is large and most of the carriers are shunted, i.e. lost to interfacial recombination.

3.2.4. Tunneling Barriers

Consider the semiconductor–insulator–semiconductor sandwich shown in Fig. 3.5 where the presence of the thin insulator introduces a barrier in the valence band. If this barrier is sufficiently thin, then holes may tunnel through. A more rigorous treatment of tunneling can be found in [2]; for our purposes a semiclassical

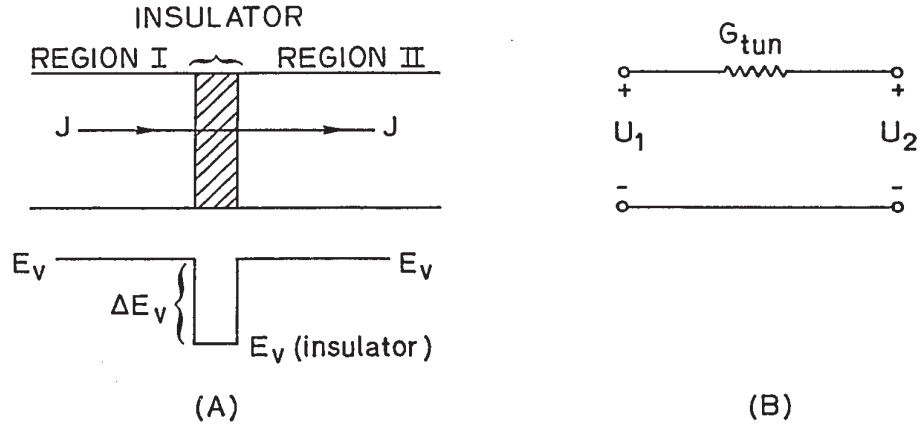


Figure 3.5: Insulating layer and associated hole tunnel barrier (A), equivalent circuit (B)

approach will suffice. Suppose that there is a net hole current flowing in the $+x$ direction through the SIS structure. For holes which follow a Boltzman distribution, the mean velocity per carrier is [3]

$$v_{th} = \sqrt{\frac{8kT}{\pi m^*}} \quad (3.27)$$

Since we have assumed the effective mass m^* to be isotropic, the flux of holes in the $+x$ direction is simply given by

$$\frac{1}{4} v_{th} p_1 \quad (3.28)$$

If there is negligible voltage drop across the oxide (see Chapter 7) then we can assign a constant tunnel probability T_h so that the $+x$ directed current is given by

$$|J_+| = \frac{q}{4} v_{th} T_h p_1 \quad (3.29)$$

Similarly for holes travelling in the $-x$ direction

$$|J_-| = \frac{q}{4} v'_{th} p_2 T_h \quad (3.30)$$

where v'_{th} may differ from v_{th} if the effective mass is different in region II. The net tunnel current must then equal the hole current J flowing through this structure.

$$\begin{aligned} J &= J_+ - J_- = \frac{q}{4} T_h (v_{th} p_1 - v'_{th} p_2) \\ &= \frac{q}{4} T_h \left[v_{th} \frac{n_{io}^2}{N_{eff}} u_1 - v'_{th} \frac{n_{io}^2}{N_{eff,2}} u_2 \right] \end{aligned} \quad (3.31)$$

where $N_{eff,2}$ must take into account the band structure of region II. If regions I and II are both silicon then this reduces to

$$J = \frac{q n_{io}^2}{N_{eff}} \frac{v_{th}}{4} T_h (u_1 - u_2) \quad (3.32)$$

which has the trivial representation of Fig. 3.5b where

$$G_{tun} = \frac{q n_{io}^2}{N_{eff}} T_h \frac{v_{th}}{4} \quad (3.33)$$

If the hole tunnel probability is of order unity then transport is minimally affected. On the other hand if $T_h \ll 1$ then the effective transport impedance becomes quite large resulting in a large drop in the hole imref even for small values of J . If the semiconductors in regions I and II differ, then a more complicated representation is needed; unfortunately, this makes the utility of the equivalent circuit representation somewhat marginal.

3.2.5. Integration into a Network

A general emitter and contact structure is constructed as a stack of several different types of layers. In order to integrate the equivalent circuit elements into an equivalent network, the boundary conditions must be examined.

Consider the infinitesimal element dx represented in Fig. 3.6a. From charge conservation, in steady state

$$J_l = J_r$$

since

$$J_p = \frac{q n_{io}^2}{N_{eff}/D_p} \nabla u \quad (3.34)$$

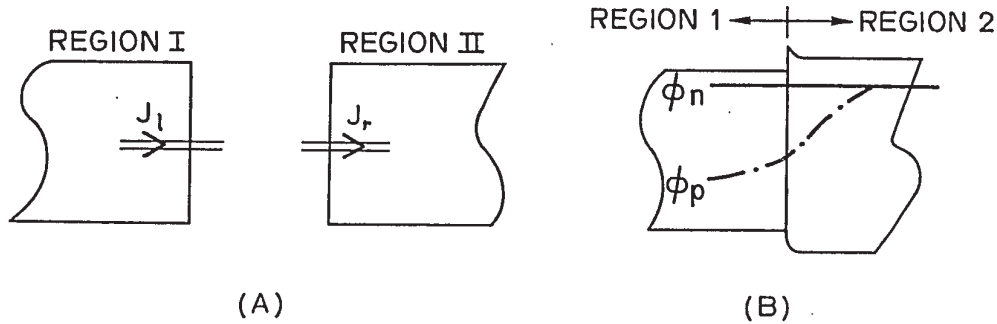


Figure 3.6: boundary conditions for matching adjacent regions

u must be continuous since J_p must remain finite. As for the gradient in u , this is determined by the requirement that J_p be continuous. In general, it will experience a change in slope as a boundary is crossed (see Fig. 3.6b); however, for the purposes of the equivalent circuit representation, only the continuity of u is relevant. Therefore, the equivalent circuit models can be sequentially linked in the same order as the various layers in the emitter and emitter contact appear.

3.3. Practical Emitter Structures

A generalized emitter+contact structure is illustrated in cross section in Fig. 3.7a. An $N+$ diffusion in the monocrystalline region is separated by an interfacial layer from some other semiconductor "X" (which may be polycrystalline silicon) to which the metal makes an ohmic contact. From this diagram, it can be seen that J , the total hole current flowing into the contact is composed of several components reflecting recombination in the various regions. J_1 is due to bulk recombination in the $N+$ diffusion in the monocrystalline silicon. J_2 represents recombination at the interface between the silicon and the semiconductor "X" due to interfacial states. J_3 arises from recombination in the bulk of semiconductor "X" and J_4 represents the current due to carriers that have survived the transit process to recombine at the semiconductor "X"-metal interface.

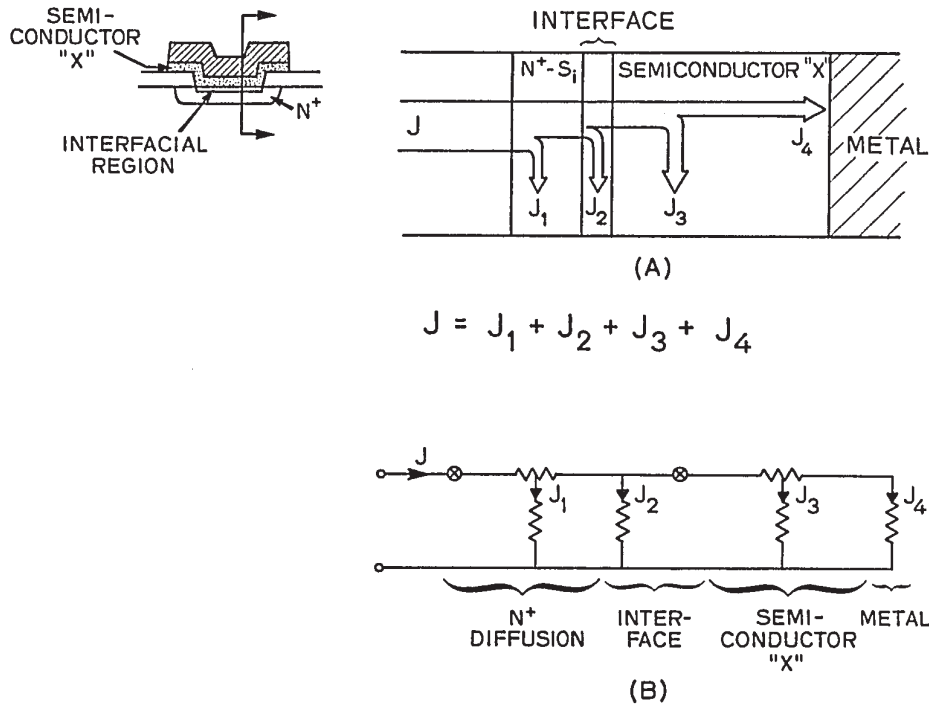


Figure 3.7: generalized contact structure (A), and equiv. circuit (B)

Assembling the various elements yields the equivalent circuit for the generalized HJC structure as shown in Fig. 3.7b. In the following section, simplified subsets of this representation will be used to examine the dominant recombination mechanisms in various emitter structures. For each special case that is considered, the results of numerical simulations obtained by using the algorithm outlined in [1] are presented to corroborate the qualitative trends deduced from the equivalent circuit approach. Details regarding the BGN and lifetime parameters used in the simulations are presented in Appendix D.

3.3.1. Homojunction Emitter

An emitter fabricated entirely in monocrystalline silicon with a metallic contact is referred to as the "conventional" emitter structure as it is the one commonly used

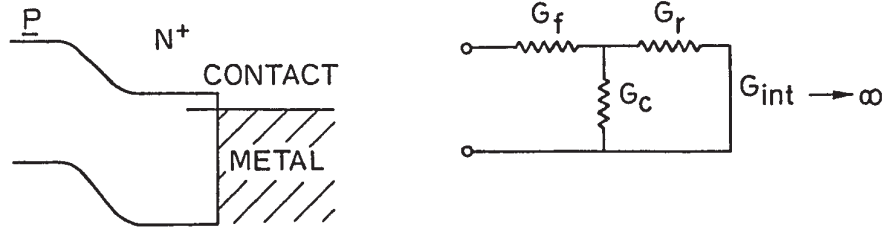


Figure 3.8: homojunction emitter: band structure (A), equivalent circuit (B)

in contemporary devices. A cross-section and equivalent circuit are presented in Fig. 3.8. The impedance corresponding to S_{metal} is taken in the limit as $S_{metal} \rightarrow \infty$. Although finite in reality, its effective impedance is negligible compared to the transport impedance through the $N+$ diffused layer. G_f, G_r, G_c have the values given by eqn. (3.23). Since the large surface recombination velocity at the metal essentially forces $u = 0$ at the metal-silicon interface, the resulting emitter current J_{oe} simply equals J_{of} as can be confirmed by substitution of values for G_f, G_r, G_c . The physical interpretation of the various elements in the equivalent representation has already been discussed; however, it may be profitable at this point to discuss J_{oe} from a slightly different viewpoint. J_{oe} is essentially a flux and as such, can be decomposed into the product of a density term and a transport velocity term, i.e.

$$J_{oe} = \text{charge} \cdot \text{density} \cdot \text{transport velocity}$$

Minimizing J_{oe} requires minimization of this product subject to the constraints of a particular application. The transport velocity in a diffused region is essentially given by the diffusion velocity of carriers. This can be written as

$$v_{diff} = \frac{D}{\eta L} \quad (3.35)$$

where D is the diffusion coefficient and L is the emitter depth for a transparent emitter or some effective diffusion length for an opaque emitter. η is a factor of

order unity which reflects the effect of built-in quasi-fields for graded profiles. It will be assumed to be unity for the sake of simplifying the discussion. It can be seen that for the transparent case, apart from its dependence on D , the diffusion velocity is fixed by the emitter depth. The choice of dopant profile does not strongly influence v_{diff} apart from the dependence of D on the doping level. Therefore, the burden of minimizing J_{oe} for a transparent emitter falls on minimization of the equilibrium hole density if the value of L is constrained by other considerations. For applications where deeper emitters can be tolerated, the use of opaque structures where $W_e > L$, can reduce the diffusion velocity. The maximum reduction that can be achieved is limited by the value of D/L . Since

$$v_{diff} \geq \frac{D}{\eta L} = \frac{1}{\eta} \sqrt{\frac{D}{\tau_p}} \quad (3.36)$$

we see that the relevant parameter of interest is D/τ_p where τ_p is the hole life-time. For the heavier doping levels encountered in the emitter, D displays a weak dependence on the dopant density. However, if τ_p becomes limited by Auger recombination mechanisms, L will fall rapidly with increasing doping level. The diffusion velocity will then increase almost proportionately with dopant density once Auger effects dominate. These trends can be seen in Appendix D.

The density term is given by

$$p_o \approx \frac{n_{io}^2}{N_{eff}}$$

where N_{eff} is the “effective” dopant density. N_{eff} shows a deviation from N for doping densities above $1 \times 10^{17} \text{ cm}^{-3}$. The value of N_{eff} increases with N until an ultimate value of $1 \times 10^{18} \text{ cm}^{-3}$ is reached for $N \approx 3 \times 10^{19} \text{ cm}^{-3}$. Further increases in N will not increase N_{eff} and depending upon the particular BGN model used, N_{eff} may even decrease.

The role of these two terms in determining J_{oe} can be illustrated by examining the dependence of J_{oe} for a constant doping profile emitter with a fixed junction depth of 10μ as a function of dopant concentration. The curves shown in Fig. 3.9,10,11 display $\overline{v_{diff}}$ and $\overline{p_o}$ and J_{oe} where these quantities are defined by

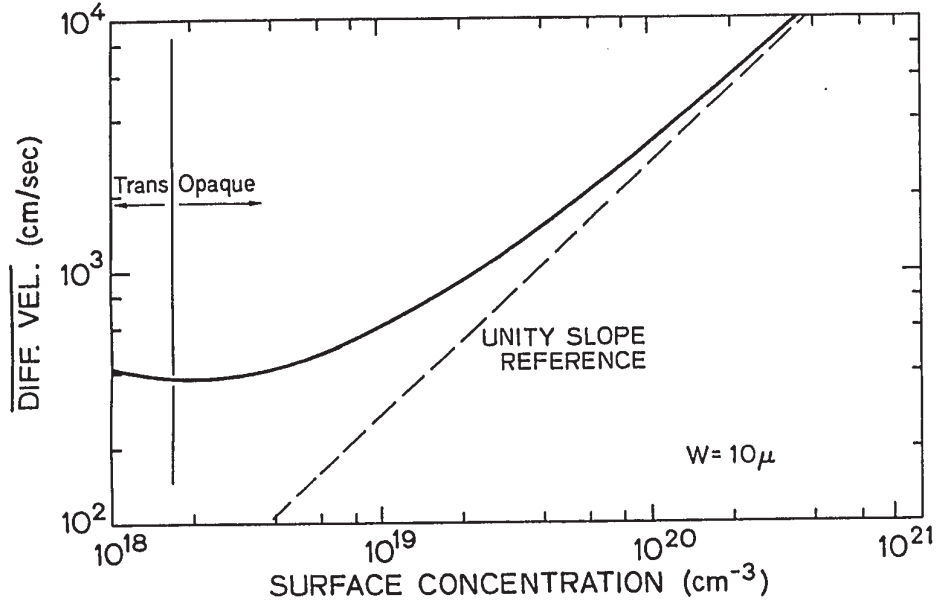


Figure 3.9: diffusion velocity for a uniform homojunction emitter

$$\bar{p}_o = \frac{n_{io}^2 W}{\int_0^W N_{eff}(x) dx} \quad (3.37)$$

$$\bar{v}_{diff} = \frac{J_{oe}}{q \bar{p}_o} \quad (3.38)$$

The emitter is considered transparent if $\alpha > 0.67$. These curves were calculated using the algorithm and BGN models in [1]. In the transparent regime for $N < 2 \times 10^{18} \text{ cm}^{-3}$, $\bar{v}_{diff}$ decreases with doping density. This reflects the assumed form for the dependency of D on N (see Appendix D). Once the emitter becomes opaque, the behavior of L controls $\bar{v}_{diff}$. As can be seen from Fig. 3.9, for $N > 1 \times 10^{19} \text{ cm}^{-3}$, $\bar{v}_{diff}$ increases rapidly with N , approaching a linear dependence for $N > 1 \times 10^{20} \text{ cm}^{-3}$ as expected for the Auger dominated regime. The value of N for which $\bar{v}_{diff}$ begins to increase depends primarily on the choice of the Auger coefficient. To first order, the behavior of $\bar{v}_{diff}$ is insensitive to the BGN model used.

Looking at the thermal equilibrium hole density term in Fig. 3.10, $\bar{p}_o$ decreases

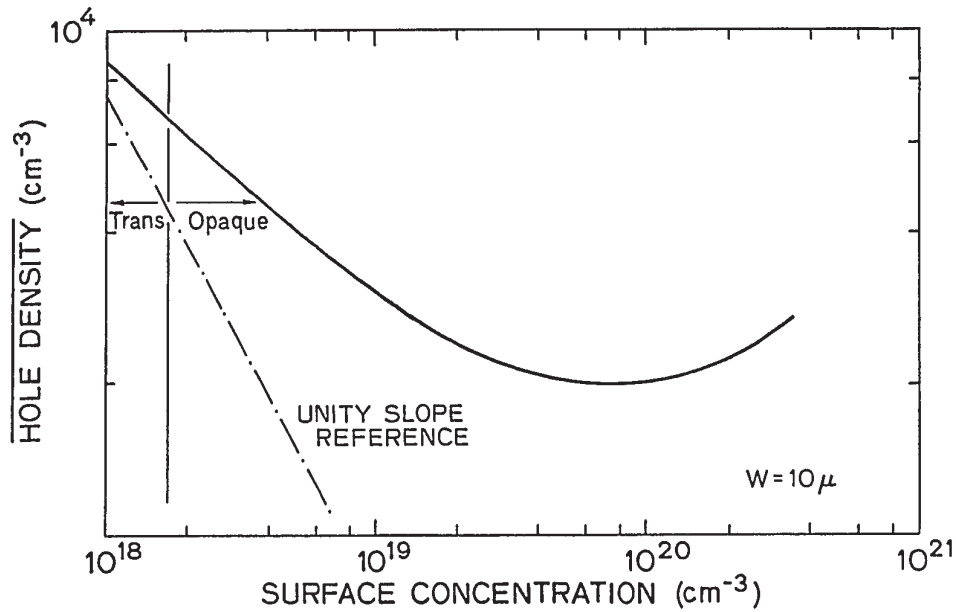


Figure 3.10: effective hole density for uniform homojunction emitter

with N for $N < 4 \times 10^{19} \text{ cm}^{-3}$ although much more slowly than the inverse linear dependence predicted by ignoring BGN effects. It is clear that increasing doping much above $6 \times 10^{19} \text{ cm}^{-3}$ does not reduce $\bar{p}_o$; in fact this BGN model predicts an increase in $\bar{p}_o$ for $N > 1 \times 10^{20} \text{ cm}^{-3}$.

Since J_{oe} is given by the product of the two terms, the minimum J_{oe} in Fig. 3.11 is seen to occur at $N \approx 7 \times 10^{18} \text{ cm}^{-3}$ with a value of approximately $3 \times 10^{-13} \text{ A/cm}^2$. Although the hole density has not yet attained its minimum value, the rapid increase in diffusion velocity produces the minimum in J_{oe} at this value of N . Analogous results for a Gaussian emitter are shown in Fig. 3.12 where the doping concentration at the fixed junction depth of 10μ is kept constrained at $1 \times 10^{16} \text{ cm}^{-3}$. The variable is the surface doping concentration. Comparing Fig. 3.9-11 and 3.12, the corresponding quantities display similar behavior. The high doping effects and J_{oe} saturation for the Gaussian emitter are deferred to higher surface concentrations

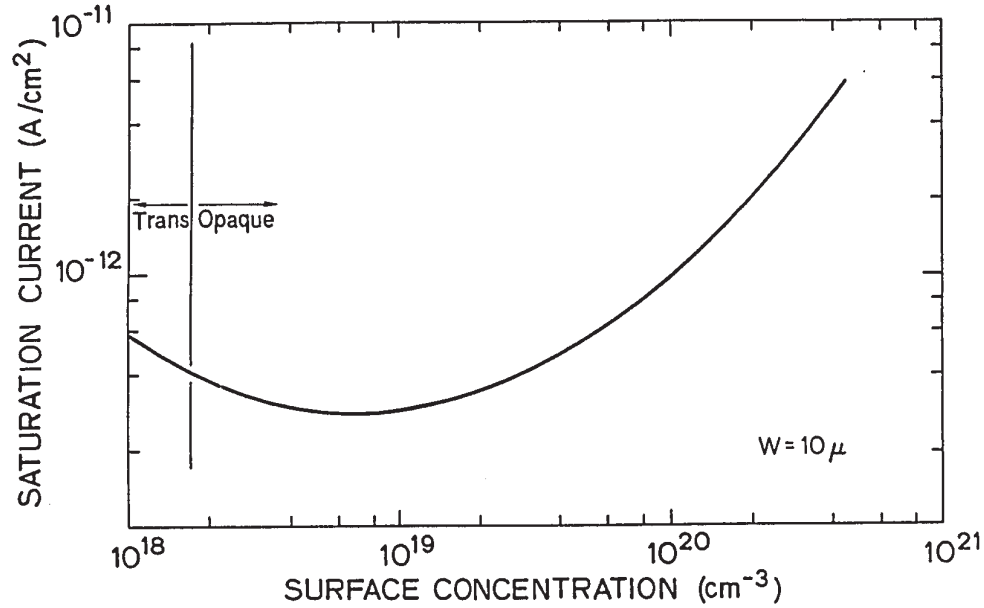


Figure 3.11: J_{oe} for a uniform homojunction emitter

compared to the constant doping profile case. The diffusion velocity also tends to saturate for $N_s > 1 \times 10^{20} \text{ cm}^{-3}$ which contrasts with the steadily increasing trend seen for the constant profile emitter. Both observations can be explained by the presence near the junction of a lightly doped region for the Gaussian (or similar profile) which is not present for the constant profile case. This lightly doped layer acts to “buffer” the junction from the more heavily doped regions near the surface. When $N_s > 1 \times 10^{20} \text{ cm}^{-3}$, the surface region has a very short hole lifetime and the resulting high diffusion velocity for holes makes this region for all practical purposes an extension of the metallic contact. The only influence of N_s on v_{diff} at this point is the slight narrowing of the “transparent” layer effected by the changing profile.

Computations of J_{oe} for a constant, Gaussian, and complementary error function profile emitter as a function of emitter depth with surface concentration as a parameter are shown in Fig. 3.13,14,15 respectively. For practical structures

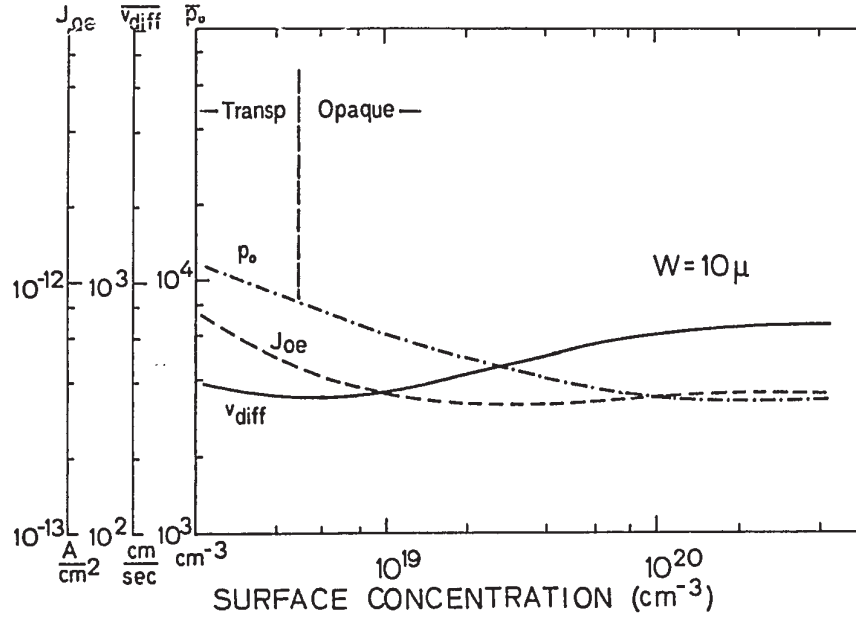


Figure 3.12: diffusion velocity, hole density and J_{oe} for Gaussian emitter

($W_e < 20\mu$) it can be seen that J_{oe} is limited to values above $2 \times 10^{-13} \text{ A/cm}^2$. For applications such as high speed bipolar structures where the emitter depth is constrained to be shallow ($\ll 1\mu$), the rapid increase of J_{oe} with decreasing junction depth is apparent. For this transparent regime, it is seen that surface concentrations above $5 \times 10^{19} \text{ cm}^{-3}$ are needed, but that further increases do not materially affect the value of J_{oe} . The necessity of maintaining surface concentrations high in order to meet the requirements for low majority carrier contact resistance is seen to be consistent with minimizing J_{oe} for all the cases examined with the exception of the deep constant profile emitter where a two step profile is needed to maintain a high surface concentration while keeping the bulk dopant concentration below $1 \times 10^{19} \text{ cm}^{-3}$.

To summarize the results for the homojunction case, we can note that $J_{oe} < 1 \times 10^{-13} \text{ A/cm}^2$ are not possible for practical emitter structures regardless of the

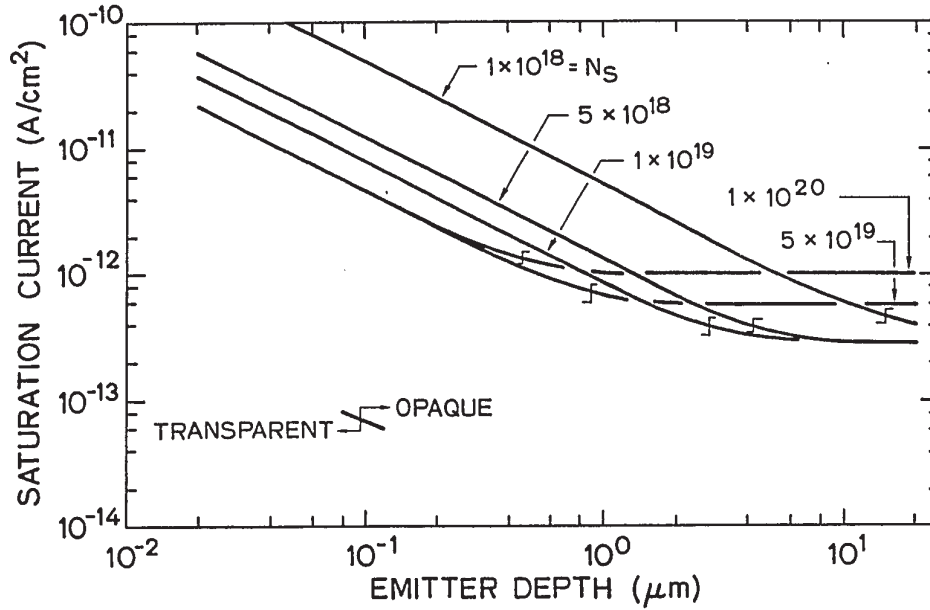


Figure 3.13: J_{oe} vs. emitter depth for uniform emitter

doping profile used. The deeper emitter junctions depths required to achieve the lower values of J_{oe} require high temperature diffusions. In the solar cell, this can lead to degradation of the bulk lifetime properties which could nullify any advantages offered by a lower emitter J_{oe} . For shallow emitter structures, the increase in J_{oe} predicted by using the simplified assumptions in Chapter 2 are supported by the numerical simulations.

3.3.2. Heterojunctions

By interposing a semiconductor other than silicon between the $N+$ diffusion and the metal, the equivalent circuit becomes Fig. 3.16. In addition to the “ T ” network representing semiconductor X (which may be silicon), there is a shunt element G_{int} which models the action of interfacial recombination. Looking at the circuit, it can be seen that J_{oe} can be reduced below the value attainable for the homojunction case if G_{int} is sufficiently small and if the transport impedance of X is greater than

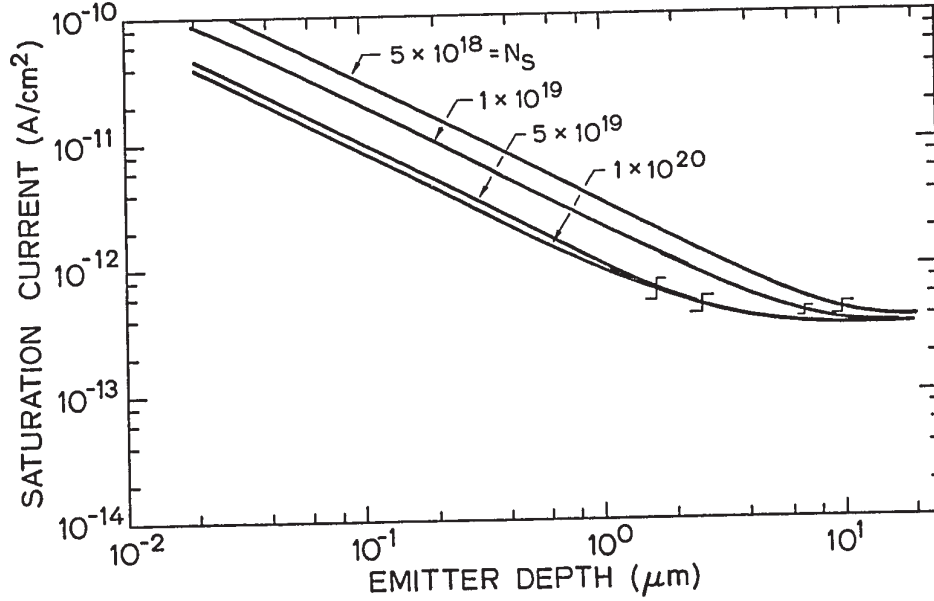


Figure 3.14: J_{oe} vs. emitter depth for Gaussian emitter

that attainable for monocrystalline silicon. More precisely

$$J_{oe} = J_{of} \left[1 - \alpha_f \alpha_r \frac{J_{or}}{J_{or} + J_{o,X}} \right] \quad (3.39)$$

where $\alpha_f, \alpha_r, J_{of}, J_{or}$ refer to the monocrystalline parameters, and $J_{o,X}$ is the saturation current of region X. To minimize the value of $J_{o,X}$, the same considerations apply as for monocrystalline silicon except that now, we are no longer constrained by the materials properties of crystalline silicon. Since

$$J_{o,X} = \text{charge} \cdot \text{density} \cdot \text{transport velocity}$$

and since the transport velocity for a given hole density is minimized by using thicknesses sufficient to ensure opacity, the equation becomes

$$J_{o,X} = \frac{q n_{ie,X}^2}{N_X} \cdot \frac{D_X}{L_X} \quad (3.40)$$

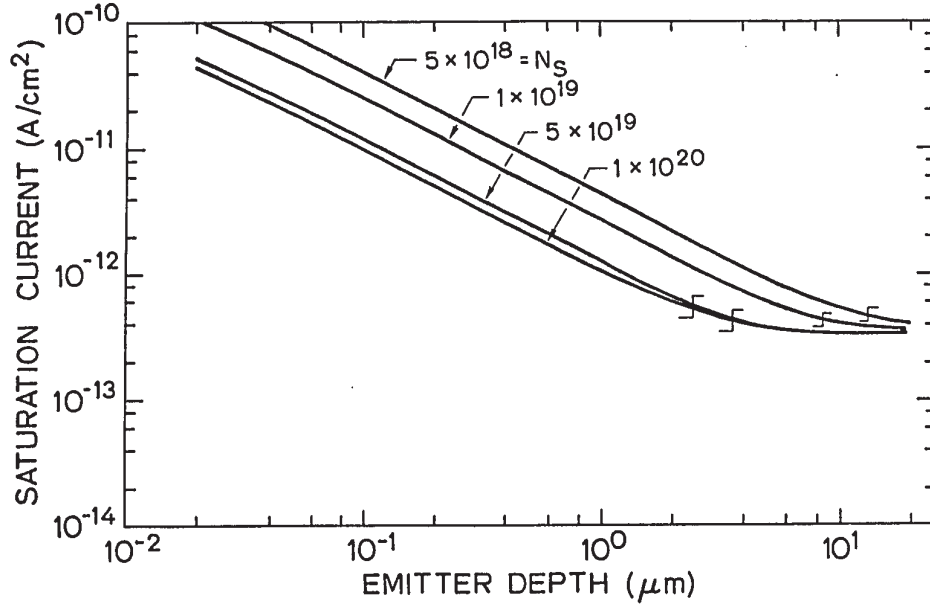


Figure 3.15: J_{oe} vs. emitter depth for ERFC emitter

where the subscript X refers to the parameters that apply to the region “X” and a constant doping profile has been assumed. As in the monocrystalline case, there are two generic approaches to reducing $J_{o,X}$. The “barrier” approach seeks to reduce $n_{ie,X}^2/N_X$ by reducing the value of $n_{ie,X}^2$ by using a wider gap semiconductor. The transport velocity approach seeks the same objective by reducing the value of $(D/L)_X$ or equivalently, $(D/\tau)_X$. For applications where the AC performance of the emitter is of concern, the barrier approach is to be preferred since it minimizes minority carrier storage. It also does not rely on the maintenance of minority carrier lifetime in region “X”, which may prove difficult. In any case, we can write

$$\begin{aligned}
 J_{surf} &= J_{int} + J_X \\
 qS_c p_{surf} &= qS_{int} p_{surf} + q \frac{n_{ie,X}^2}{N_X} \sqrt{\frac{D_X}{\tau_X}} \\
 S_c &= S_{int} + \frac{n_{ie,X}^2}{n_{io}^2} \frac{N_{surf,eff}}{N_X} \sqrt{\frac{D_X}{\tau_X}}
 \end{aligned} \tag{3.41}$$

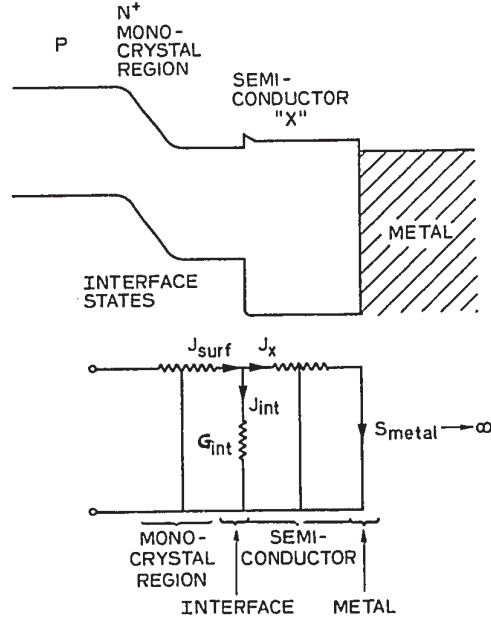


Figure 3.16: heterojunction emitter+ contact structure (A), equivalent circuit (B)

Therefore, we can model J_{oe} of the composite structure by imposing a suitable value for S_c , the effective (contact) recombination velocity at the surface of the monocrystalline region. There may be an additional complication if momentum selection rules are operative at the Si - X interface. There will be a finite value for the quantum mechanical reflection coefficient, but since it relies on the band structure and is independent of the hole density, it can be incorporated into the $n_{ie,X}^2$ term. For low frequencies where S_c is a constant, the transport velocity and barrier approaches are seen to be equivalent. It is not possible from DC measurements of J_{oe} alone to deconvolve the effects of D_X , $n_{ie,X}^2$, and S_{int} unless the physical mechanisms are well characterized and can be identified e.g. by a unique temperature dependence. Since analogous observations can be made for the SIS structure, numerical simulations where S_c is varied as a parameter will be presented in the next section after SIS structures are introduced.

Before proceeding, however, it may be of interest to examine the case of polysilicon contacted emitters. These are of current technological interest because they offer several processing advantages for shallow emitter structures. They are self-aligned if poly is used as the diffusion source, reduce the sidewall contribution to emitter base junction capacitance, and improve process yields by reducing the incidence of base-emitter shorts that are otherwise problematical for very shallow emitters. The measured values for J_{oe} vary widely [4,5,6] and this lack of consistency is most likely due to varying deposition conditions and predeposition cleaning sequences [6]. In this analysis “intimate” poly contacts are assumed where the interfacial oxide, if present, exerts no influence over minority carrier transport.

In order to compute J_{oe} for a poly contacted emitter, $J_{o,x}$, the saturation current for the poly region must be determined. A first principles calculation is not possible with the current state of understanding of minority carrier transport in heavily doped poly. Even the semi-empirical models [7,8] assume an idealized, regular one-dimensional structure for the poly where the lateral grain size is often assumed for the effective grain size in the vertical direction. The microstructure of poly, especially in the crucial near-interface region is not well known; for thicker films, preferential orientation and columnar growth imply that the simple one-dimensional models are too simplistic. For our purposes, it is sufficient that hole recombination mechanisms be linear in the hole density. Then the concept of a surface recombination velocity is useful as a modeling concept even if its value can not be predicted from first principles at this point.

These caveats notwithstanding, a naive calculation of J_{oe} assuming that with the exception of the diffusion coefficient, minority carrier properties in the poly are the same as for crystalline silicon with the same doping concentration. The resulting values of $J_{o,x}$ for a poly region with constant doping of $1 \times 10^{20} \text{ cm}^{-3}$ when D_X is assumed equal to and smaller than D_o , the value assumed for crystalline silicon by factors of 3 and 10, are shown in Fig. 3.17. This reduction in D_X is assumed to parallel the reduction in majority carrier mobility for heavily doped poly [9]. For

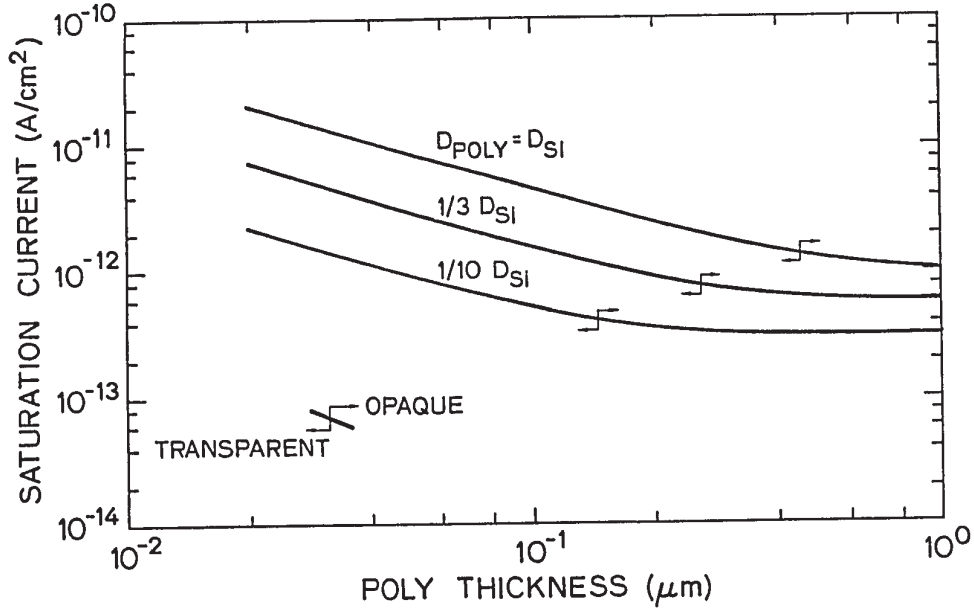


Figure 3.17: J_{oe} vs. POLY thickness for $N+$ POLY emitter

the transparent regime

$$J_{o,X} \approx \frac{qD}{W} \quad (3.42)$$

where W is the poly thickness, $J_{o,X}$ decreases linearly with D_X . For the opaque region where

$$J_{o,X} \approx q \sqrt{\frac{D_X}{\tau_X}} \quad (3.43)$$

$J_{o,X}$ scales with the square root of D_X . These dependencies can be seen in Fig. 3.17. Using this value for $J_{o,X}$, the J_{oe} for a 2000ÅERFC (complementary error function) emitter with concentrations of $1 \times 10^{20} \text{ cm}^{-3}$ and $1 \times 10^{18} \text{ cm}^{-3}$ at the surface and emitter-base junction respectively and an overlying poly layer of varying thickness with $N_X = 1 \times 10^{20} \text{ cm}^{-3}$ is shown in Fig. 3.18. These values were computed using eqn. (3.37). The J_{oe} of the ERFC emitter alone is $3.4 \times 10^{-12} \text{ A/cm}^2$. The maximal reduction in J_{oe} achieved by adding the poly is 6 if $D_X = D_o/3$ where D_o is the hole diffusion coefficient in monocrystalline silicon of the same doping, is assumed.

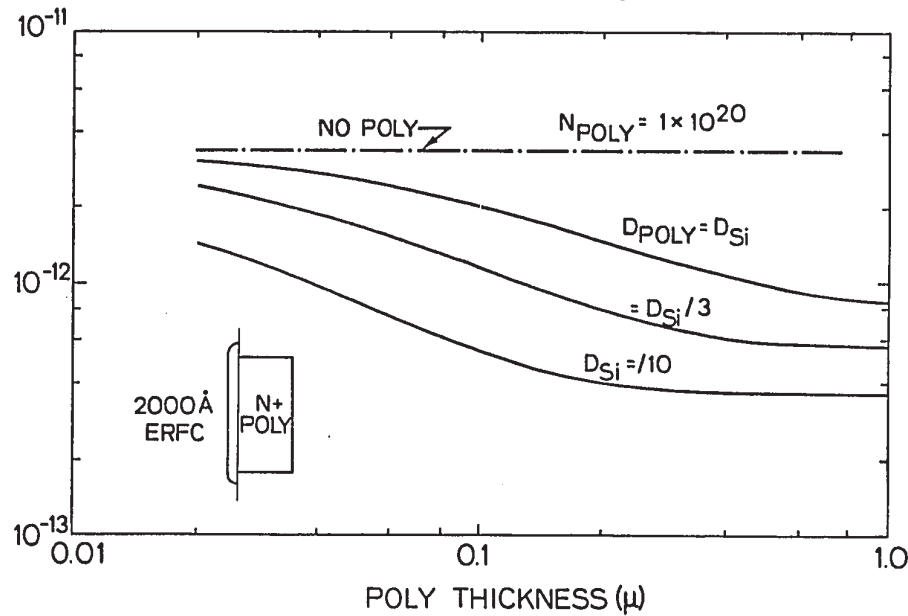


Figure 3.18: J_{oe} for mono+polycrystalline emitter

This seems a remarkable reduction, but is in fact somewhat misleading. The valid comparison would be to a monocrystalline emitter of the same total thickness and doping profile as the composite structure. In Fig. 3.18, this is represented by the $D_X = D_o$ curve; from this it can be seen that the addition of poly only provides a factor of 1.5 relative reduction. In the next chapter, it will be seen that poly emitters are still plagued by minority carrier storage which may limit their frequency performance in some applications.

HJC's fabricated with compound semiconductors have achieved a degree of success as illustrated by devices where the base can have a larger total integrated dose than the emitter; yet because of the difference between the widegap emitter and the narrow gap base, injection efficiencies of near unity are still possible. For silicon, HJC's have not met with success because of the difficulty of obtaining good interfacial characteristics. This results in a high value for S_{int} ; the large inter-

facial recombination currents negate the barrier action of the widegap emitter by providing a parallel shunt path for recombination.

3.3.3. SIS contacts

The $Si-SiO_2$ interface is relatively free of gap states; this self-passivating nature of the $Si-SiO_2$ interface suggests a modification to the previously described structure in which a thin interfacial oxide is grown between the underlying monocrystalline silicon and the overlying layer (which could be poly). If the insulator is made sufficiently thin, majority carriers will be able to tunnel through without a significant voltage drop. Minority carriers would also see a barrier, although in general a different tunneling probability would apply.

The equivalent circuit for this structure adds the transport impedance due to the tunneling barrier as shown in Fig. 3.19. The other change upon going from the HJ to the SIS structure is the replacement of G_{int} by G_{ox} , where G_{ox} is proportional to the effective interface recombination velocity at the Si -oxide and oxide-X interfaces.

There are two particular cases of interest for the SIS system. In the first, the transport impedance of the insulating layer is too small to materially reduce hole transport, but the reduction of G_{int} to G_{ox} which occurs if $S_{ox} \ll S_{int}$ may allow J_{oe} reduction to take place by use of a wider gap material for X. In this case

$$S_c = S_{ox} + \frac{qn_{ie,X}^2}{N_X} \sqrt{\frac{D_X}{\tau_X}} \quad (3.44)$$

where again, S_c is the surface recombination velocity seen by holes approaching the silicon-oxide interface from the monocrystalline emitter. In the second case, the tunneling impedance is sufficiently large that it dominates the transport impedance. In this case

$$S_c = S'_{ox} + \frac{qn_{io}^2}{N_{eff}(0)} \frac{v_{th}}{4} T_h \quad (3.45)$$

where T_h , the hole tunnel probability satisfies $T_h \ll 1$, and S'_{ox} reflects the fact that S_{ox} may be changed by the fact that there is now a substantial drop in the hole imref across the insulator.

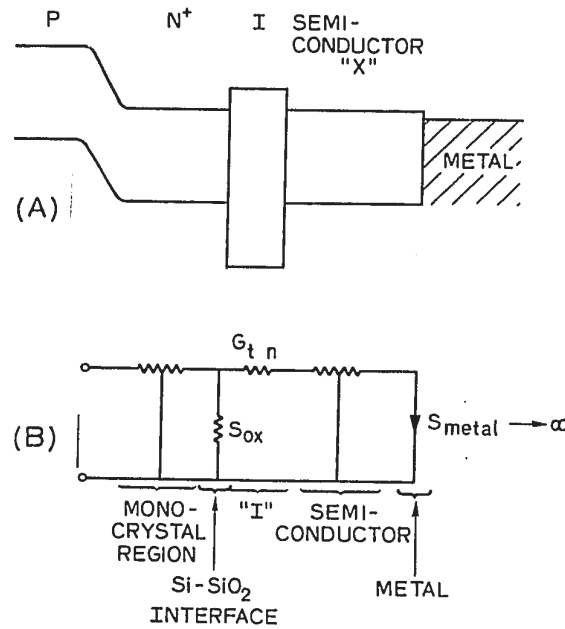


Figure 3.19: SIS contact structure (A) equivalent circuit (B)

For both cases, it can be seen that modeling the "I" and "X" regions by a constant effective surface recombination velocity is appropriate as was the case for the HJ structure. Again, the current level of understanding does not permit a derivation of these quantities from first principles. For example, the value for T_h can be calculated by assuming a barrier height that applies for bulk silicon and silicon dioxide and using thicknesses measured by ellipsometry or a more direct method such as HRTEM. However, the electronic structure of these thin ($< 20 \text{ \AA}$) oxides is not known, and is very likely strongly dependent upon processing history, method used to grow the oxide, etc. These uncertainties make a more detailed modeling at this point of limited practical value.

Tailoring the profile in the monocrystalline region to minimize J_{oe} for the HJ and SIS structures depends upon the value of S_c . For the homojunction case, it was shown that BGN restricts $N_{eff} < 1 \times 10^{18} \text{ cm}^{-3}$. Therefore the minimum values for

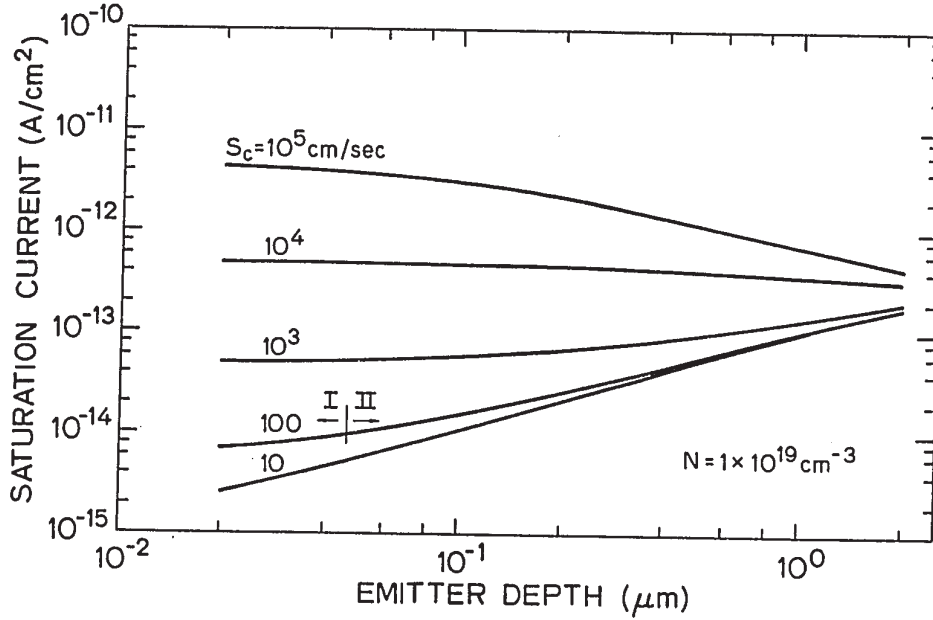


Figure 3.20: J_{oe} vs. emitter depth, S_c for uniform emitter

J_{oe} of $2 \times 10^{-13} \text{ A/cm}^2$ imply that a lower limit on the diffusion velocity corresponds to

$$J_{oe} = \frac{qn_{i0}^2}{N_{eff}} \cdot v_{diff}$$

$$v_{diff} \geq 6 \times 10^3 \text{ cm/sec}$$

If S_c is much greater than this value, the HJ and SIS structures degenerate into the homojunction case with a metallic contact. If S_c can be reduced significantly below this value, the possibility arises of reducing J_{oe} significantly below that which is attainable using the homojunction. Simulations using the same algorithm used for the homojunction case of a constant doping profile emitter as a function of emitter depth for various values of S_c , are shown in Fig. 3.20. The choice of this doping level follows from considerations mentioned in the homojunction case and from optimization considerations to be presented. It can be seen that J_{oe} asymptotically approaches a value of $3 \times 10^{-13} \text{ A/cm}^2$ for emitter depths in excess of a micron inde-

pendent of the value of S_c . This is to be expected since the J_{oe} of an opaque emitter is independent of the surface boundary conditions. For $S_c > 1 \times 10^4 \text{ cm/sec}$, J_{oe} increases with decreasing emitter depth, W_c , due to the increase in diffusion velocity in the transparent regime as the emitter depth is reduced. The opposite trend is noted for $S_c < 1 \times 10^4 \text{ cm/sec}$ where J_{oe} decreases for smaller values of W_c . This behavior can be explained by examining the equivalent circuit in Fig. 3.19. For small values of S_c , the equivalent impedance of the "I" and "X" regions becomes large enough that the equivalent circuit is reduced to that of Fig. 3.19c. The only remaining elements are those representing the monocrystalline region, and recombination at the interface. The recombination current in the bulk of the emitter can be written as

$$J_{bulk} = \frac{qn_{io}^2 W}{N_{eff} \tau_p} \quad (3.46)$$

where W is the emitter depth and τ_p is the hole lifetime. This assumes that the hole imref is flat across the emitter which is valid for S_c sufficiently small. The interfacial recombination is given by

$$J_{surf} = \frac{qn_{io}^2}{N_{eff}} S_c \quad (3.47)$$

When $W/\tau_p < S_c$, then surface recombination dominates and

$$J_{oe} = \frac{qn_{io}^2 S_c}{N_{eff}} = \text{constant} \quad (3.48)$$

For the $S_c = 100 \text{ cm/sec}$ curve, this corresponds to region I in Fig. 3.20. As $W/\tau_p > S_c$,

$$J_{oe} = \frac{qn_{io}^2}{N_{eff} \tau_p} W \quad (3.49)$$

so J_{oe} increases linearly with W . This corresponds to region II. For $W > L$, where L is the hole diffusion length, J_{oe} saturates at the value

$$J_{oe} = \frac{qn_{io}^2}{N_{eff}} \sqrt{\frac{D}{\tau_p}}$$

this corresponds to an emitter thicker than 2.8μ . Minimizing J_{oe} for low S_c structures evidently relies on a different set of guidelines. Since we can write

$$J_{oe} = qn_{io}^2 \left(\frac{S_c}{N_{eff}} + \frac{W}{\tau_p(N)N_{eff}} \right) \quad (3.50)$$

for this regime, we can see that there is a conflicting requirement to use low doping densities to keep the Auger lifetime high, and the need to keep the surface concentration high to reduce surface recombination by minimizing the hole density at the surface. In reality, the interfacial recombination velocity may depend upon the surface doping density, as is apparently the case for the silicon-silicon dioxide interface [10], but as no functional dependence is known, especially for these thin oxides, it will be assumed that S_c has no such dependence.

The emitter thickness must exceed the mean free path for electron-phonon interactions. This will ensure that the minority carrier density at the surface is given by the quasi-equilibrium value dictated by the separation between the band edge and the hole imref. If the emitter is made too thin, one must take into account the possibility that a one dimensional treatment may not suffice if statistical fluctuations in dopant atom distribution become significant. Using the BGN and Auger coefficients from [1], J_{oe} for the constant profile emitter as a function of concentration is plotted in Fig. 3.21. The emitter depths were fixed at 500 Å and 2000 Å. The curves display a broad minimum. For the lower doping densities where the surface recombination component dominates, J_{oe} decreases slightly with increasing N reflecting the small decrease in the surface hole density. When Auger recombination becomes dominant, the bulk component increases rapidly. For doping densities approaching $1 \times 10^{20} \text{ cm}^{-3}$, the J_{oe} is insensitive to the value of S_c . For the heavy doping parameters used in this simulation, J_{oe} values of $1 \times 10^{-14} \text{ A/cm}^2$ can only be attained with very shallow emitters and well passivated (i.e. low S_c) interfaces. It can be seen that outdiffusion of even 500 Å of dopant at $1 \times 10^{20} \text{ cm}^{-3}$ density can result in J_{oe} of $1 \times 10^{-13} \text{ A/cm}^2$. A similar analysis for a 500 Å ERFC emitter profile is shown in Fig. 3.22. Except for a shift of a factor of two to higher surface concentrations, the characteristics are quite similar. This shift reflects the fact that the constant profile emitter has a much greater dose than an ERFC profile with the same surface concentration and junction depth; it would not be present if the profiles had been chosen instead to have the same total dopant dose. For the ERFC

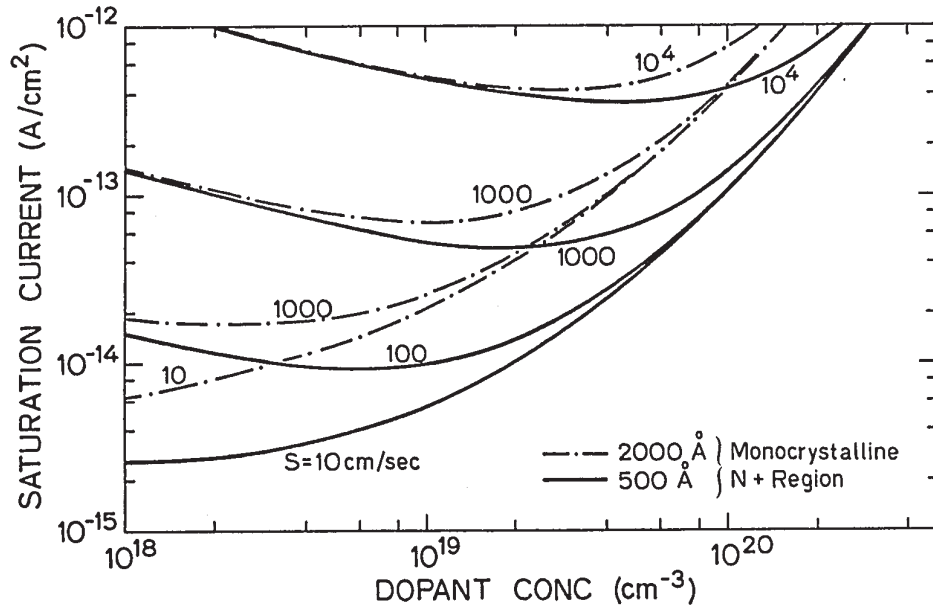


Figure 3.21: J_{oe} vs. N , S_c for uniform emitter

profile, the surface concentration should be kept below $1 \times 10^{20} \text{ cm}^{-3}$. Such shallow profiles are most conveniently produced by outdiffusion from the overlying X layer; this is a loose justification for use of the ERFC profile.

3.3.4. Majority Carrier Series Resistance

Bipolar devices operating at high current densities such as the point contact solar cell at its design point of 500 suns, or scaled bipolar VLSI, are very sensitive to emitter series resistance. For homojunctions, the series resistance is dominated by the contact resistance. For HJ structures, there is the possibility that interfacial layers can induce potential barriers to majority carrier flow; a discussion can be found in Milnes & Feucht [11]. In this section, only the contribution of the insulating layer to the majority carrier series resistance of the SIS structure will be discussed.

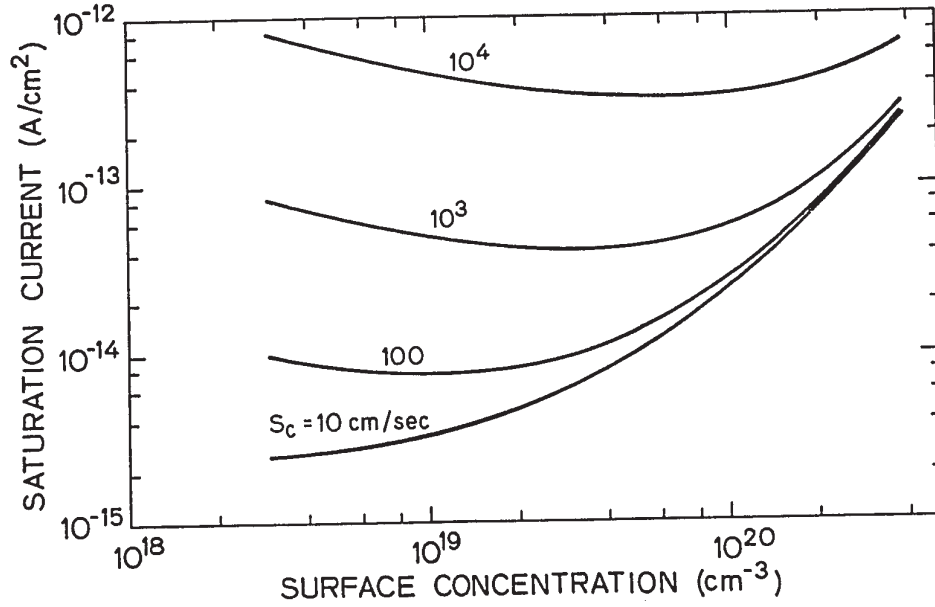


Figure 3.22: J_{oe} vs. N_{surf} , S_c for ERFC emitter

Looking at the equivalent circuit for the SIS structure

$$G_{tunnel} = \frac{qn_{io}}{N_{eff}} \frac{v_{th}}{4} T_h$$

Clearly, the smaller T_h is made, the lower the S_c that can be achieved until S_{ox} becomes dominant. T_h can be made negligibly small by simply increasing the insulator thickness. Unfortunately, this decreases the tunnel probability for majority carrier electrons as well. Using the same semiclassical treatment as for the minority carriers, and assuming for simplicity that the majority carriers also possess a Maxwell-Boltzmann distribution of velocities, we can write the net tunnel current as the difference between current in the $+x$ and $-x$ directions,

$$\begin{aligned} J_n &= J_+ - J_- \\ &= qN \frac{v_{th}}{4} T_n \left(\frac{q\Delta\phi_n}{kT} \right), \quad \text{for } \Delta\phi_n < \frac{kT}{q} \end{aligned}$$

where N is the smaller of the majority carrier densities on either side of the insulating layer, T_n is the majority carrier tunneling probability, and $\Delta\phi_n$ is the drop in majority carrier imref across the insulating layer. The specific contact resistance is then given by

$$R_c = \frac{\Delta\phi_n}{J_n} = \frac{kT}{q^2 N} \frac{4}{v_{th}} \frac{1}{T_n} \quad (3.51)$$

To keep $R_c < 1 \times 10^{-6} \text{ ohm} - \text{cm}^2$,

$$\frac{4kT}{q^2 v_{th}} \frac{1}{NT_n} < 10^{-6}$$

$$T_n > \frac{6 \times 10^{16}}{N}$$

For $N = 1 \times 10^{19} \text{ cm}^{-3}$, $T_n > 6 \times 10^{-3}$ and for $N = 1 \times 10^{20} \text{ cm}^{-3}$, $T_n > 6 \times 10^{-4}$ is required. For the minority carriers, to attain a value of 10^3 cm/sec

$$S_c = \frac{v_{th}}{4} T_h < 10^3$$

we see that if N_{eff} attains its maximum value of $1 \times 10^{18} \text{ cm}^{-3}$, then T_h must be less than 4×10^{-4} to meet this goal. A value for the thermal velocity for both holes and electrons of $1 \times 10^7 \text{ cm/sec}$ was used in these calculations. Whether these tunneling probabilities can be simultaneously attained in practice depends upon the insulator used, its thickness, its preparation conditions and processing history. Since the tunneling barrier heights are not well known, the only recourse at this point is to rely on experimental measurements of J_{oe} and R_c .

3.4. Summary

In this chapter the low frequency performance of several types of emitter structures has been evaluated. Using the BGN and Auger coefficients from [1], it was determined that homojunction emitters with metallic contacts could not achieve J_{oe} values less than about $2 \times 10^{-13} \text{ A/cm}^2$ regardless of the profiles used. The qualitative predictions stated in Chapter 2 regarding the increase of J_{oe} for transparent

emitters as the junction depth was reduced was confirmed by numerical simulations. Analysis of HJ and SIS structures indicate that J_{oe} values of $1 \times 10^{-14} A/cm^2$ and lower are attainable, but that for such low saturation currents, care must be taken to avoid excessive outdiffusion of impurities from the overlying films. Dynamic effects which were neglected in this analysis will be the subject of the next chapter and experimental results from SIPOS and POLY SIS structures that exhibit J_{oe} of $1 \times 10^{-14} A/cm^2$ are presented in Chapter 7.

Chapter 4

Dynamic Behavior of Emitter Structures

4.1, Introduction

For finite frequencies, the minority carrier distribution is nonstationary and time varying minority carrier currents must flow as a result of charge conservation. Quasi-neutrality then requires an equal majority carrier current to flow to maintain charge neutrality. Unlike their static counterparts, these dynamic components of junction current are in general not in phase with the junction voltage. Junction admittances then become complex functions of frequency and operating point.

The charge control approach affords a simplified description of dynamic behavior. Since it is predicated on the assumption of a quasi-static carrier distribution, it is limited to time scales that exceed the time required to establish internal equilibrium within the junction. Fortunately a large fraction of device applications fall within this domain.

In this chapter, the frequency dependence of the saturation current is examined. It will be shown that the static J_{oe} has little bearing on the high frequency behavior. The implications for BJT performance, especially for those utilizing low J_{oe} emitter structures will be investigated using the algorithm of del Alamo and Swanson [1] extended to handle finite frequencies.

4.2. Time Dependent Transport Equations

For finite frequencies, the hole transport equations (3.13,14) must be generalized to include explicit time dependence in the current continuity equation. For the frequencies that will be considered, quasi-equilibrium in the sense that a local imref can be assigned and that carrier recombination is adequately described by a local lifetime will be assumed to hold. Then we have

$$J_p(\omega) = \frac{qn_{i0}^2}{N_{eff}D_p} \nabla u \quad (4.1)$$

$$\nabla \cdot J_p(\omega) = - \left(\frac{qn_{i0}^2}{N_{eff}\tau_p} + j\omega \frac{qn_{i0}^2}{N_{eff}} \right) u \quad (4.2)$$

where we have taken the Fourier transform. These will form the basis for the numerical simulations of AC junction characteristics.

4.3. Charge Control

Charge control offers a simplified alternative to solution of the transport equations. Using an integral approach and the assumption of quasi-statically varying charge distributions provides a physically intuitive model for dynamic behavior which is adequate for many purposes.

The basic concept of charge control relies on charge conservation as expressed through the current continuity equation. Since

$$\frac{1}{q} \nabla \cdot \vec{J}_p = -\frac{\partial p}{\partial t} + G - R = -\frac{\partial p'}{\partial t} - \frac{p'}{\tau_p} \quad (4.3)$$

where p' is the excess hole density. Integrating over the relevant volume of interest yields,

$$\frac{1}{q} \int_V \nabla \cdot \vec{J}_p dv = - \int_V \frac{\partial p'}{\partial t} dv - \int_V \frac{p'}{\tau_p} dv \quad (4.4)$$

Making a key approximation by defining $\tau_{p,eff}$ as an effective lifetime for this region gives,

$$\frac{1}{q} \int_S \hat{n} dS \cdot \vec{J}_p = -\frac{d}{dt} \int_V p' dv - \frac{1}{\tau_{p,eff}} \int_V p' dv \quad (4.5)$$

and since the surface integral is simply the net current flowing out of the region,

$$-I_p^{net} = \frac{-d\overline{Q_p'}}{dt} - \frac{\overline{Q_p'}}{\tau_{p,eff}} \quad (4.6)$$

where $I_{p,net}$ represents the net terminal current **into** the region, and $\overline{Q_p'}$ is the total integrated excess hole charge in the region. Charge control further assumes that using a quasi-static value for the integrated hole charge, i.e. one that depends only upon the instantaneous values of the terminal potentials, is a good approximation. Since we will be dealing with one-dimensional situations, it is convenient to deal with these quantities on a per unit area basis. Equation (4.6) becomes

$$J_p^{net} = \frac{dQ_p'}{dt} + \frac{Q_p'}{\tau_{p,eff}} \quad (4.7)$$

where Q_p' now refers to the quasi-static excess hole charge per unit area. The quantities of interest, Q_p' and $\tau_{p,eff}$ can be computed using the static hole profiles. This is one of the key advantages to using charge control, namely that dynamic performance can be estimated from parameters deduced from purely static measurements.

4.3.1. Constant Profile Emitter

Use of an emitter region with a constant dopant profile allows illustration of high frequency behavior while still yielding tractable analytical solutions. Derivation of charge control results will be followed by a comparison to the exact analytic solutions to evaluate the limitations imposed by the approximations made in charge control.

In the following analysis, three types of emitter structures will be considered — transparent emitters with $S_c = 0$ and $S_c = \infty$, and an opaque emitter. Since charge control requires knowledge of the static solutions, these will be derived before incorporation into the charge control formalism.

Case 1: Transparent Emitter, $S_c = \infty$: Referring to Fig. 4.1, the thin transparent emitter case assumes $W \ll L_p$ the hole diffusion length and that $S_c \rightarrow \infty$. Then the hole current density is given by

$$J_p^{static} = \frac{qn_{i0}^2}{N_{eff}} \frac{D_p}{W} (e^{V/V_T} - 1) \quad (4.8)$$

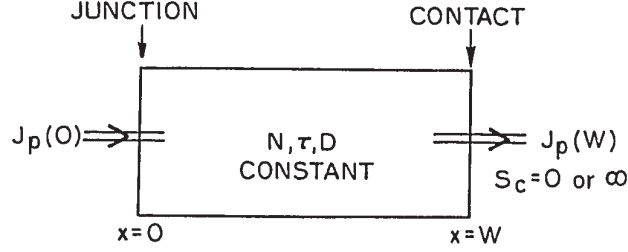


Figure 4.1: symbol definition for uniform emitter

where V is the junction bias, and V_T is the thermal voltage. We can rewrite this as

$$J_p^{\text{static}} = \frac{qn_{i0}^2 W}{2N_{eff}} \frac{1}{W^2/2D_p} (e^{V/V_T} - 1) \quad (4.9)$$

and since

$$\frac{qn_{i0}^2 W}{2N_{eff}} (e^{V/V_T} - 1) = qp_o W/2 (e^{V/V_T} - 1) = Q'_{p1} \quad (4.10)$$

where p_o is the thermal equilibrium hole density, we can write

$$J_p^{\text{static}} = \frac{Q'_{p1}}{\tau_f}, \quad \text{where } \tau_f = \frac{W^2}{2D_p} \quad (4.11)$$

Case 2: Transparent Emitter, $S_c = 0$: The emitter depth W still satisfies $W \ll L_p$ but S_c , the surface recombination velocity, is sufficiently small that bulk recombination dominates. Under these conditions, the hole Fermi level is essentially flat across the region, and we can write for this emitter

$$\begin{aligned} J_p^{\text{static}} &= \frac{qn_{i0}^2 W}{N_{eff} \tau_p} \\ &= \frac{Q'_{p2}}{\tau_p} \end{aligned} \quad (4.12)$$

where τ_p is the effective hole lifetime.

Case 3: Thick Opaque Emitter: W satisfies the relation $W > L_p$. Then the static hole current is given by

$$J_p^{\text{static}} = \frac{qn_{i0}^2 D_p}{N_{eff} L_p} \quad (4.13)$$

Now since

$$p' = p'_{(x=0)} e^{-x/L_p} \quad (4.14)$$

$$\begin{aligned} Q'_{p3} &= q \int_0^\infty p'_{(x=0)} e^{-x/L_p} dx \\ &= \left(\frac{qn_{i0}^2}{N_{eff}} \right) L_p (e^{V/V_T} - 1) \end{aligned} \quad (4.15)$$

Then

$$J_p^{static} = \frac{Q'_{p3}}{L_p^2/D_p} \quad (4.16)$$

since $L_p^2 = D_p \tau_p$

$$J_p^{static} = \frac{Q'_{p3}}{\tau_p} \quad (4.17)$$

4.3.2. Application of Charge Control

From eqn. (4.7)

$$J_p^{net} = \frac{dQ'_p}{dt} + \frac{Q'_p}{\tau_{p,eff}}$$

taking the Fourier transform yields

$$J_p^{net}(\omega) = j\omega Q'_p + \frac{Q'_p}{\tau_{p,eff}} \quad (4.18)$$

Referring to Fig. 4.1, this becomes

$$J_p(0) - J_p(W) = j\omega Q'_p + \frac{Q'_p}{\tau_{p,eff}} \quad (4.19)$$

For case 1, we have that

$$J_p(W) \approx J_p(0) = \frac{Q'_{p1}}{\tau_f} \quad (4.20)$$

since the transport factor for a transparent emitter is unity. Therefore, eqn. (4.18) can be written as

$$J_{p1}(\omega)_{x=0} = Q'_{p1} \left(\frac{1}{\tau_f} + \frac{1}{\tau_{p,eff}} + j\omega \right). \quad (4.21)$$

Since we are considering thin transparent emitters for which the relation $\tau_{p,eff} \gg \tau_f$ holds, this simplifies to

$$J_{p1}(\omega)_{x=0} = \frac{Q'_{p1}}{\tau_f} (1 + j\omega\tau_f) \quad (4.22)$$

For case 2, $S_c = 0$ implies that $J_p(W) = 0$; the same holds for case 3, because the hole density is negligible at the “surface” (which is at ∞) of an opaque emitter. Therefore $J_p(W) = 0$ which allows eqn. (4.7) to be written as

$$\begin{aligned} J_{p2,3}(\omega)_{x=0} &= j\omega Q'_{p2,3} + \frac{Q'_{p2,3}}{\tau_{p,eff}} \\ &= \frac{Q'_{p2,3}}{\tau_{p,eff}} (1 + j\omega\tau_{p,eff}) \end{aligned} \quad (4.23)$$

From this point on, when charge control is invoked, the subscript $x = 0$ will be dropped and τ_p will be used in place of the more correct expression $\tau_{p,eff}$ to simplify notation; that the effective hole lifetime is being used (as opposed to a local lifetime) will be implicit.

4.3.3. Validity of Charge Control

The time required for a disturbance in the hole density to propagate from one side of the junction to the other is given by

$$t_{equil} = W/v_{diff} \quad (4.24)$$

where v_{diff} is the diffusion velocity. As a first approximation to the equilibration time, it sets a limit to the frequencies for which charge control can be applied. For the thick opaque emitter, which can be regarded as a semi-infinite region for hole transport, there is no equivalent quantity. Disturbances introduced at the junction continue to propagate into the bulk; however, since the bulk of the recombination comes from within a diffusion length of the junction, this can be used as an arbitrary definition of an effective junction depth. The equilibration time is then

$$t_{equil} = \frac{W^2}{2D_p} \quad (\text{case 1}) \quad (4.25)$$

$$= \frac{W^2}{2D_p} \quad (\text{case 2}) \quad (4.26)$$

$$= \tau_p \quad (\text{case 3}) \quad (4.27)$$

In summary then, charge control predicts

$$J_{p1}(\omega) = \frac{qn_{io}^2 D_p}{N_{eff} W} (1 + j\omega\tau_f) (e^{V/V_T} - 1); \quad \omega < \frac{1}{\tau_f} \quad (\text{case1}) \quad (4.28)$$

$$J_{p2}(\omega) = \frac{qn_{io}^2 W}{N_{eff} \tau_p} (1 + j\omega\tau_p) (e^{V/V_T} - 1); \quad \omega < \frac{1}{\tau_f} \quad (\text{case2}) \quad (4.29)$$

$$J_{p3}(\omega) = \frac{qn_{io}^2 D_p}{N_{eff} L_p} (1 + j\omega\tau_p) (e^{V/V_T} - 1); \quad \omega < \frac{1}{\tau_p} \quad (\text{case3}) \quad (4.30)$$

It can be seen that for cases 1 and 2, charge control is restricted to frequencies less than the inverse of the transit time. Since in general, $\tau_p > \tau_f$, for case 2, there may be a considerable frequency range where J_p increases linearly with ω (a simple zero) until $\omega = 1/\tau_f$, at which point charge control is no longer applicable. For the thick opaque case, charge control is limited to frequencies below the reciprocal of the minority carrier lifetime, a frequency which is usually much smaller than for the first two cases.

So far, only a plausibility argument has been presented for charge control. There is no apriori reason to expect that use of the static distributions should yield the correct dynamic currents. In fact, after the exact solutions are presented in the next section, it will be shown that charge control does not always yield the correct answer; however, the qualitative behavior is modeled correctly within the frequency limits mentioned above, and the corner frequencies are in error only by a factor of order unity. Because of its physically intuitive nature, charge control is still valuable when seeking to evaluate approximate limits to device behavior. To examine behavior beyond the approximations inherent in charge control, the exact solutions to the transport equations are needed.

4.4. Exact Calculations for the Constant Profile

Using frequency domain transport eqns. (4.1,2) changes the solution for the hole profile in a region with a constant profile from

$$p = p_o(Ae^{-x/L_p} + Be^{x/L_p}) \quad (4.31)$$

where L_p^o is the static hole diffusion length, to

$$p = p_o(Ae^{-x/L_p^*} + Be^{x/L_p^*}) \quad (4.32)$$

where L_p^* is the complex diffusion length given by

$$\frac{1}{L_p^*} = \frac{1}{L_p^o} \sqrt{1 + j\omega\tau_p} \quad (4.33)$$

and p_o is the equilibrium hole density. The only change required is replacing the static diffusion length by its complex counterpart (of course, this implies that all calculations must now be done in the complex domain). For $\omega \rightarrow 0$, $L_p^* \rightarrow L_p^o$, but for $\omega > 1/\tau_p$, $L_p^* \rightarrow L_p^o/\sqrt{j\omega\tau_p}$ so the diffusion length decreases as $\omega^{-1/2}$ above $1/\tau_p$. The exact solution for the three cases under consideration yield the saturation currents

$$J_{o1}(\omega) = \frac{qD_p p_o}{L_p^o} \sqrt{1 + j\omega\tau_p} \left[\frac{1 + \exp(-2W/L_p^*)}{1 - \exp(-2W/L_p^*)} \right] \quad (\text{case 1}) \quad (4.34)$$

$$J_{o2}(\omega) = \frac{qD_p p_o}{L_p^o} \sqrt{1 + j\omega\tau_p} \left[\frac{1 - \exp(-2W/L_p^*)}{1 + \exp(-2W/L_p^*)} \right] \quad (\text{case 2}) \quad (4.35)$$

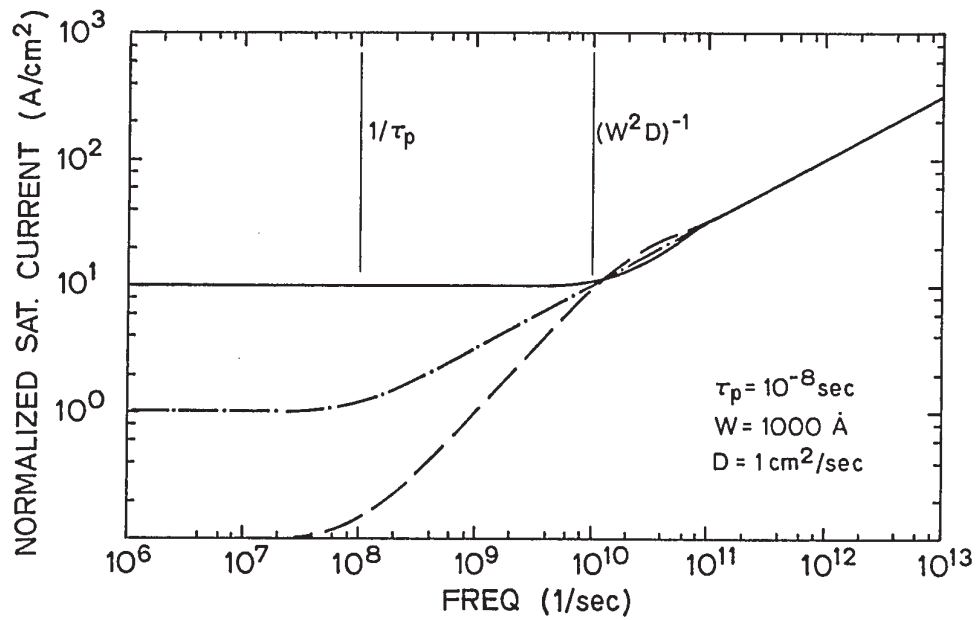
$$J_{o3}(\omega) = \frac{qD_p p_o}{L_p^o} \sqrt{1 + j\omega\tau_p} \quad (\text{case 3}). \quad (4.36)$$

The magnitude and phase components for all three cases are plotted in Fig. 4.2a,b where the J_o values have been normalized to

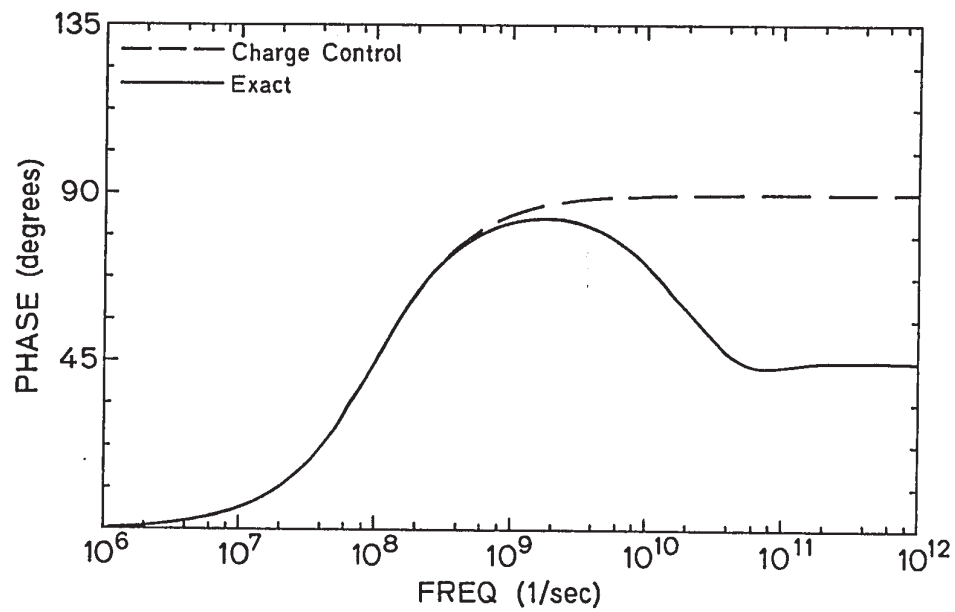
$$\frac{qD_p p_o}{L_p^o}$$

The thin transparent emitter has a static J_o two orders of magnitude larger than for case 2, the thin opaque emitter, but its J_o remains constant to $\omega = 1/\tau_f$ after which it increases as $\omega^{1/2}$. The phase response remains near zero until $1/\tau_f$ at which point it rapidly approaches its limiting value of 45° reflecting the $\sqrt{j}$ factor in the phase. The reason for this increase in J_o becomes apparent by examining the behavior of the complex diffusion length. For $\omega = 1/\tau_f$

$$\frac{1}{L_p^*} = \frac{1}{L_p^o} \sqrt{1 + j\frac{\tau_p}{\tau_f}}$$



(A)



(B)

Figure 4.2: exact $J_{sc}(\omega)$ solutions for magnitude (A), phase (B) for uniform emitter

since $\tau_p/\tau_f = 100$

$$\begin{aligned}\frac{1}{L_p^*} &= \frac{1}{L_p^o} \sqrt{j \frac{2\tau_p D_p}{W^2}} \\ &= \sqrt{j} \sqrt{\frac{2}{W^2}}\end{aligned}$$

i.e. $|L| \approx W$ so the diffusion length is reduced from its static value of $10W$ to being roughly comparable to W at $\omega = 1/\tau_f$. As ω increases beyond this point, L falls as $\omega^{-1/2}$, which explains the increase in J_o . In this regime, the emitter which was transparent at lower frequencies has become opaque.

For case 2, the thin emitter with the same width as case 1, but with the low surface recombination velocity, the static J_o is smaller than the best that can be attained with a metallic contact (given by case 3) by a factor of 10, but as predicted by the charge control eqn. (4.29), J_o increases proportionately to ω for $1/\tau_p < \omega < 1/\tau_f$. The phase response approaches 90° which is characteristic of a zero in the admittance function. For $\omega > 1/\tau_f$, the behavior approaches that of case 1 as expected since at this frequency, the emitters are becoming opaque, and opaque emitters by definition are insensitive to the surface boundary condition. For case 3, J_o increases for $\omega > 1/\tau_p$ and the phase response assumes a value of 45° . The curves for all three cases merge for $\omega > 1/\tau_f$ since they are all opaque above this frequency. In this region, dynamic components dominate and J_o assumes a value of

$$J_o(\omega) \rightarrow \frac{qn_{i0}^2}{N_{eff}} \sqrt{D_p} \omega^{1/2} \quad (4.37)$$

independent of its static value. Since N_{eff} is limited by BGN and since D_p displays only a weak dependence on the doping level, there is little that can be done to limit the value of J_o once N_{eff} has reached its limiting value of $1 \times 10^{18} \text{ cm}^{-3}$. The effects of varying the lifetime and the junction depth on the frequency dependence of J_o are displayed in Fig. 4.3 a,b. For case 1, increasing the lifetime leaves the characteristic unchanged since it is by assumption transparent. For case 2, this results in a lowering of both the static J_o and $1/\tau_p$ corner frequency until the surface recombination velocity dominates. Then both quantities saturate. Similar considerations apply

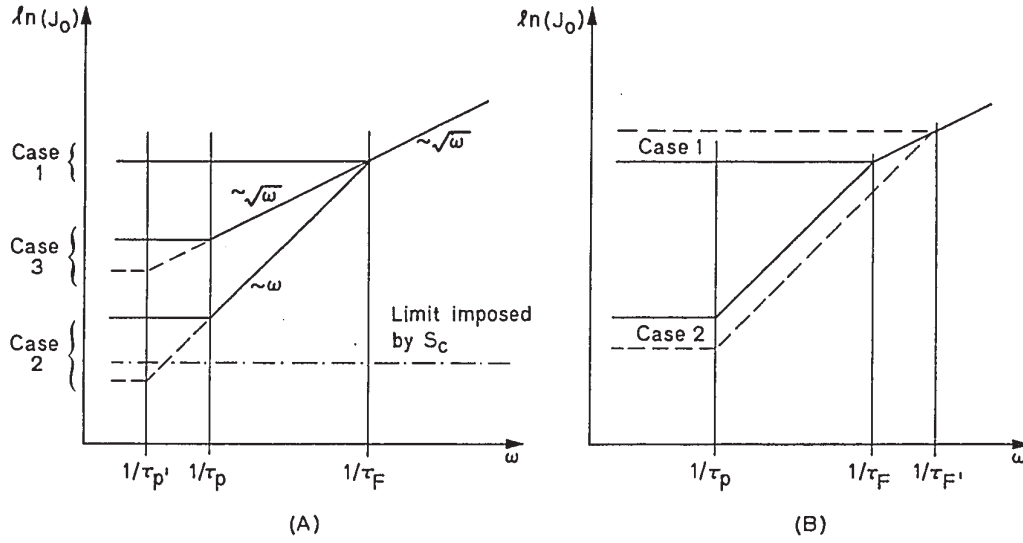


Figure 4.3: effect of changing hole lifetime (A) and junction depth (B) on frequency dependence of J_{oe}

for case 3. Reducing the junction depth for the thin transparent emitter results in an increase in both the static J_o and $1/\tau_f$ corner frequency in such a manner that the

$$J_o \cdot \tau_f^{-1/2} = \text{constant}$$

A larger bandwidth can be obtained at the expense of larger static values of J_o . This is a type of “gain-bandwidth product” for transparent junctions. For the thin opaque emitter, reducing W does not impact the lower $1/\tau_p$ corner frequency; however, as the $1/\tau_f$ frequency is scaled up as W , the entire low frequency is shifted down by a factor equal to the reduction in W , unless the value of the surface recombination velocity is high enough to dominate at low frequencies. This analysis indicates that while there is a tradeoff between static J_o and bandwidth for the thin transparent emitter, for low J_o structures using low S_c contacts, a reduction in W is beneficial to both static and dynamic characteristics,

4.5. Comparison with Charge Control

J_{oe} curves for the uniformly doped region resulting from the exact calculations are compared to their charge control counterparts for cases 1,2,3 in Fig. 4.4,5,6 respectively. For cases 1 and 3, charge control gives good prediction of J_{oe} for frequencies less than the inverse of the transit time and inverse of the minority carrier lifetime respectively. The phase response, which reflects the growing magnitude of the imaginary component of junction current is not so well predicted by charge control. Instead of the $\sqrt{j\omega}$ dependence predicted by the exact solution, charge control only predicts a simple zero behavior. This is an unavoidable consequence of assuming a quasi-static charge distribution which forces the real component of J_{oe} to remain constant, independent of frequency, when in reality at high frequencies, the real part increases at the same rate as the imaginary term.

For case 2, the single zero behavior is accurately predicted by charge control in both magnitude and phase. At the second frequency breakpoint, where $|L| < W$, the approximation loses validity as noted above. Instead of reverting back to a 45° phase response, charge control predicts a constant phase at 90°. These observations can be quantified by approximating the exact solutions by their Taylor series expansions and comparing them to the charge control counterparts. For case 1, expanding the exact solution to fourth order in W/L

$$J_{o1} = \frac{qD_p n_{io}^2}{N_{eff}W} \left(1 + \frac{\lambda^2}{3} - \frac{1}{45}\lambda^4\right) \quad \text{where } \lambda \equiv W/L_p^* \quad (4.38)$$

if $W < L_o/6$, then to within 1 %

$$\begin{aligned} J_o(\omega) &= \frac{qD_p n_{io}^2}{N_{eff}W} \left(1 + j\omega\tau_p \frac{W^2}{3L_p^2}\right) \\ &= \frac{qD_p n_{io}^2}{N_{eff}W} \left(1 + j\omega \frac{W^2}{3D_p}\right) \end{aligned} \quad (4.39)$$

This compares to the charge control result (see eqn. 4.28) which gives

$$J_{o1}(\text{chargecontrol}) = \frac{qD_p n_{io}^2}{N_{eff}W} \left(1 + j\omega \frac{W^2}{2D_p}\right) \quad (4.40)$$

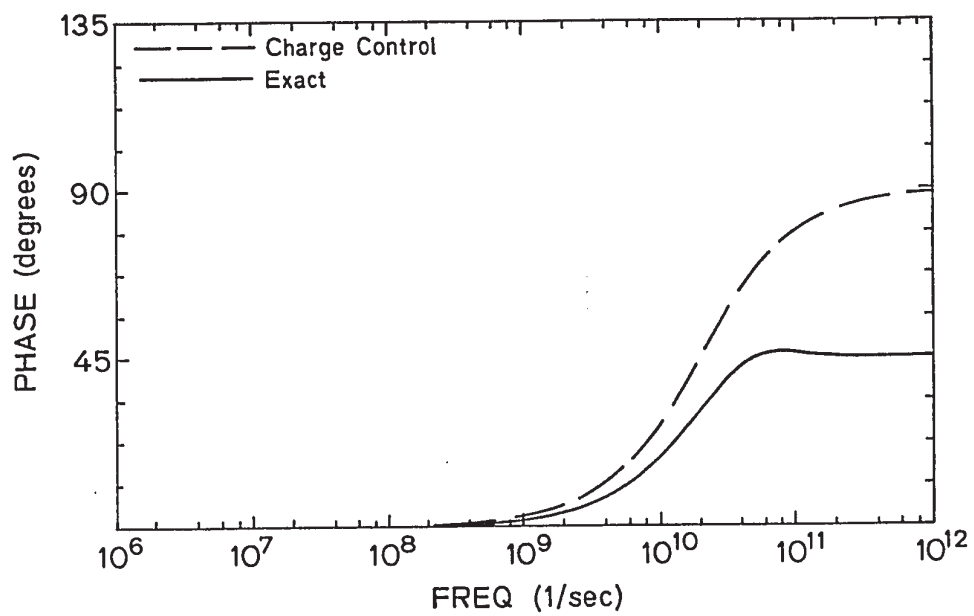
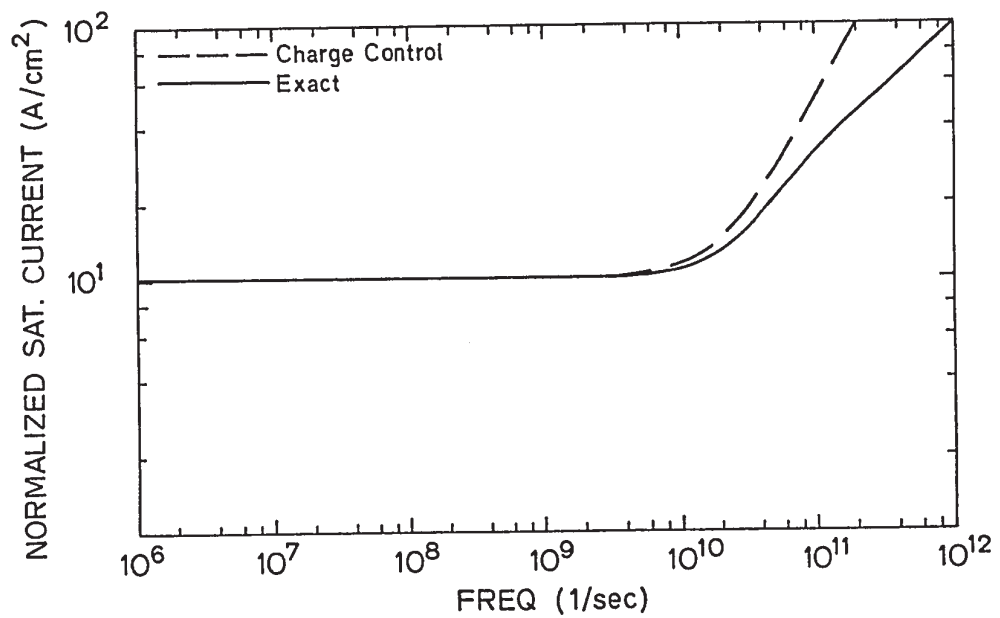


Figure 4.4: comparison of charge control and exact result, case 1; magnitude (A), phase (B)

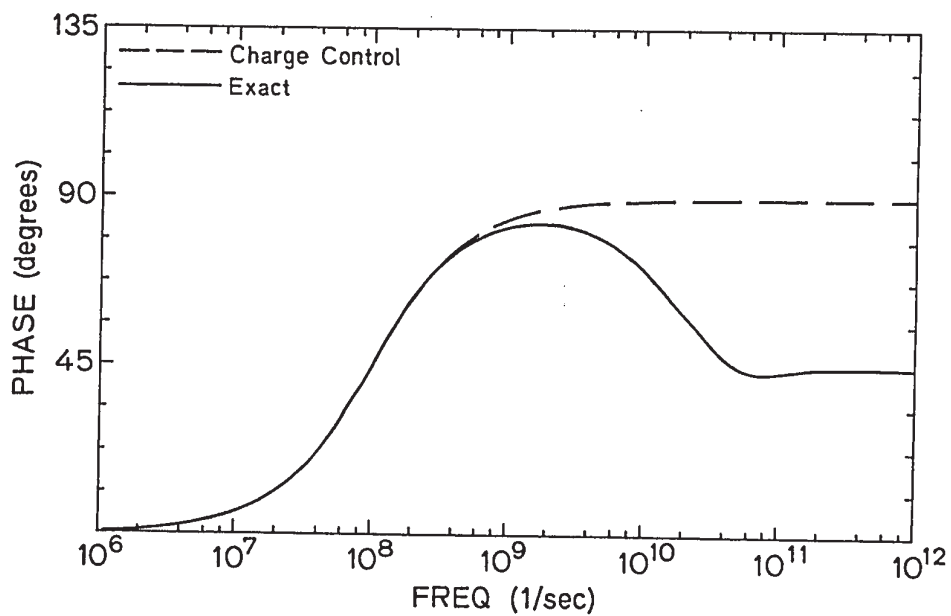
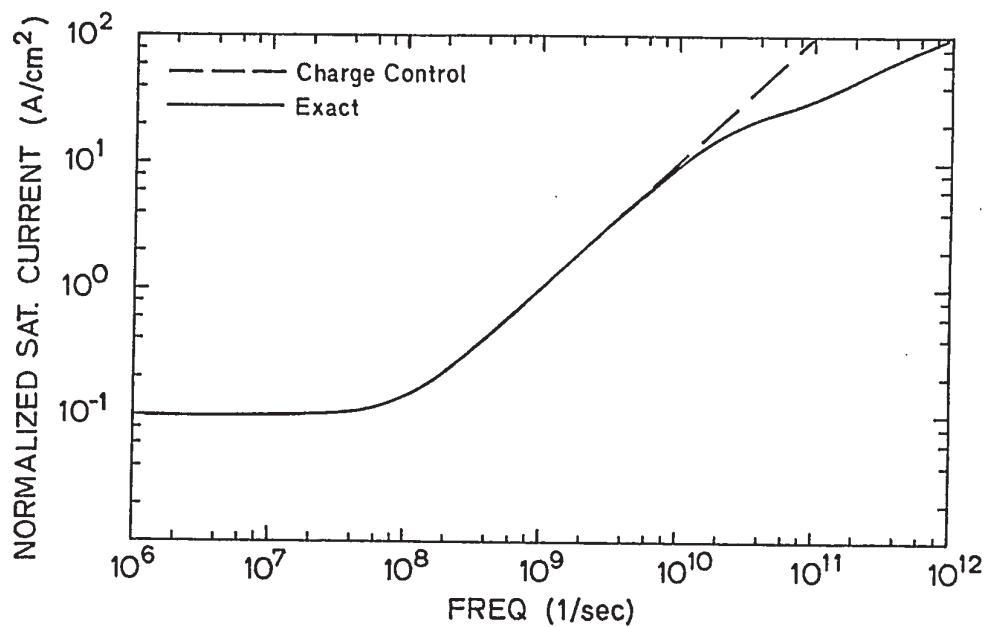


Figure 4.5: comparison of charge control and exact result, case 2; magnitude (A), phase (B)

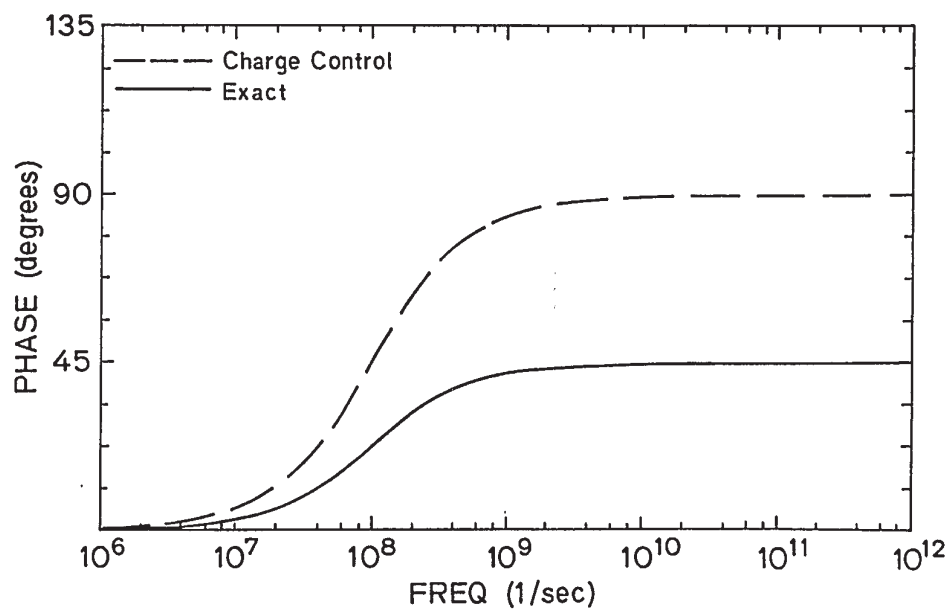
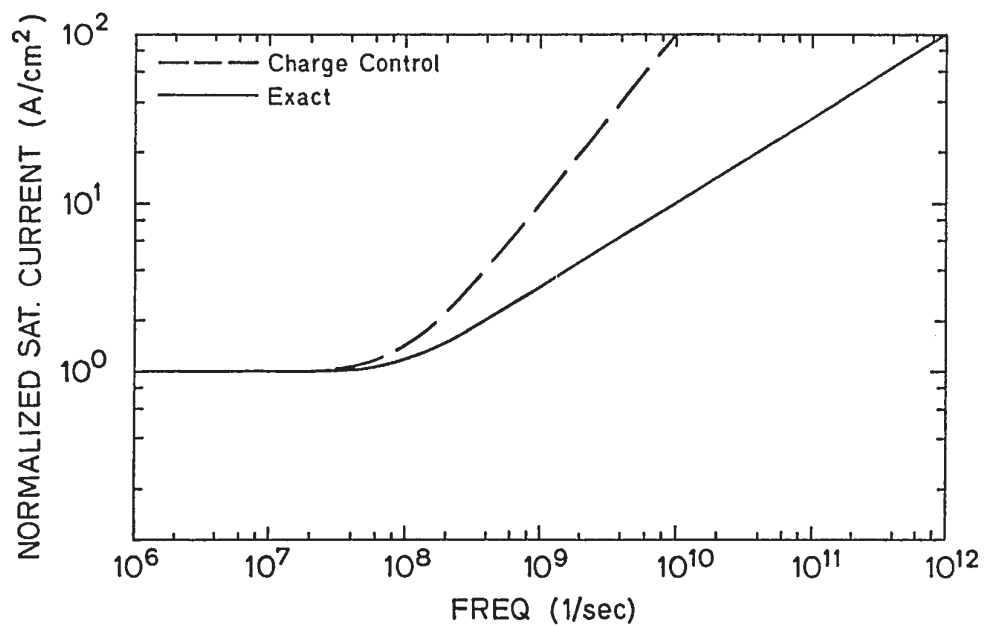


Figure 4.6: comparison of charge control and exact result, case 3; magnitude (A), phase (B)

where J_{o1} is the saturation current for case 1, related to J_{p1} by

$$J_{p1}(\omega) = J_{o1}(\omega)e^{V/V_T}$$

where V is the junction voltage. We can see that there is a difference in the imaginary components, where the factor of three in the denominator has been replaced by two in the charge control result. The fractional error in the imaginary component is seen to be independent of frequency which reflects the inadequacy of assuming that quasi-static conditions hold even for the limit of very small ω . The source of this discrepancy is identified in Appendix E.

For case 2, where $W < |L|$ but $S_c \rightarrow 0$, expansion of the exact solution gives

$$J_{o2}(\omega) = \frac{qD_p n_{i0}^2}{N_{eff}W} \lambda^2 \left(1 - \frac{\lambda^2}{3} + \frac{2}{15} \lambda^4\right) \quad (4.41)$$

where again $\lambda = W/L_p^*$ which upon expansion and regrouping, gives

$$J_{o2}(\omega) = \frac{qn_{i0}^2 W}{N_{eff}\tau_p} \left[1 - \frac{1}{3} \left(\frac{W}{L_p^o}\right)^2 (1 - \omega^2 \tau_p^2) + j\omega\tau_p\right] \quad (4.42)$$

The corresponding charge control expression (see eqn. 4.29) is

$$J_{o2}(\omega) = \frac{qn_{i0}^2 W}{N_{eff}\tau_p} (1 + j\omega\tau_p) \quad (4.43)$$

At $\omega = 1/\tau_p$ the expressions are identical. In fact charge control gives essentially the same solution as the exact expression until

$$1 \approx \frac{1}{3} \left(\frac{W}{L_p^o}\right)^2 (1 - \omega^2 \tau_p^2)$$

solving for ω

$$\omega = \frac{1}{\sqrt{\tau_f' \tau_p}} \quad \tau_f' = \frac{W^2}{3D_p}$$

As can be seen from Fig. 4.5b, the phase response deviates at the geometric mean of $1/\tau_p = 1 \times 10^8 \text{ sec}^{-1}$ and $1/\tau_f' = 1 \times 10^8 \text{ sec}^{-1}$. For case 3, the application of the binomial expansion to the exact result gives

$$J_{o3}(\omega) = \frac{qn_{i0}^2 D_p}{N_{eff}L_p^o} \left[\left(1 + \frac{\omega^2 \tau_p^2}{8}\right) + j\omega\tau_p \left(\frac{1}{2} + \frac{3}{48} \omega^2 \tau_p^2\right) \right] \quad (4.44)$$

For $\omega\tau_p < 1$, we can ignore the higher order terms which gives

$$J_{o3}(\omega) = \frac{qn_{io}^2 D_p}{N_{eff} L_p^o} \left(1 + j \frac{\omega\tau_p}{2}\right) \quad (4.45)$$

This compares with the charge control expression (see eqn. 4.30)

$$J_{o3}(\omega) = \frac{qn_{io}^2 D_p}{N_{eff} L_p^o} (1 + j\omega\tau_p) \quad (4.46)$$

The imaginary terms differ by a factor of two.

This analysis indicates that charge control gives passable results for the magnitude of J_{oe} up to the frequency limits of validity; however, the phase response which is sensitive to the error in the imaginary term, is rather poorly predicted. Charge control is therefore seen to be most useful as a tool that permits approximate determination of the dominant zeros (poles) and their linkage to specific physical mechanisms without yielding accurate numerical values especially of the phase response.

4.6. Nonconstant Profiles

Although diffused or implanted profiles are graded in some fashion, their overall qualitative behavior is quite similar to that of the constant profile. The corner frequencies may be altered by factors of order unity due to the effect of retarding fields for example, and the phase response can not be expected to be given by the constant profile results. Another significant distinction between the two cases is the presence of a lightly doped region near the junction in the graded profile. Due to the lower dopant densities, minority carrier storage is exacerbated and it may prove beneficial to use higher surface concentrations to reduce the effective width of this lightly doped region. The J_{oe} for an ERFC profile emitter with a fixed surface recombination velocity of 100 cm/sec with various surface concentrations is plotted in Fig. 4.7. J_{oe} exhibits a simple zero behavior increasing linearly with frequency up to the inverse of the transit time, which in this case lies in the neighborhood of

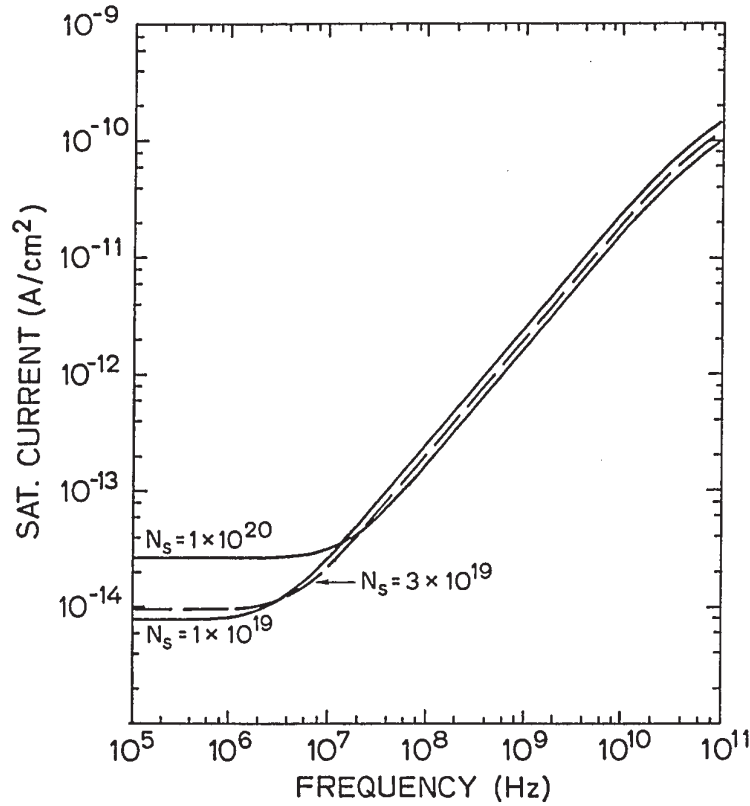


Figure 4.7: $J_{sc}(\omega)$ for ERFC emitter, parameter is surface concentration

$4 \times 10^{10} \text{ Hz}$. If extended to higher frequencies, J_{sc} would approach a $\sqrt{\omega}$ behavior as in the constant profile case. A surface concentration of $1 \times 10^{19} \text{ cm}^{-3}$ yields the best static J_{sc} but due to the more gradual dopant profile, the lightly doped region near the junction is wider than for the other curves where the surface concentrations are higher. This translates into extra minority carrier storage which increases the high frequency J_{sc} relative to the other curves. Increasing the surface concentration to $3 \times 10^{19} \text{ cm}^{-3}$ degrades the static J_{sc} slightly while improving the high frequency characteristic somewhat. A further increase to $1 \times 10^{20} \text{ cm}^{-3}$ marginally improves high frequency response, but at the expense of a substantial increase in the static

J_{oe} . For this particular structure, a surface concentration of $3 \times 10^{19} \text{ cm}^{-3}$ is seen to provide a good balance between low and high frequency performance, but for the majority of applications, the marginal improvement in high frequency response probably would not warrant the sacrifice in static J_{oe} . These results also indicate that abrupt doping profiles are to be preferred since they minimize the extent of this lightly doped region near the junction.

4.7. Validity of Assuming Constant S_c

Implicit in this discussion was that the dynamics in the X region for the low S_c structures could be neglected. The low values of S_c ($< 100 \text{ cm/sec}$) required to achieve J_o of $1 \times 10^{-14} \text{ A/cm}^2$ and less, are probably dominated by interfacial recombination. Even for the Si-SiO_2 interface, S_{ox} increases rapidly from a value of approximately 10 cm/sec at lower dopant densities for a well passivated interface as the dopant level is increased [2]. This implies that the contribution to the effective surface recombination velocity from transport in the X region must be even smaller than 100 cm/sec in practice.

Since the effective surface recombination velocity is given by

$$S_c = S_{ox} + \frac{p_X}{p_{Si}} v_{trans}$$

where v_{trans} is the transport velocity in the "X" region, p_X is the hole density in X, and p_{Si} the hole density at the surface of the surface of the monocrystalline region, we can see that the product

$$\frac{p_X}{p_{Si}} v_{trans}$$

must be significantly less than 100 cm/sec . For the types of materials we are considering, the transport velocity is comparable to that in the monocrystalline silicon. It is possible to conceive of materials where the reduction in J_o of the X region is due to a drastic reduction in the diffusion coefficient; however, as the diffusion velocity is given by D/τ , there must not be a concomitant decrease in the

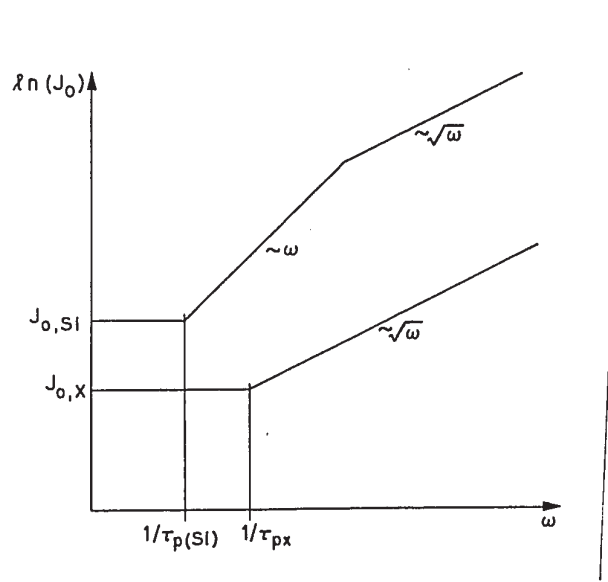


Figure 4.8: influence of J_{oX} on $J_{oe}(\omega)$ for POLY SIS contact

lifetime of the region, i.e. the defects that induce scattering must not generate or act as active recombination sites. For materials such as polysilicon, this is apparently not the case, and the transport velocity is quite similar to that in monocrystalline silicon. The reduction in J_o in region X must then be attributable to a significant reduction in the injected minority carrier density. This would be the case in a POLY SIS or barrier HJ contact structure (see Chapter 3).

Since it is not possible to prove in general that dynamics in the X region are not significant simply from the fact that hole transport in the X region does not impact static J_o , the case of the POLY SIS structure will be discussed to provide a concrete example. Referring to Fig. 4.8, J_{oX} , the saturation current of the X region is shown as a function of frequency. The static J_{oX} is shown much smaller than the static J_{oe} since interfacial recombination is assumed to dominate. The corner frequency is given by the inverse of the hole lifetime in the POLY. This lifetime is believed to be shorter than in the monocrystalline region of comparable doping, so that the corner frequency is higher for the POLY than for the monocrystalline

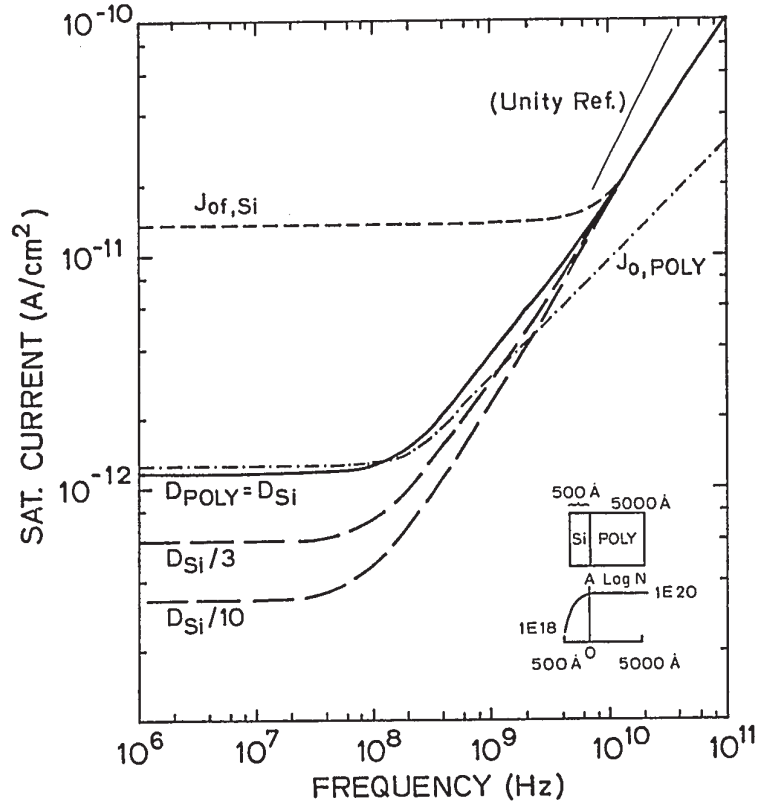


Figure 4.9: $J_{oe}(\omega)$ for a 5000 Å POLY + outdiffused emitter

region. The J_o in the monocrystalline silicon increases at a lower frequency than in the POLY and increases with a faster slope of unity as opposed to the 1/2 slope assumed for the POLY since it is an opaque region. It is clear from the plot that J_{oX} will not affect J_{oe} significantly. For cases such as the HJ where J_o reduction relies on a reduction of the transport velocity in the X region (as in “intimate” POLY contacts), the dynamics of each region must be calculated and the J_o of the overall structure computed from the two port parameters using the equivalent circuit. A calculation of J_{oe} for an intimate POLY contact using this latter approach is shown in Fig. 4.9. This structure consists of a 5000 Å POLY layer from which

dopant has outdiffused into the monocrystalline region to form a 500 Å junction with an ERFC profile. The POLY region was modeled assuming the same minority carrier transport constants in monocrystalline silicon, the only difference being the assumption of diffusion coefficients equal to and smaller than the monocrystalline region by factors of 3 and 10. In the low frequency regime, the transport impedance of the POLY layer dominates and since the POLY region is essentially opaque at the assumed doping density of $1 \times 10^{20} \text{ cm}^{-3}$, J_{oe} falls as the square root of the diffusion coefficient. Once the frequency exceeds the inverse of the hole lifetime, J_{oe} might be expected to increase as the square root of frequency as would be the case if only the POLY layer were present. Instead, it increases at a much faster rate, a behavior that can be understood by realizing that the 500 Å outdiffused region has significant minority carrier storage due to its effectively lower doping. Looking at Fig. 4.9, the J_o of the POLY and ERFC regions are plotted separately and for frequencies above $1 \times 10^{10} \text{ Hz}$, it is clear that the monocrystalline region is essentially “shorting out” the POLY layer from the viewpoint of minority carrier transport. In this case, outdiffusion is detrimental to performance. Of course, this conclusion only obtains if the neglect of interfacial recombination is valid. If this is not the case, optimization must be carried out using the equivalent circuit with the interfacial recombination element included. In that case, some outdiffusion may be beneficial, although its optimal extent depends upon the value of the interfacial recombination velocity and the relative importance of low and high frequency performance.

4.8. SUMMARY

In this section, the high frequency behavior of J_{oe} has been examined using the charge control approximation and numerical simulations. It has been demonstrated that with the onset of high frequency characteristics, J_{oe} is controlled by the dynamic terms and eventually approaches the same asymptotic behavior independent of its static value. The fundamental goal in optimizing high frequency performance

is seen to be the minimization of minority carrier storage, a strategy which is not always compatible with the attainment of the lowest values of static J_{oe} . Profile optimization is a global problem especially when other considerations such as depletion layer charge storage and junction capacitances, which were not considered here, are also taken into account. In the next section, the effect of the frequency dependence of J_{oe} on the frequency performance of the BJT will be examined to see if emitter dynamics, which are ignored in most treatments can become significant when low J_{oe} emitters are used in these structures.

4.9. Bipolar Junction Transistors

The applications of low J_{oe} emitter structures to bipolar transistors in low frequency applications was outlined in chapter 2. It remains to be seen how high frequency performance is affected by the incorporation of such structures. There are two key figures of merit which in some sense characterize the suitability of a transistor for high frequency operation. These are f_T , the unity current gain frequency and f_{max} , the maximum frequency of oscillation. In this section, a limited analysis of the impact of low J_{oe} emitter structures on these parameters will be discussed.

4.10. Current Gain

At low frequencies, the current gain for the common emitter structure, the configuration assumed in this analysis, was shown to be determined by recombination mechanisms in various regions of the device. As the frequency of operation increases, the base current is dominated by the dynamic components that flow in response to time varying distributions of charge throughout the device. The base current must supply the time varying hole charge in the quasi-neutral emitter and emitter-base space charge region, and must neutralize the time varying electron charge in the quasi-neutral base and base-collector space charge regions. The analysis in this

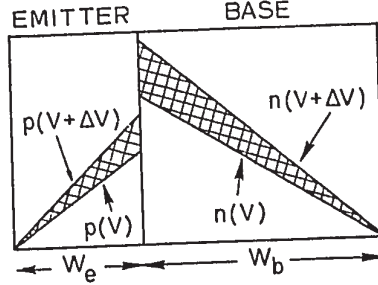


Figure 4.10: simplified model for charge control analysis of BJT

section will be limited in the sense that only the components corresponding to the quasi-neutral emitter and base regions will be considered. This implies neglect of junction capacitances, and space charge region storage. Although they are not negligible in general, they do not fall within the realm of this discussion.

4.11. Charge Control

Referring to Fig. 4.10, we can see the two components of base current that are under consideration. The solid lines represent hole and electron profiles in the quasi-neutral emitter and base regions respectively at two different emitter-base voltages. The cross-hatched regions represent the amount of charge difference between the two distributions. The net change of hole charge in the QN (quasi-neutral) emitter must be supplied by injection from the base region. The change in electron charge in the QN base must be neutralized by an equal amount of holes, again supplied by the base terminal.

Using charge control, we can write the total base current density as

$$J_b = \frac{Q'_e}{\tau_e} + \frac{Q'_b}{\tau_b} + \frac{dQ'_e}{dt} + \frac{dQ'_b}{dt} \quad (4.47)$$

where τ_e represents the charge control time constant in the emitter and τ_b is essen-

tially the electron lifetime in the base. For modern high speed devices

$$\frac{Q'_e}{\tau_e} > \frac{Q'_b}{\tau_b}$$

reflecting the fact that static beta is limited by reverse injection into the emitter.

Now the collector current density can be written

$$J_c \equiv \frac{Q'_b}{\tau_{bf}} \quad (4.48)$$

where τ_{bf} is the base transit time. Since

$$\frac{Q'_e}{\tau_e} = \frac{1}{\beta_o} \frac{Q'_b}{\tau_{bf}} \quad (4.49)$$

therefore

$$Q'_e = \frac{1}{\beta_o} \tau_e J_c \quad (4.50)$$

substituting into eqn. (4.47), we get

$$J_b = \frac{1}{\beta} J_c + j\omega \frac{\tau_e}{\beta_o} J_c + j\omega \tau_{bf} J_c \quad (4.51)$$

note that τ_e is effectively reduced by the static beta. Since

$$J_b \equiv \frac{J_c}{\beta_o} \quad (4.52)$$

substitution yields

$$\beta(\omega) = \frac{\beta_o}{1 + j\omega(\beta_o \tau_{bf} + \tau_e)} \quad (4.53)$$

Charge control predicts that beta starts to roll off at $(\beta_o \tau_{bf})^{-1}$ or τ_e^{-1} whichever is less.

For the thin transparent emitter case, we can write

$$\tau_e \approx \frac{W_e^2}{3D_p} \quad (4.54)$$

which may be comparable to the base transit time, τ_{bf} . Since

$$\beta_o = \frac{Q'_b}{Q'_e} \frac{\tau_e}{\tau_{bf}} \quad (4.55)$$

it can be seen that for the transparent case, β relies on maintaining the ratio $Q'_b/Q'_e > 1$. For this type of emitter,

$$\beta_o \tau_{bf} = \frac{Q'_b}{Q'_e} \tau_e > \tau_e \quad \text{since } Q'_b/Q'_e > 1 \quad (4.56)$$

Therefore, the $\beta_o \tau_{bf}$ term dominates the beta rolloff and emitter dynamics do not play a pivotal role in determining device behavior. For low J_{oe} structures where τ_e is given instead by the hole lifetime in the emitter, we have

$$\beta_o = \frac{\tau_e}{\tau_{bf}} \frac{Q'_b}{Q'_e} \quad (4.57)$$

and since $\tau_e \gg \tau_{bf}$ for high speed devices with their thin base regions, it may be possible to have $Q'_b \approx Q'_e$. If this condition is met, then

$$\beta_o \tau_{bf} = \frac{Q'_b}{Q'_e} \tau_e \approx \tau_e \quad (4.58)$$

and emitter dynamics may in fact be an important factor. To see when this condition is met, recall that the lifetime predicted using the model of del Alamo [1] for a doping level of $1 \times 10^{19} \text{ cm}^{-3}$ is $5 \times 10^{-8} \text{ sec}$. The transit time for a 1000 Å base is approximately $2 \times 10^{-11} \text{ sec}$. Therefore if the static beta using this type of emitter structure is less than 2000, then emitter dynamics may be significant.

Charge control indicates that in a conventionally contacted emitter, it is difficult for emitter dynamics to play a significant role; the beta rolloff is dominated by the current needed to neutralize the time varying component of electron charge in the QN base region. f_T , the frequency at which the common emitter current gain falls to unity, is given by

$$f_{T,1} = 1/\tau_{bf} \quad (4.59)$$

For low J_{oe} structures where bulk recombination dominates, it may be possible in nonoptimal structures to have $\tau_e > \beta_o \tau_{bf}$ in which case

$$f_{T,2} = 1/\tau_e \quad (4.60)$$

These values should only be taken as rough estimates of the actual values because of the limitations on the accuracy of charge control as pointed out in the previous section.

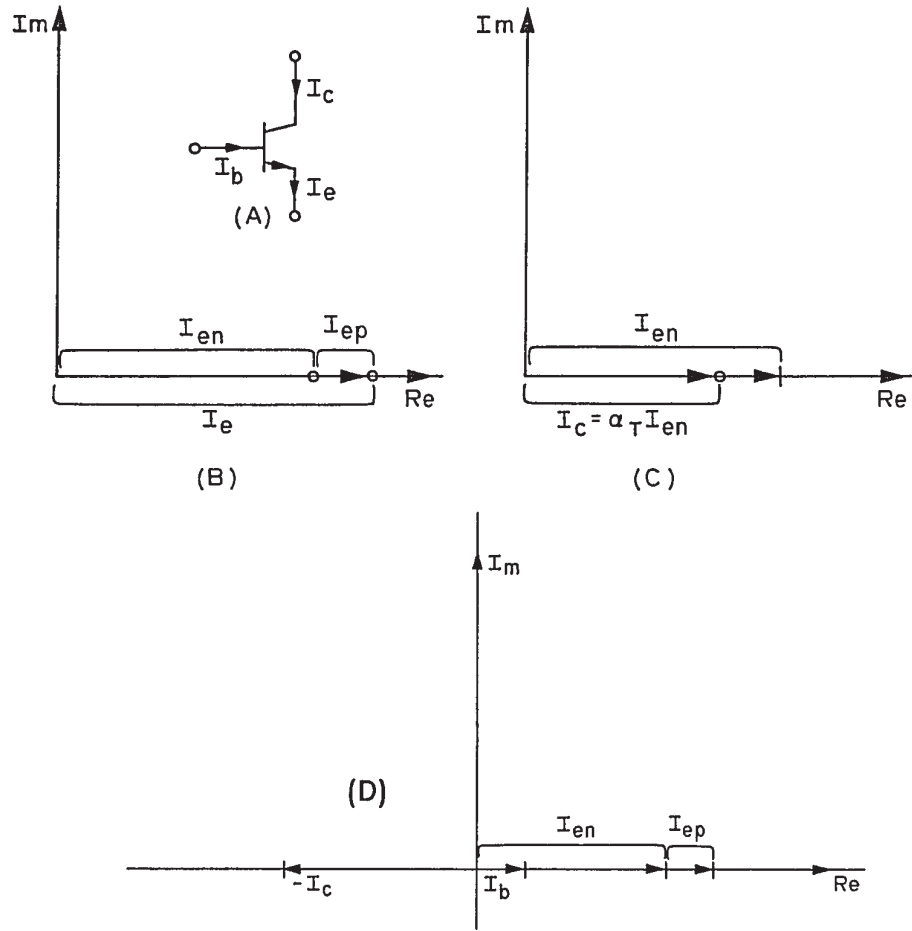


Figure 4.11: vector diagrams for BJT terminal currents: definition of currents (A), emitter current (B), collector current (C), sum of (B)+(C) to form base current (D)

4.12. Phasor Viewpoint

Another way to view the rolloff of beta at high frequencies is to look at the phasor representation for the terminal currents [3]. Referring to Fig. 4.11, we can write vectorially

$$\vec{I}_e = \vec{I}_b + \vec{I}_c \quad (4.61)$$

$$\vec{I}_b = \vec{I}_e - \vec{I}_c \quad (4.62)$$

Now $\vec{I}_c$ consists of several components of which the two under consideration are shown in Fig. 4.11b. There is a component $\vec{I}_{e,n}$ reflecting electron injection into the QN base and a component $\vec{I}_{e,p}$ reflecting reverse injection of holes into the emitter. At DC, these are all real quantities. In Fig. 4.11c, the fraction of the electron current that survives the base transit constitutes the collector current. For a thin base, this fraction is essentially unity. In Fig. 4.11d, the vectorial sum formed by taking $-\vec{I}_c + \vec{I}_e$ yields the base current as shown, where the magnitude of $\vec{I}_b$ is exaggerated for clarity. As the frequency increases to $1/(\beta_o \tau_{bf})$, the collector current $\vec{I}_c$ exhibits a phase shift with respect to $\vec{I}_{e,n}$ which is given to first order by

$$e^{j\omega \tau_{bf}}$$

This term represents a linear phase which is equivalent to a pure time delay of τ_{bf} which is consistent with the interpretation of τ_{bf} as a “transit time” through the base. When $\omega = 1/(\beta_o \tau_{bf}) \equiv \omega_{\beta 1}$ then the phase shift is given by

$$e^{j/\beta_o}$$

Referring to Fig. 4.12a we can see that the base current needed to neutralize the time varying electron charge is jI_c/β_o which has the same magnitude as the static value for the thin transparent emitter case. Since this is a quadrature component, beta falls by $\sqrt{2}/2$ as expected from the definition of $\omega_{\beta 1}$ as the beta rolloff corner frequency. Note that since

$$\tau_e < \beta_o \tau_{bf}$$

emitter dynamics do not play a role as indicated by the representation of $\vec{I}_{e,p}$ by a real phasor with zero phase.

For the thin opaque emitter where τ_e is given by the hole lifetime, if $\beta_o \tau_{bf} < \tau_e$ then for $\omega \equiv \omega_{\beta 2} = 1/\tau_e$ the phasor diagram is shown in Fig. 4.11b. At this frequency, $\vec{I}_{e,p}$ exhibits a 45° phase shift. Since the frequency is assumed to be much less than $1/(\beta_o \tau_{bf})$, I_c does not exhibit a significant phase shift with respect

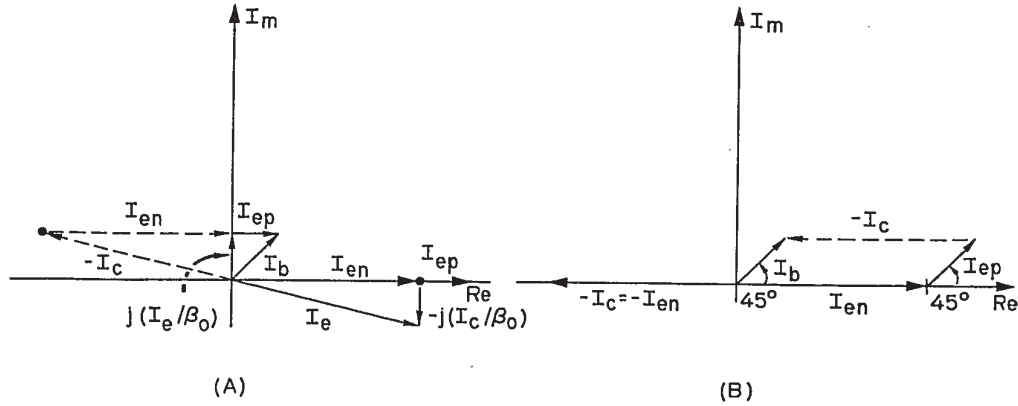


Figure 4.12: vector diagrams for BJT at $f = f_T$: base transparency limited (A), emitter limited (B)

to $\vec{I}_{e,n}$. The vectorial sum is given by $\vec{I}_b = \vec{I}_{e,p}$ since the base transparency has been assumed to be unity. Since

$$I_b = (1 + j)I_b(\omega = 0) \quad \text{at } \omega = 1/\tau_e \quad (4.63)$$

beta has fallen off by $1/\sqrt{2}$ as expected at the corner frequency which is now $1/\tau_e$. If the condition $\beta\tau_{bf} < \tau_e$ is not satisfied, then the beta rolloff is dominated by the phase lag in the base transport factor just as in the thin transparent emitter case.

4.13. Impact of Emitter and Base Structures on Device Performance

To see if emitter dynamics play a significant role in practice, numerical simulations of four device structures were executed. The relevant parameters for the base and emitter regions are listed in Table 4.1. For each structure, the emitter and base saturation currents were computed as a function of frequency. Then, assuming a saturated drift velocity of $1 \times 10^7 \text{ cm/sec}$ for electrons at the edge of the

base-collector space charge region, the current gain was calculated from

$$\beta = \frac{\alpha_b J_{ob}}{J_{oc} + J_{ob}(1 - \alpha_b)} \quad (4.64)$$

where the base transparency is given by

$$\alpha_b = \frac{\alpha_{bf} J_{bs}}{J_{bs} + J_{ob,r}(1 - \alpha_{bf} \alpha_{br})} \quad (4.65)$$

is the base transparency, J_{ob} is the base region saturation current given by:

$$J_{ob} = J_{ob,f} \frac{J_{bs} + J_{ob,r}(1 - \alpha_{bf} \alpha_{br})}{J_{bs} + J_{ob,r}}. \quad (4.66)$$

J_{bs} represents the effect of the velocity saturation which can be modeled as a “contact” to the base region with $s = 1 \times 10^7 \text{ cm/sec}$, i.e. as an equivalent

$$J_{bs} = \frac{qn_{i0}^2}{N_b W_b} v_{sat} \quad (4.67)$$

$J_{ob,f}$ and $J_{ob,r}$ are the forward and reverse saturation currents that pertain to the neutral base region, and α_{bf} and α_{br} are the corresponding transparency factors (see ref. [1]). To evaluate the relative influence of loss of emitter injection efficiency and loss of base transparency to beta rolloff, the “emitter efficiency” limited beta which assumed a unity base transport coefficient

$$\beta_\gamma \equiv \frac{\gamma}{1 - \gamma} \quad (4.68)$$

where

$$\gamma = \frac{J_{ob}}{J_{ob} + J_{oc}} \quad (4.69)$$

and a “base transport” limited beta which assumed a unity emitter efficiency

$$\beta_\gamma \equiv \frac{\alpha_b}{1 - \alpha_b} \quad (4.70)$$

were also computed. For each device, the magnitude and phase of the current gain and the saturation currents for the base and emitter are presented. It should be noted that this analysis neglects storage in the depletion regions as well as the

displacement currents due to the junction capacitances. Proper analysis of depletion layer dynamics is a difficult problem and lies outside the realm of this analysis. Another deviation from reality is the assumption of a monotonically decreasing analytic form for the base region doping profile. In practice, there may be a thin region near the emitter base junction that introduces a retarding field over part of the QN base; this will act to effectively increase the transit time, but the effect is small [5] if the retarding region represents a small fraction of the total base width. Neglect of these factors implies that the f_T presented here are upper bounds. To what extent f_T in practice falls short of these values depends upon the particular device geometry and operating point.

The emitter doping profile for device A which is representative of device structures currently in use, is shown in Fig. 4.13. It utilizes an emitter profile that was obtained from a 900°C 30 minute POCl_3 predep [6] and a $3000 \text{ \AA}$ wide base. Inspection of the beta vs frequency characteristic in Fig. 4.14a,b reveals that the DC β is limited by reverse injection into the emitter. If base dynamics were negligible, then the current gain would follow the emitter injection efficiency limited curve. This curve remains relatively flat until $\omega \approx 1/\tau_e$ at which point the injection efficiency falls because J_{oe} starts to increase. At $\omega \approx 1/\tau_{bf}$ the base starts to lose transparency, and J_{oe} starts to increase which causes the injection efficiency to eventually level off. Of course, once the base loses transparency the actual current gain lies far below this upper limit. In practice, as indicated by the charge control results, the base current increase is controlled by a dominant pole. Emitter dynamics are not significant in this device structure as can be seen by the magnitude and phase of β above the corner frequency. The role of the emitter in this case is simply to set the DC β , and hence the β rolloff corner frequency.

Seeking to increase the f_T of the device, narrower base regions will be used. The profile for device B narrows the base width to $1000 \text{ \AA}$ and increases the dopant density to keep the base sheet resistivity relatively constant as shown as shown by the entries in Table 4.1. The emitter junction depth is reduced to $2000 \text{ \AA}$ and

Device	A	B	C	D
EMITTER				
N_{surf} (cm^{-3})	4×10^{20}	1×10^{20}	3×10^{19}	1×10^{20}
$N_{E-B \text{ junction}}$ (cm^{-3})	3×10^{17}	2×10^{18}	2×10^{18}	2×10^{18}
Profile	Fig. 4.13	ERFC	ERFC	ERFC
Total Dose (cm^{-2})	3.8×10^{15}	6.7×10^{14}	6.2×10^{13}	1.7×10^{14}
$\rho_s(\Omega/\square)$	18	100	980	400
Contact	Metal	Metal	SIS or HJ $S_c = 100$	5000ÅPOLY $N = 10^{20}$ $D = D_{Si}/3$
Junction Depth (Å)	Fig. 4.13	2000	500	500
BASE				
$N_{E-B \text{ junction}}$ (cm^{-3})	3×10^{17}	2×10^{18}	2×10^{18}	2×10^{18}
$N_{B-C \text{ junction}}$ (cm^{-3})	1×10^{16}	1×10^{16}	1×10^{16}	1×10^{16}
Profile	Fig. 4.13	EXP	EXP	EXP
Total Dose (cm^{-2})	2.5×10^{12}	3.7×10^{12}	3.7×10^{12}	3.7×10^{12}
$\rho_s(K\Omega/\square)$	7.7	10	10	10
Current Gain				
Static β	123	43	13000	225
$f_\beta(Hz)$	7×10^6	7×10^7	2×10^5	1.3×10^7
$f_T(Hz)$	8×10^8	3×10^9	3×10^9	3×10^9

Table 4.1: parameters and results for BJT simulations

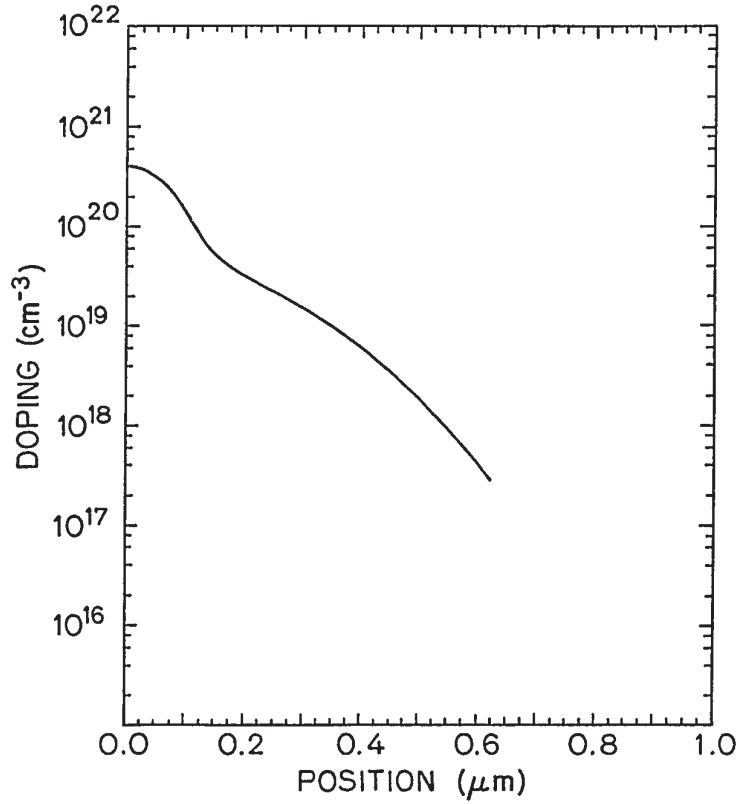


Figure 4.13: emitter dopant profile for Device A simulation

is modelled by an ERFC profile. As the β characteristic in Fig. 4.16 shows, the dynamic limitations imposed by the base region are relaxed to higher frequencies, but the thinner emitter also defers emitter dynamics to higher frequencies. As the characteristic reveals, the bandwidth of the device has been substantially increased albeit at the expense of static current gain. This is the trend discussed in Chapter 2. The plots of J_{oe} and J_{ob} for this device shown in Fig. 4.17 indicate that although the emitter has a lower corner frequency than the base due primarily to the presence of a retarding field in the emitter as opposed to the presence of an accelerating field in the base, the division of τ_e by β_o as indicated by eqn. (4.51) from the charge control analysis implies that the base region still controls the beta frequency

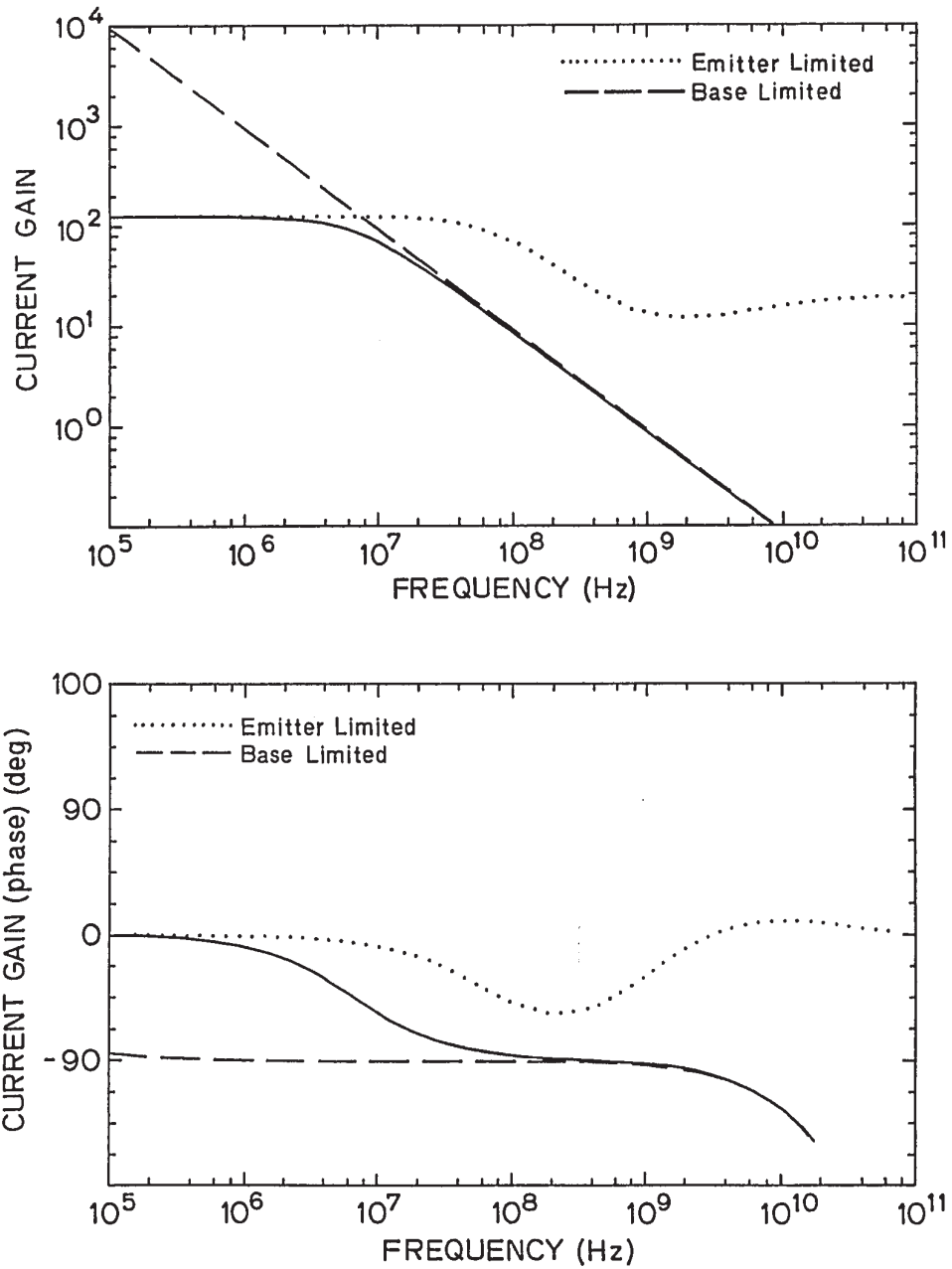


Figure 4.14: simulation of β vs. freq. for Device A: magnitude (A), phase (B)

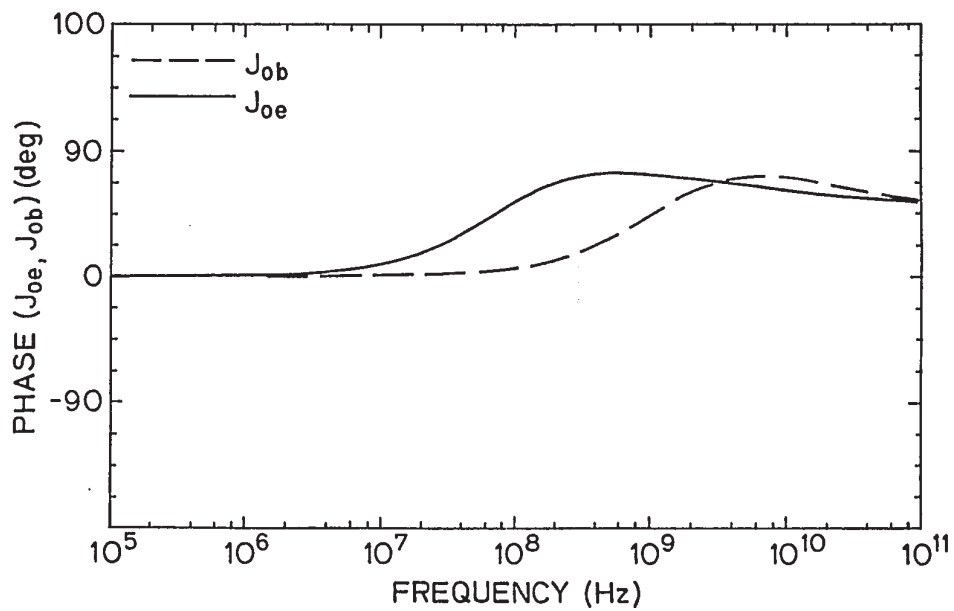
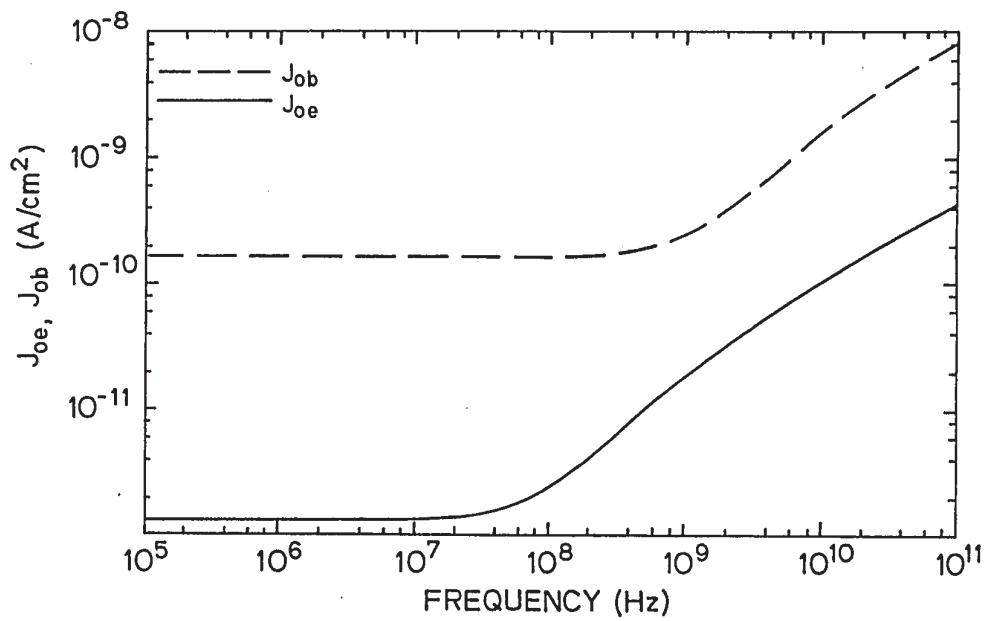


Figure 4.15: simulation of J_{oe}, J_{ob} vs. freq. for Device A

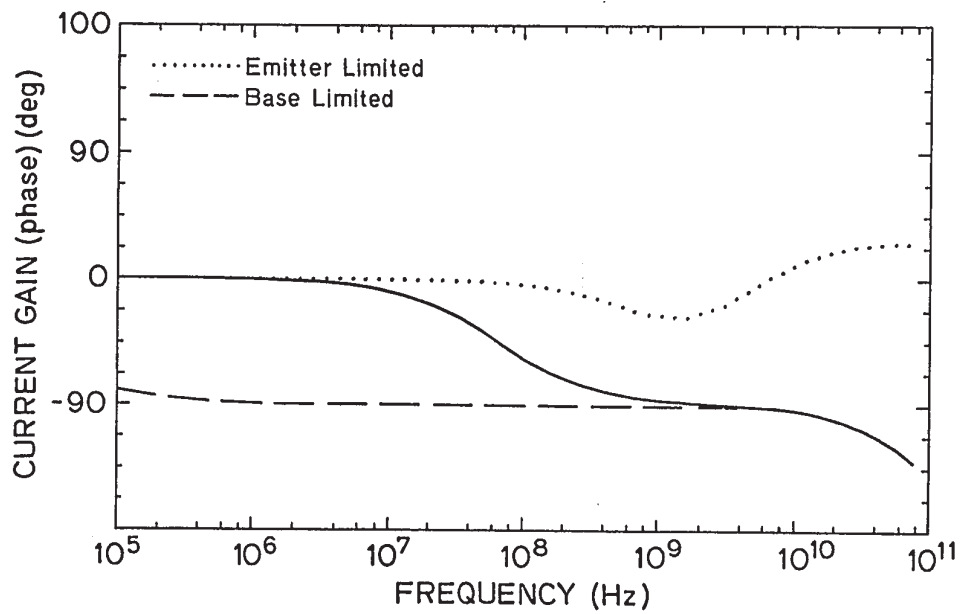
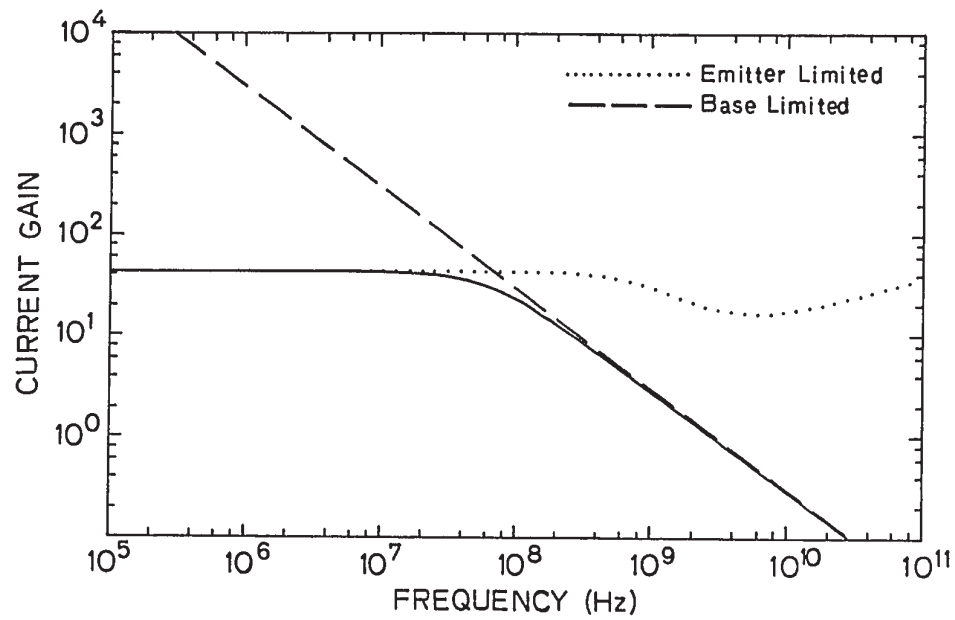


Figure 4.16: simulation of β vs. freq. for Device B: magnitude (A), phase (B)

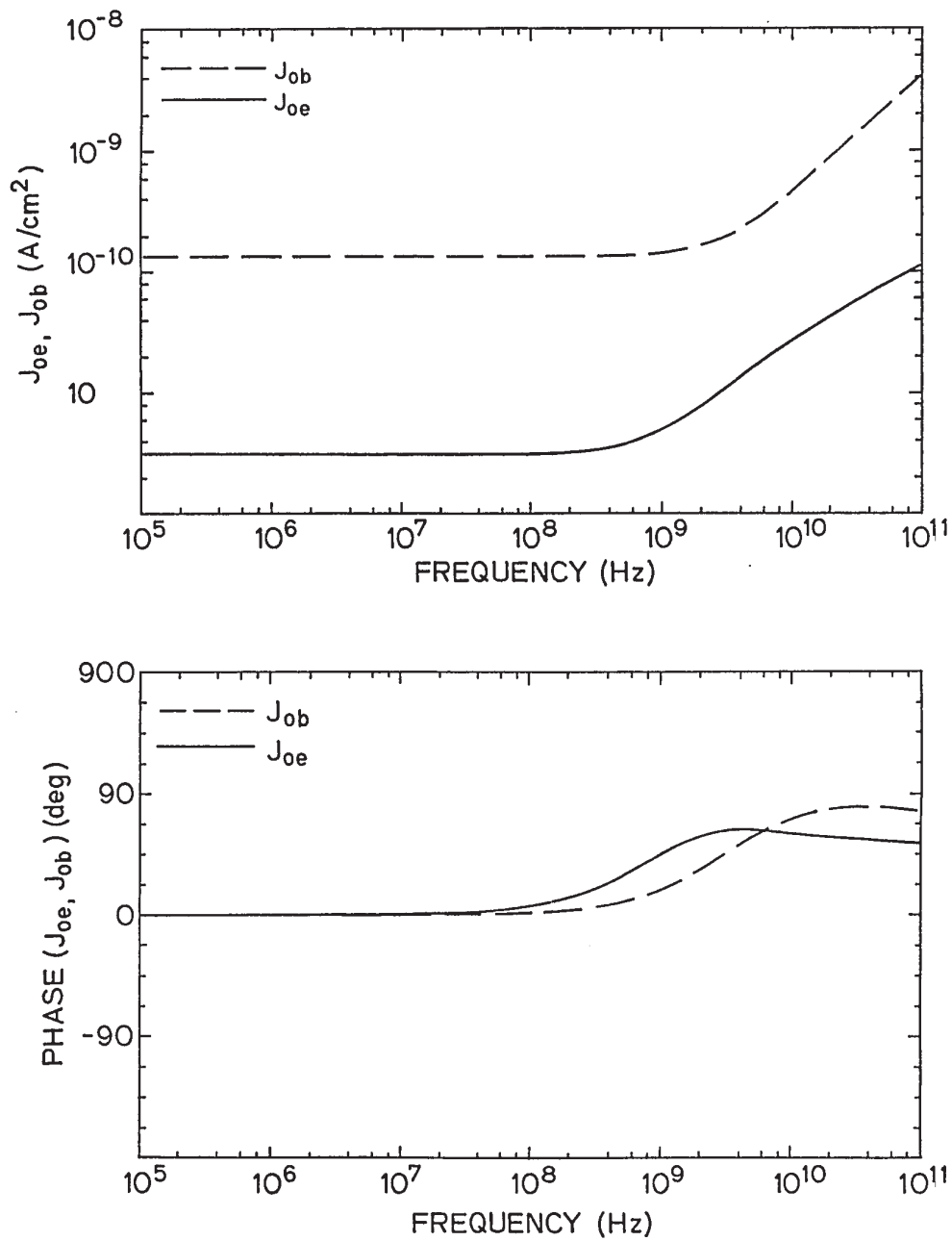


Figure 4.17: simulation of J_{oe}, J_{ob} vs. freq. for Device B: magnitude (A), phase (B)

characteristic. Reduction of the base width by a factor of three only results in an increase in f_T by about a factor of 3.3 instead of the factor of 9 expected. This is attributable in part to the reduction of the electron diffusion coefficient associated with the increased base doping. The effects of BGN are also becoming apparent at these dopant densities; their net effect is to reduce the magnitude of the built-in field which results in a smaller f_T than would have otherwise been expected.

The use of selective contacts in low J_{oe} emitter structures can regain this static beta loss as can be seen from the β plots for Device C in Fig. 4.18, but the role of emitter dynamics must be more closely scrutinized since J_{oe} now begins to increase at a much lower frequency. Device C uses the same base profile as device B, but the emitter uses a much shallower 500 ÅERFC profile with an effective 100 cm/sec contact recombination velocity. Looking at the plot of J_{oe} and J_{ob} in Fig. 4.19 reveals a striking difference in the emitter behavior between the two cases. J_{oe} now increases rapidly at 4×10^6 Hz; yet although $1/\tau_{bf}$ is about three orders of magnitude higher, division of τ_e by β_o which exceeds 1×10^4 implies that even in this case, emitter dynamics are not important. It is possible to construct a device where the emitter can limit frequency performance simply by doping the emitter much more lightly, thereby increasing hole storage, but this is not an optimal design for either the static or high frequency J_{oe} .

The last device structure, device D, incorporated the same base as used in devices B and C. Inspection of Fig. 4.20 reveals that the "beta envelope" is significantly enhanced over device B with respect to the static beta and over device A with respect to both static β and f_T . Yet, device C with the thin low surface recombination velocity emitter structure offers the best performance with static β over 1×10^4 . This "excess" beta can be useful in some applications as demonstrated in Chapter 2, but it is the possibility of trading off this excess beta for a larger base Gummel number (increased base doping) that presents possibilities of enhancing high frequency performance. This is an option that is present to a far lesser extent for device D and not even a viable option for device B.

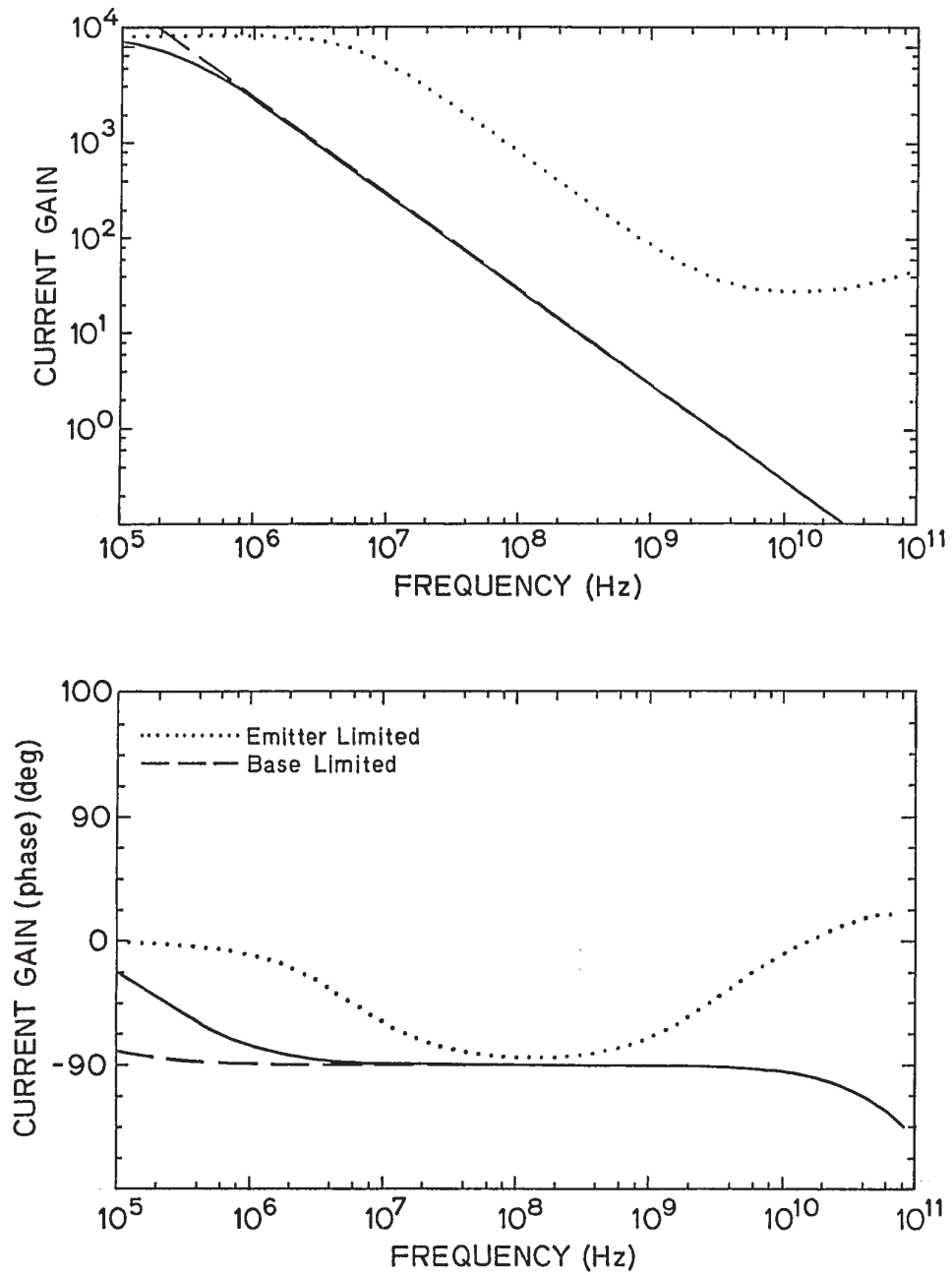


Figure 4.18: simulation of β vs. freq for Device C: magnitude (A), phase (B)

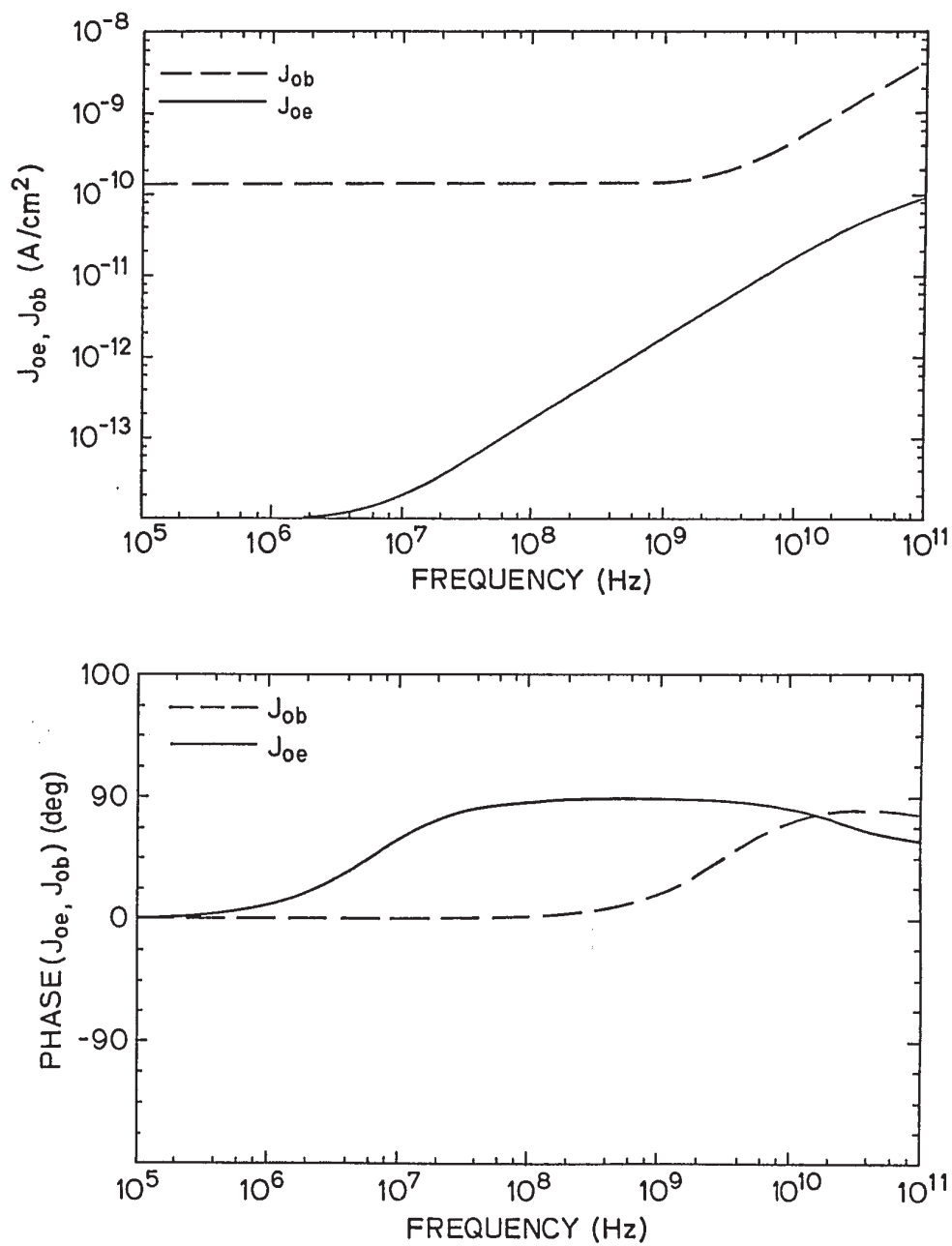


Figure 4.19: simulation of J_{oe} , J_{ob} vs. freq for Device C: magnitude (A), phase (B)

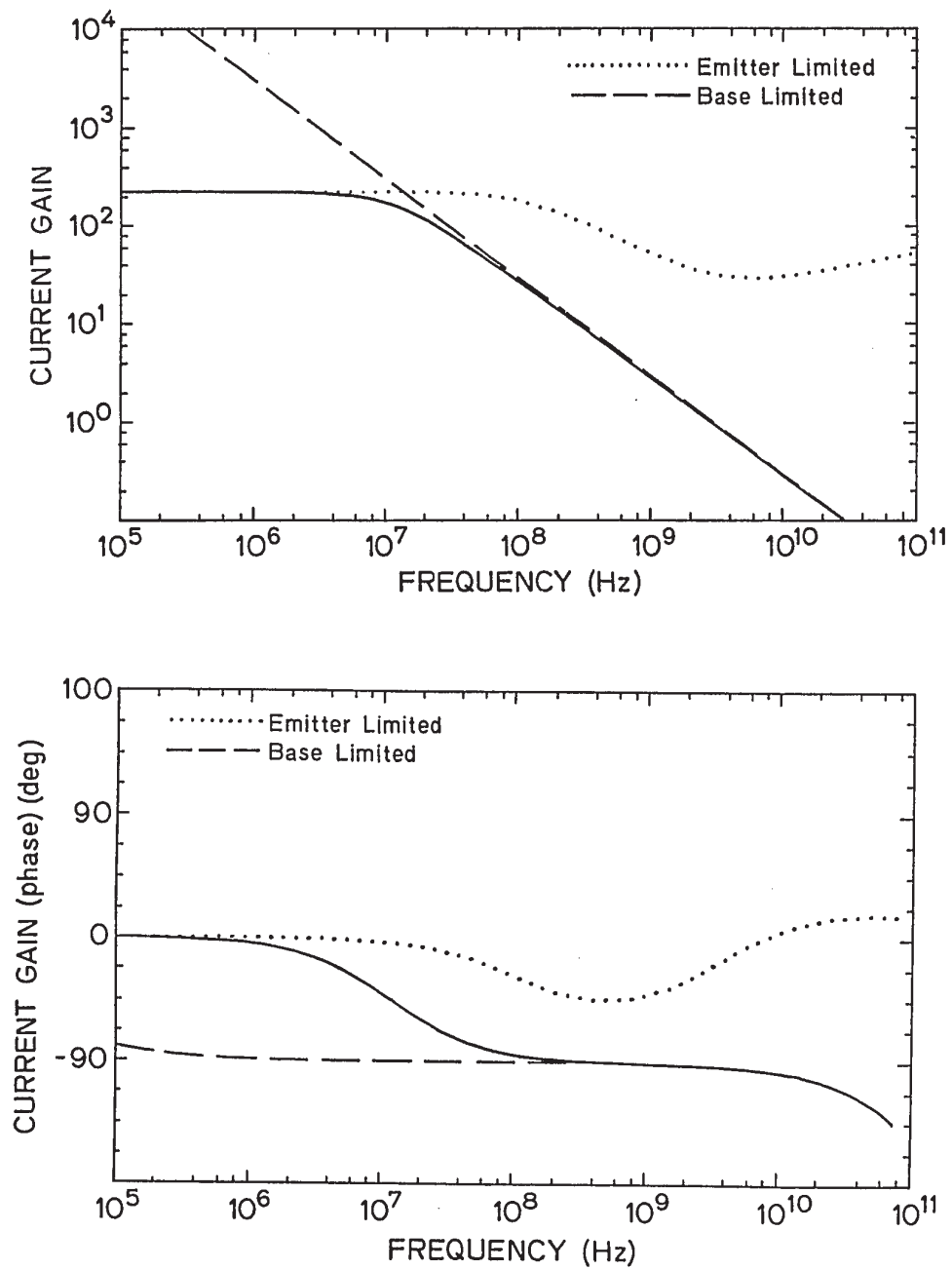


Figure 4.20: simulation of β vs. freq. for Device D: magnitude (A), phase (B)

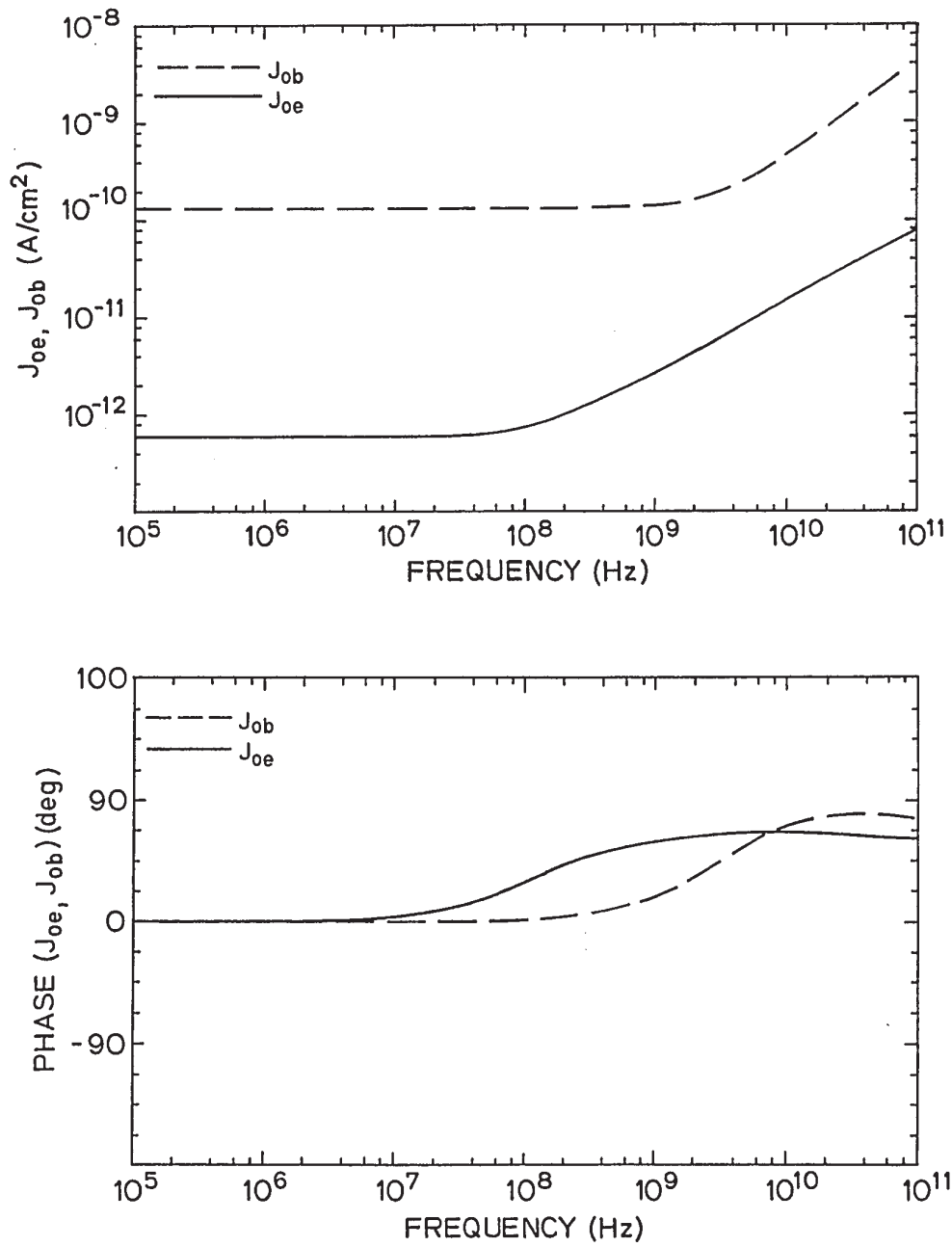


Figure 4.21: simulation of J_{oe}, J_{ob} vs. freq. for Device D: magnitude (A), phase (B)

4.14. Maximum Frequency of Oscillation

Although the common emitter current gain falls below unity for frequencies above f_T , the device is still capable of producing a useful power gain because of the difference between input and output impedances. The input impedance for this configuration is limited by r_b , the base series resistance. The real part of the output impedance is given by [7]

$$\text{Re}(Z_{out}) = \frac{1}{\omega_T C_c} \quad (4.71)$$

where

$$\omega_T = 2\pi f_T$$

and

$$C_c = C_{TC} + C_{DC}$$

is the sum of the base-collector junction and diffusion capacitances. The power gain is then given by the product of the square of the current gain and the ratio of the output to input impedance.

$$\text{Gain} = \frac{f_T}{8\pi r_b C_c} \frac{1}{f^2} \quad (4.72)$$

The frequency f_{max} at which the power gain falls to unity, also called the maximum frequency of oscillation, is given by

$$f_{max} = \sqrt{\frac{f_T}{8\pi r_b C_c}} \quad (4.73)$$

Optimization of this quantity requires inclusion of factors such as junction capacitances and depletion layers transit times which have not been included in our analysis of f_T . Yet it is possible to make some general statements regarding the design choices made available through the use of low J_{oe} structures. From equation (4.73), it is evident that potential improvements in f_{max} lie in the reduction of the base series resistance. For a given geometry, C_c is determined by the width of the lightly doped collector epi region independent of the base doping. Reducing base

series resistance requires that the base doping be increased while maintaining or even attempting to reduce the base width to keep the base transit time small. This has several effects on f_T ; emitter-base junction capacitance is increased and the mobility reduction for electrons as the base doping is increased both tend to reduce f_T ; however, the base sheet resistance falls more rapidly than f_T is compromised, so that there is a net gain in going to higher doping levels. The ultimate increase in base doping will be limited by other factors such as the reduction in base-emitter breakdown voltage. There are also potential reductions in the Johnson noise associated with the lower base resistivity, although this has not been investigated.

4.15. SUMMARY

In this section, the high frequency behavior of the BJT was investigated using charge control. Potential improvements in high frequency performance that accrue through the use of selective contacts to produce low static J_{oe} emitters were illustrated. Although this is an interesting application for these structures, the remainder of the study will be restricted to the low frequency domain since application of these contact structures to a photovoltaic cell is the principle goal of the project.

Chapter 5

SIPOS

5.1. Introduction

SIPOS, an acronym for **S**emi-**I**nsulating **P**olycrystalline **S**ilicon has become a generic name for a class of silicon oxides SiO_x , $x < 2$ prepared by CVD. The physical properties of these films such as refractive index and the electrical conductivity can be varied over a wide range by adjusting the oxygen content. By introducing phosphine or diborane into the deposition ambient, these films can be doped *N* or *P* type *in situ*; implantation of phosphorous or boron offers an alternative means for introducing these dopants. These films, pioneered at SONY, found initial application in their intrinsic form to the passivation of high voltage devices [1]. Later, doped films were successfully applied to selective contact structures, resulting in a fifty fold reduction of J_{oe} for *N*-type contacts [2,3]. This chapter will briefly survey the current state of SIPOS technology and will discuss aspects of proposed microstructural models that have relevance to determining the nature of carrier transport when these films are applied to selective contact structures.

5.2. SiO_x Films

SiO_x films prepared by vacuum evaporation or sputtering of silicon or silicon monoxide in a reactive ambient containing controlled amounts of oxygen have been studied by Phillips and others [4]. These films, whose refractive index and absorp-

tion edge can be varied to span the continuum from amorphous silicon to vitreous silica simply by varying the oxygen partial pressure in the ambient may have applications to optimized antireflection coatings. Due to the ease of control over the refractive index, for example, it may be possible to fabricate continuously graded films for wide bandwidth applications.

Phillips noted that the optical properties of these evaporated films could be explained on the basis of a random bonding model where the probability of incorporating an oxygen atom into the nearest neighbor coordination shell of a silicon atom was determined statistically (i.e. at random without consideration of energetics). The resulting probability distribution of $Si - Si_4$, $Si - Si_3O$, $Si - Si_2O_2$, $Si - SiO_3$, $Si - O_4$ tetrahedra satisfactorily modeled the spectral dependence of reflectivity from these films. The key assumption that the optical response is dominated by short range order (i.e. nearest neighbor environment) is plausible considering that crystalline and amorphous SiO_2 have remarkably similar optical spectra despite the loss of long range order in the latter. The conclusion of these studies was that oxygen seems to be homogeneously incorporated in evaporated films implying the existence of tetrahedra in which the silicon is not fully four-fold coordinated with oxygen.

5.3. CVD SIPOS

SIPOS films are prepared by chemical vapor deposition (CVD) at a much higher substrate temperature. The reaction is one between silane and an oxygen bearing species such as nitrous oxide [5] or carbon dioxide [6]. Both low pressure CVD (LPCVD) and atmospheric pressure CVD (APCVD) have been prepared. The microstructure of CVD SiO_x is quite sensitive to the deposition conditions. For example, LPCVD films with their characteristically lower deposition rates have displayed more SiO_2 character (i.e. a greater fraction of the oxygen is present in $Si - O_4$ tetrahedra) than the more rapidly deposited APCVD films of comparable

oxygen content. For the lower deposition rate, energetics may be an important factor if the reaction is closer to equilibrium. Silicon atoms already bonded to oxygen possess a slight positive net charge due to the greater electronegativity of the oxygen to which it is bonded. They can compete more effectively for incoming oxygen atoms and this tends toward complete oxygen coordination if the arrival rate is sufficiently low [7]. If there is insufficient time for lateral diffusion, then it would be expected that the oxygen incorporation would be determined at random, and the resulting distribution of oxygen atoms per tetrahedron would resemble that of the evaporated films.

Studies of as-deposited SIPOS films relying on XPS and UPS spectra indicate the presence of tetrahedra with less than four-fold oxygen coordination. There is some variance in the conclusions arrived at by several investigators [8,9] arising from disagreements in the values assigned to shifts of the silicon 2*p* peak induced by oxygen bonding. Furthermore, the role of bond strain in determining chemical shifts is not well known and it is possible that most of the shifts can be explained on the basis of structural and not compositional disorder [9]. Yet another complication arises from sputter induced damage. Since the escape depth for these techniques is very small, on the order of few monolayers, they tend to be surface sensitive, and it is likely that the surface morphology bears little resemblance to the bulk.

In applications such as high voltage device passivation, these issues may be of relevance. For selective contacts, however, these films receive subsequent high temperature steps. These annealing steps may be necessitated by the need to activate the dopant atoms, to anneal implant damage, or to form a shallow emitter by out-diffusion of dopant. In any case, the structure of these films is drastically altered by these thermal treatments, and the controversy surrounding the morphology of as-deposited films is of peripheral interest for contact applications. Unfortunately, little work has been done on doped annealed films. Hamasaki [10] in a seminal work investigated the evolution of the microstructure of undoped films as a function of heat treatment. At deposition temperatures below 650°C, no crystalline regions

could be found via TEM using Si< 111 > reflection. An upper limit for the size of the microcrystals, if present was set at 10 Å , the resolution limit of the instrument. This was supported by measurements of the peak broadening of X-ray diffraction peaks (for silicon) induced by the finite grain size. Raman backscattering [11] indicates that silicon is present in the form of amorphous silicon granules in as-deposited films. This implies that for CVD films, oxygen is not homogeneously distributed on an atomic scale, but rather that separate phases are already present.

Annealing these films results in crystallization of these amorphous silicon regions. The broad peaks in Raman spectra sharpen to those characteristic of crystalline silicon and these silicon grains become visible in TEM pictures. The transformation temperature depends upon the oxygen content of the film but is much higher than that for bulk silicon. The grain size also increases with increasing anneal temperature and appears to be a function of the ultimate temperature independent of the past history of the sample and only weakly dependent on its oxygen content. This transformation occurs in two steps. A rapid conversion of the amorphous regions into crystalline regions occurs for films annealed for a few minutes. Continued annealing results in a slow growth of the grains whose rate depends upon the oxygen content of the surrounding matrix. This is probably accompanied by an expulsion of oxygen in excess of solid solubility from the growing silicon grains. This process creates an oxygen rich shell surrounding the grain. Once a significant number of SiO_4 tetrahedra form, their inherent stability tends to retard further diffusion and grain growth ceases. The resulting structure in many ways resembles that of a cermet.

The driving forces behind the phase separation of SIPOS are the conversion of the oxygen rich SiO_x regions to SiO_2 and the crystallization of the amorphous silicon regions. This latter tendency is dependent upon the size of the region as well as its chemical environment . Raman spectra indicate a completed crystalline transformation at 1000°C for a film with 32 at% oxygen [11]. Films with 54 at% oxygen however do not show complete crystallization until temperatures of 1150°C

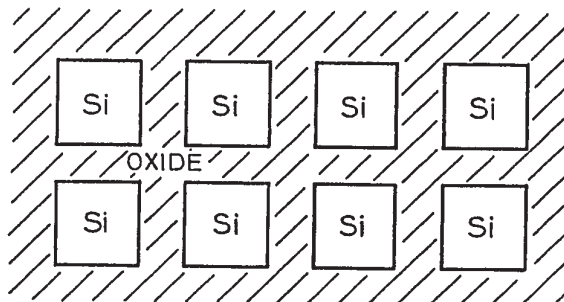


Figure 5.1: shell model of SIPOS

[12]. For the higher oxygen content films with their smaller silicon grains, interfacial energies play a more significant role and may inhibit the transformation. However, no consistent physical mechanism has been proposed for this dependence.

No comparable structural studies have been performed on doped SIPOS films. Although phosphorous incorporation in the *N*-type films is of the order of one atomic percent [5], its impact on the resulting microstructure and annealing kinetics has not been investigated. Although dopant enhanced mechanisms such as occur for doped polysilicon may be operative, the overall morphology, namely that of silicon microcrystals isolated by an insulating matrix is still expected to apply.

5.4. Device Applications for SIPOS Films

The uniform dispersion of silicon crystallites in an oxide matrix results in a large effective internal interfacial area. This coupled with the observation that the interface between the silicon microcrystal and the oxide-like matrix is likely to be of inferior quality to that achieved by thermal oxidation would indicate that the surface properties of the silicon grains are an important factor in determining the electronic properties of these films. As an example, consider the idealized model of an annealed SIPOS film, as shown in Fig. 5.1. This “shell” model, proposed by Ni [13] consists of cubical silicon grains surrounded by an oxide shell. If the average

density of interface states is $N_{ss} (cm^{-2}eV^{-1})$ then the effective volumetric density of defect states is simply $6N_{ss}/d$ where d is the size of the silicon grain and associated oxide wall. For a value of N_{ss} of $1 \times 10^{11} cm^{-2}eV^{-1}$ which is a typical value for a thermal oxide prior to low temperature hydrogen anneal, by assuming a typical grain size of $100\text{\AA}$, this translates into an effective volumetric defect density of $6 \times 10^{17} cm^{-3}eV^{-1}$. If N_{ss} is of the order of $1 \times 10^{13} cm^{-2}eV^{-1}$ which still represents only 1% of the available surface sites, the volumetric defect density increases to $6 \times 10^{19} cm^{-3}eV^{-1}$. If these interface states remain in equilibrium with the bulk of the grains so that the assumption of a flat Fermi level is justified, externally applied fields can be screened in an effective Debye length given by:

$$L_D = \sqrt{\frac{\epsilon_{SIPOS}}{q^2 N(E_f)}}$$

where ϵ_{SIPOS} is the effective static dielectric constant of the SIPOS film, and where $N(E_f)$ is the volumetric density of screening states per eV at the Fermi level. For the range of $N(E_f)$ considered above, L_D can range from $10\text{\AA}$ to $600\text{\AA}$. Even for the lower interfacial densities, the screening length can be quite small. Therefore, a thin layer of SIPOS can serve to effectively screen external fields assuming of course that the interfacial states remain in equilibrium. For static conditions, this condition is satisfied, but for time varying fields, charge must be transported into and out of these states. This flow which constitutes a current must be transported from grain to grain. Since the grains are isolated with oxide-like insulating shells, the current may be limited by mechanisms such as tunneling. These barriers to current flow present an effective series resistance and the dynamic response of these films may be limited by this effective RC time constant.

To ascertain the validity of this cermet-like structure, Ni [13] measured the conductivity of SIPOS films as a function of their oxygen content. The crystallite size was fixed at $45\text{\AA}$ typical of SIPOS annealed at $900^\circ C$, and was taken as constant independent of the oxygen content. This simplification fails at either extreme of oxygen content but the weak dependence of crystallite size on oxygen content for

intermediate values [10] lends this assumption some credibility. By assuming that the intergrain material was SiO_2 , and by using bulk values for the mass density of silicon and silicon dioxide, the thickness of the oxide shell can be computed from the oxygen content. Using the standard WKB formalism for the tunneling current, a good fit is found for oxygen contents above 20 at%. For lower values, the assumption of a constant grain size and the accompanying prediction of oxide shell thicknesses less than 4 Å is questionable. The structural integrity of such thin shells is doubtful and the effective barrier height is probably substantially reduced from that of bulk SiO_2 [14]. For higher oxygen contents, the exponential increase in resistivity with oxygen content as predicted by the model is confirmed.

The field screening properties of SIPOS and its adjustable resistivity make it suitable for several applications where screening of external fields is desirable. The effective threshold voltage of an MOS device with an overlying SIPOS layer can far exceed any anticipated circuit potentials even with a thin capping oxide. This obviates the need for a thick field oxide to accomplish this function. A planarized CMOS technology has been proposed [15] where SIPOS is selectively oxidized to SiO_2 in the gate regions. The SIPOS screening action in the field regions makes a channel-stopperless technology possible with attendant reductions in chip area.

In power devices lightly doped regions which are necessary to sustain the high voltages are susceptible to surface band bending by circuit potentials and by trapped oxide charge. Because of its finite resistivity, charges injected into the SIPOS layer are bled off. In addition to the field screening action, there is a lateral ohmic current that flows due to device potentials as illustrated in Fig. 5.2. This ohmic flow serves to relax the surface potentials thereby increasing the junction breakdown voltage. In this application, where the SIPOS acts as a “leaky” field plate, there is a tradeoff between the extent of field relaxation and the resulting leakage currents. Lower oxygen content films are more effective due to their higher conductivity, yet incur the penalty of increased leakage. This tradeoff has been quantified by Matsushita [1] who found the optimal value to be approximately

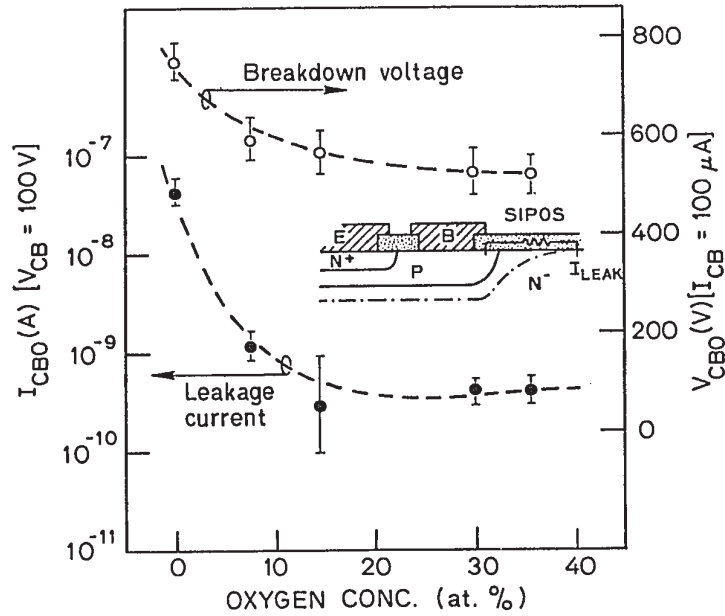


Figure 5.2: SIPOS passivation of high voltage devices, from [1]

15 at%. Interestingly, the component of leakage due to generation in the depleted region near the *Si*-SIPOS interface is much lower for SIPOS than for undoped polysilicon implying that a lower density of surface states for the SIPOS-*Si* interface as compared to the polysilicon-*Si* interface. The significance of this observation will become more apparent when the application to selective contacts is discussed.

5.5. Application to Selective Contacts

SIPOS can be doped by adding phosphine to the deposition ambient for *N*-type films and diborane for *P*-type films. The electronic properties of as-deposited *N*-type SIPOS show little difference from their undoped counterparts [16]. Presumably, the as-deposited film is highly defective; dopant atoms may not be incorporated into tetrahedral sites or they may be compensated by defect states. Annealing these films drastically lowers their resistivity. Film sheet resistivity as a function of

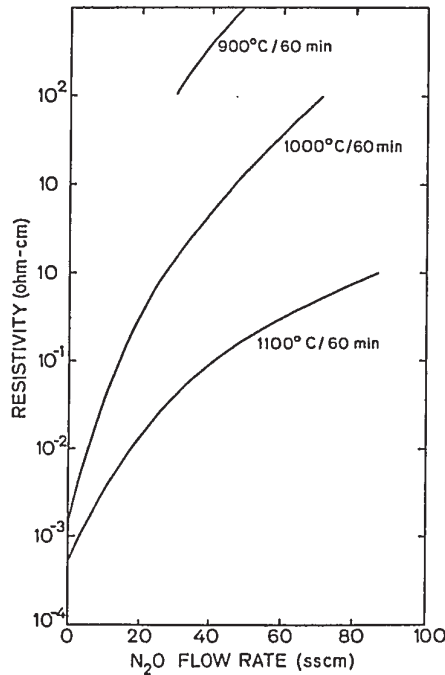


Figure 5.3: SIPOS resistivity vs. N_2O flow rate and anneal temp., from [5]

various process parameters is shown in Fig. 5.3 [5]. Crystallization of the silicon regions activates the impurities and the accompanying sharpening of the interfacial region may also serve to reduce the effective transport barrier. For these heavily doped films ($\approx 1 \text{ at\% P}$) the silicon grains can be assumed to be degenerately doped. Hence a true cermet structure may be achieved in the sense that small metallic particles are suspended in an insulating matrix. For such a situation, a percolation model seems reasonable for majority carrier transport [17]. Ni argued against such a model because a percolation threshold was not observed as the oxygen content was reduced to zero in the films that he studied. However, in those undoped films with their readily depleted grains, the transport is likely limited by the presence of induced barriers. In such a case, a sharp threshold would not be expected to occur. For heavily doped films, induced barriers are difficult to sustain

unless very large values for the interfacial state density are assumed. Therefore, the picture of metallic grains surrounded by an insulating shell seems reasonable. For majority carriers, the transport mechanisms, although not worked out in detail, seem straightforward. The principle parameters of interest are the effective resistivity and for dynamic applications also the dielectric relaxation time.

5.6. Minority Carriers

For minority carriers, the situation is very complicated and until a definitive understanding of the microstructure and the recombination efficiency of grain surface states is obtained, modeling can at best only indicate which mechanisms are most likely. By using limiting values for various parameters of interest, it should be possible to obtain a semiquantitative understanding.

At first glance, SIPOS with its large internal effective surface area would seem an unlikely candidate for a selective contact. Yet experimental contacts to silicon indicate a fifty fold reduction in J_{oe} , the saturation current [2,3]. Matsushita invoked a heterojunction contact model to explain this reduction. As discussed in Chapter 3, the effective surface recombination velocity looking toward the heterojunction contact is given by (3.41)

$$S_c = S_{int} + \frac{n_{ie,X}^2}{n_{io}^2} \frac{N_{surf,eff}}{N_{x,eff}} \sqrt{\frac{D_X}{L_X}}$$

Although D_X and L_X for SIPOS are unknown, these quantities can be estimated if one constructs a model structure and assumes values for the capture cross-section, density, etc. Matsushita [3] estimated a value for the diffusion length of 300 Å. Therefore, in order to obtain an effective contact recombination velocity of 100 cm/sec, a valence band edge discontinuity ΔE_v of 0.2 eV is required, assuming a unity value for D_{eff} and assuming that S_{int} is negligible. This is in fact the value for the barrier to reverse hole injection estimated by Matsushita from measurements of the diffusion potential of a SIPOS-Si *n-p* junction. The significance of the observation that the Si-SIPOS interface appears relatively free of midgap

states that was made in the context of passivation of high voltage devices is now apparent. The drawback of previous attempts to fabricate heterojunction contacts to silicon was the high value for S_{int} associated with active defect states arising from lattice mismatch; presumably these prove difficult to passivate. Although SIPOS possesses a disordered structure, interfacial accommodation is apparently achieved since the deduced value for S_c are of the order of $100\text{ cm}^2/\text{sec}$.

The application of a HJC model to SIPOS contacts seems dubious in view of the microstructural studies which reveal a multiphase structure. Yet a CFO (Cohen, Fritzsche, Ovshinsky) density of states, which postulates the presence of a "U"-shaped continuum of localized states in the bandgap of disordered materials, was invoked by Hamasaki [18] to explain the dependence of field effect conductance on gate voltage and temperature. The apparent activation energy for conductivity was taken to be a measure of the mobility gap in the CFO model. The key point is the apparent dependence of "bandgap" on oxygen content. Theoretical calculations for the density of states for SiO_x based on cluster calculations [19] do in fact reveal an expanding gap as the oxygen is increased. As more oxygen atoms are added between formerly bonded silicon atoms, the $Si-Si$ bonding states which formed the silicon valence band are gradually replaced by $Si-O$ bonds. The silicon valence band is gradually attenuated and the apparent bandgap increases very slowly which can be rationalized on the basis of decreasing coordination to other silicon atoms. As x approaches 1.5 the gap increases rapidly, ultimately to reach the value of SiO_2 as $x \rightarrow 2$. Care must be taken before applying these simplified calculations to SIPOS. The small clusters used in the calculations did not weight various configurations based on energetics. That is to say, the lowest energy bonding configurations were not preferentially represented. In this sense, it is a random bonding model. No bond distortions were considered (such model structures as assumed for the calculations could not be realized in real space without bond distortion) which would likely give rise to gap states. In any case, for SIPOS films of technological interest, phase separation is evident and oxygen is segregated to the intergrain regions. It is here

that the random bonding model which assumes that oxygen is uniformly distributed on an atomic basis loses its validity.

The issue of which model to use is of more than passing interest. If oxygen could be uniformly incorporated, there arises the possibility that the effective gap could be tailored by adjusting the oxygen content. Since this is accompanied by an increase in film resistivity there would appear to be a tradeoff between attaining large bandgaps by using large oxygen contents and the penalty that must be paid in terms of series resistance. Oxygen content would be a key process parameter in such a scenario since it would be the primary determinant of J_{oe} reduction. This is the rationale behind the use of SIPOS with 45 at% oxygen [2,3], the implication being that such high oxygen contents were necessary to produce the large bandgap that was responsible for the hole barrier. Since the resistivity of such a film is of the order of 10 $\Omega\text{-cm}$, it rules out applications such as concentrator solar cells and high speed bipolar devices where the excessive series resistance would negate the benefits that would otherwise accrue from the use of SIPOS contacts. If on the other hand, the SIS model pertained, then the degree of J_{oe} reduction would be less sensitive to the bulk properties of the SIPOS itself, and effort would best be directed at optimizing the interfacial oxide layer.

The current state of knowledge in regard to the microstructure of SIPOS films is clearly inadequate to permit first principles modeling of minority carrier transport in these films. Yet such an understanding will be necessary if these films are to be optimized for applications such as high frequency devices where their dynamic characteristics will be dictated by the nature of the transport. For example if suppression of hole transport is via self-trapping of holes, the dynamic performance will be different from the case where holes are repelled from the bulk of the SIPOS by a tunnelable interfacial oxide. Another issue that can be experimentally tested is the dependence of J_{oe} on the oxygen content of the sample. If it is insensitive to bulk SIPOS properties, then it seems unlikely that the heterojunction model is applicable; in that case, another model such as the SIS structure may possess more

relevance.

5.7. SIPOS Deposition

Published experimental results had already proven that SIPOS films were capable of blocking hole transport when deposited onto $N+$ regions. The sensitivity of the anticipated applications to series resistance dictated that primary consideration be given to quantifying the tradeoff between J_{oe} reduction and film resistivity i.e. to find the minimum amount of oxygen in the film necessary to achieve suppression of hole transport. Other key process parameters as pointed out in Chapter 3 would involve surface (interfacial) treatments prior to SIPOS depositions and the thermal annealing schedule necessary to tailor the dopant profile outdiffused into the monocrystalline region.

Before devices could be fabricated to evaluate J_{oe} reduction as a function of these variables, a deposition technology for SIPOS films had to be developed. The desire for *in situ* doping and the flexibility required to accommodate a wide range of film compositions dictated the development of an in-house capability. The silane-nitrous oxide system, for which the bulk of the literature is available, was chosen primarily to duplicate the conditions reported by Matsushita [3].

A horizontal cold wall RF heated APCVD reactor was used for the initial test runs. The resulting film uniformity remained poor despite attempts to optimize carrier flow rates, susceptor position, and deposition temperature. In retrospect, the small susceptor size ($7\text{ cm} \times 23\text{ cm}$) and the sensitivity of the composition of this non-stoichiometric film to reaction kinetics ruled out the possibility of attaining uniform deposition in this reactor. Seaward [20] reported that for SIPOS deposited in a hot wall CVD reactor the oxygen content of the films increased significantly with downstream placement of the wafers. Although no mechanism for this observation was proposed, it seems likely that the thermal gradients along the tube introduced by the large carrier flow rates are responsible for this compositional

variation. Possible mechanisms include thermally activated dissociation of nitrous oxide in the gas stream or a temperature dependence of the oxygen incorporating surface reaction. Unfortunately, no literature pertaining to the kinetics of this reaction has been found. However, since it was felt that temperature profile control was the key determinant of compositional uniformity, deposition was transferred to a hot wall system since it offered better potential for control of this parameter.

Ideally, investigation of the role of oxygen in determining SIPOS transport properties would call for the use of an LPCVD system. If the shell model for SIPOS structure is correct, then the most interesting physics would lie in the low oxygen content regime where the shells are just forming and marking the transition from polycrystalline silicon to SIPOS. Control of residual components in the deposition ambient in LPCVD is far superior to that of APCVD since the partial pressure of the reactive gas species to active contaminants such as water vapor and oxygen is much larger in the former. Additionally, control of the temperature profile in APCVD is lost if the carrier gas stream velocity becomes excessive. This is not a drawback with LPCVD. Despite these compelling advantages, due to the lack of an available LPCVD facility, a conventional diffusion tube was modified to serve as an APCVD reactor. A circular cross-section although nonoptimal from the viewpoint of maximizing flow velocity for a given flow rate, was used to simplify construction. The first generation configuration is illustrated in Fig. 5.4; typical flow rates were nitrogen 100 *slm*, silane 30 *sccm*, and varying nitrous oxide flows up to 1 *slm* (for SiO_2).

The use of the large carrier flow rate arose from several considerations, not the least of which was a desire to maintain the deposition ambients reported in the literature [5]. The maximum partial pressure of silane is limited by the onset of spontaneous gas phase nucleation. A conservative value of 500 *ppm* was used to avoid this problem even at the highest anticipated deposition temperatures. Reactant depletion sets a lower limit to the absolute flux of reactive species. These two factors dictate the use of a large carrier flow rate; a value of 100 *slm* was chosen.

To promote vigorous mixing and uniform dilution of the reactive gases by the carrier flow a forward pointing injection nozzle/deflector assembly was utilized. In a subsequent refinement, a wad of quartz wool was placed in the front end zone to act as a large surface area preheater.

The results obtained from this geometry, while an improvement over the cold wall reactor, were still disappointing. Uniformity over a 1 cm^2 region which was deemed necessary for solar cell fabrication could not be achieved. Furthermore, the films displayed a speckled appearance with an unacceptably high defect density apparently caused by particulate formation and shedding in the upstream region of the tube. This was believed due to the formation of stagnant regions arising from flow separation at the main nitrogen inlet induced by the sharp taper and high flow velocities. In the later version, stagnant regions very likely formed in the quartz wool preheater assembly. Another possible contributing factor is the simultaneous exposure of silane and nitrous oxide to high temperatures prior to injection into the main tube.

Temperature profiles taken from the leading to trailing edge of the wafer boat as a function of several flow rates are shown in Fig. 5.4 where x is to scale with respect to the drawing. For small flow rates (5 slm) the profile is relatively flat, but at 100 slm , which is more typical of actual deposition conditions, the sag is excessive. This is not surprising in view of the inefficient heating geometry and short gas residence time.

To rectify the primary deficiencies of temperature profile sag and particulate generation, several modifications were made to the initial configuration. The modified geometry, depicted in Figs. 5.5-6 reveals the addition of an external preheater for the nitrogen carrier gas and a more elaborate mixing chamber designed to minimize stagnant region formation. The twin channel preheater is essentially a pair of quartz tubes placed on the exterior of the main tube in such a manner that the entire assembly still fits inside the furnace element. The bulk of the nitrogen carrier (100 slm) flows through the preheater before injection into the mixing chamber;

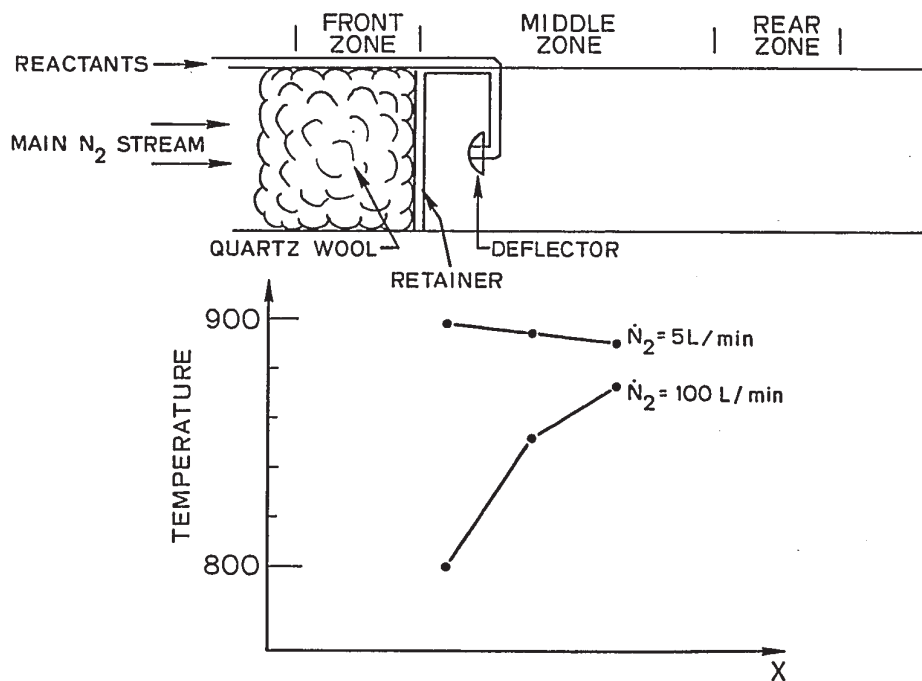


Figure 5.4: first generation SIPOS deposition apparatus

heat transfer calculations are presented in Appendix F. In initial configurations the preheater assembly was tacked down to the main tube but the thermal cycling induced by changing nitrogen flow rates resulted in the development of cracks. Supporting the preheater assembly only at the mixing baffle and allowing contraction of the tube by use of a slider arrangement as shown in Fig. 5.5 rectified this problem.

Nitrous oxide is introduced into the reactive gas line which carries the silane and phosphine in a dilutant flow of nitrogen, just prior to injection into the mixing chamber. As can be seen from Fig. 5.6, the mixing chamber is a set of concentric cylinders forming annular regions for the three separate gas flows. The innermost tube injects the reactive gas species diluted by a nitrogen carrier at room temperature. Its diameter is chosen so that the Reynolds number for the chosen flow rate lies in the turbulent regime. This turbulent jet that is formed is surrounded

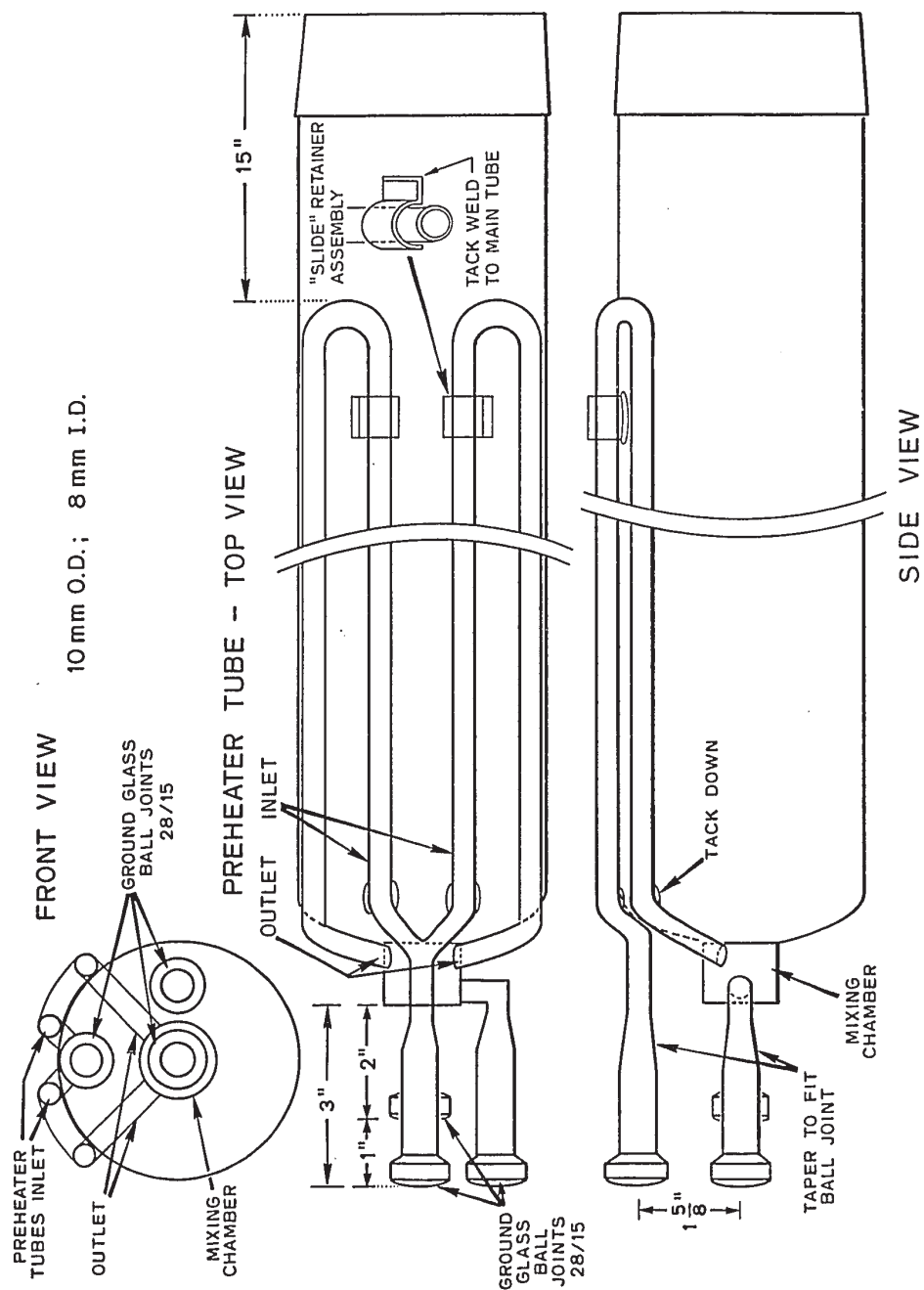


Figure 5.5: SIPOS deposition tube — overall view

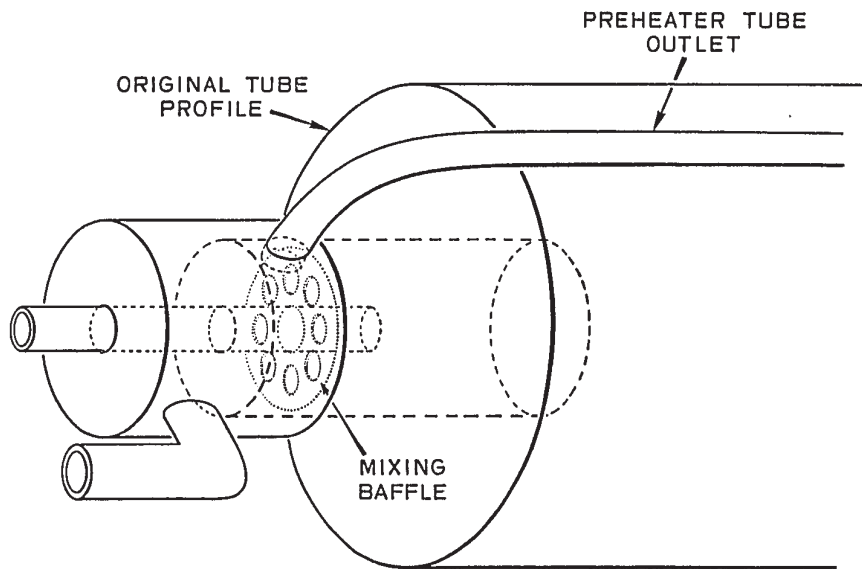


Figure 5.6: SIPOS deposition tube — closeup of mixing chamber

by the preheated nitrogen stream which constitutes the bulk of the total gas flow. The area of the annular region in which the preheated nitrogen flows is chosen to maintain a velocity ratio of .3 : 1 with respect to the reactive gas jet. This velocity differential coupled with the turbulent flow promotes thorough mixing without generating stagnant regions. The outermost annulus carries a purge flow of nitrogen; this serves to keep the front end of the tube free from eddies essentially by acting as the fluid equivalent of a shallow angle diffuser. The gas stream which at this point has been mixed and preheated then passes through a flow smoothing baffle situated in the front end zone. This close packed hexagonal array of small diameter quartz tubes helps to heat the gas stream further by virtue of its large surface area and additionally serves to minimize any angular components in the gas velocity. There is some reactant depletion due to the presence of the baffle; however, these deposits are adherent and do not result in particulate generation.

The wafer boat itself rides on a pair of sled-like rails. Since this is a hot wall CVD process, tube deposits are unavoidable. The rails serve not only to locate the horizontally placed wafers symmetrically in the deposition chamber, but also to elevate them above the particulate rich region near the bottom of the tube that results from loading and unloading. A cantilever loading system was considered, but found impractical due to other considerations.

The temperature profile for the new geometry is significantly improved exhibiting a temperature sag of 10°C over the 12" length of the wafer boat at full nitrogen flow rates of 120 *slm*. Horizontal placement of the wafers on a quartz boat with a large thermal mass results in uniform deposition. It is possible to obtain better than 10% thickness uniformity over a three inch wafer; useful deposition extends over a 6" region. Typical flow rates used are : preheater nitrogen 105 *slm*, dilutant nitrogen (carries reactive gases) 12 *slm*, tube front end purge nitrogen 6 *slm*, silane 50 *sccm*, phosphine 4 *sccm*, varying nitrous oxide up to 80 *sccm*. A plot of film resistivity as determined by four point probe is shown in Fig. 5.7 as a function of γ , the ratio of $\text{N}_2\text{O}/\text{SiH}_4$ in the deposition ambient, and annealing temperature. The phosphine partial pressure was chosen to be sufficiently large that further increases did not significantly reduce the film resistivity. From the plot, the $\gamma=0$ film behaves quite similar to POLY in that high temperature anneals lower the film resistivity to of the order of $1 \times 10^{-3} \text{ohm-cm}$, but above 900°C , dopant activation is essentially complete, and further grain growth does not materially affect the resistivity for these heavily doped films. The presence of oxygen is seen to drastically increase the film resistivity, as well as impeding the process of reconstruction. Significant changes are still occurring for the $\gamma \neq 0$ films as evidenced by the continuing reduction of film resistivity for $T_A > 1000^{\circ}\text{C}$. This trend reflects the presence of oxygen in the film which tends to inhibit film reconstruction as previously discussed.

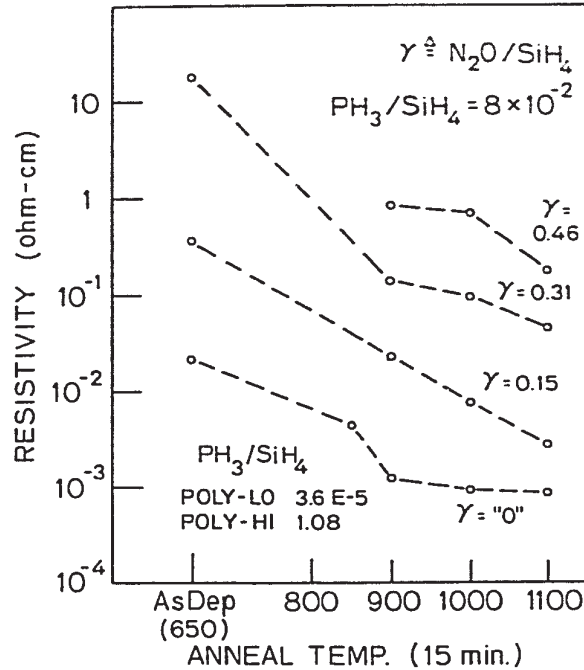


Figure 5.7: SIPOS film resistivity vs. T_A and γ

5.8. Summary

The addition of oxygen to silicon produces a class of films that have a wide variety of potential applications. For applications such as field screening where majority carrier properties are important, the cermet-like shell model may be sufficient to predict the significant film parameters. For applications to selective contact structures, minority carrier transport is of primary interest. A complete transport model requires a more extensive understanding of the microstructure of the film, and the nature of the interface that is formed when SIPOS is deposited on crystalline Si. The experiments detailed in Chapter 7 are preliminary steps in this direction.

Chapter 6

Experimental Determination of J_{oe}

6.1. Introduction

The discussion in Chapter 3 pointed out that J_{oe} , the saturation current due to minority carrier injection into the emitter region is the fundamental parameter that characterizes the effectiveness of the selective contact in blocking minority carrier transport in the low frequency limit. The task remains of determining this quantity for actual emitter structures from experimental measurements. The bipolar transistor provides a convenient test vehicle for these measurements. This Chapter will discuss the rationale behind the device structure that was chosen and will outline the procedures used to extract J_{oe} from measurements of the base current.

To extract J_{oe} , one needs only to provide a determinable concentration of excess minority carriers and to note the resulting current flow. These carriers can be provided by either electrical injection or by photo-excitation and the resulting current can be determined by direct measurement of the terminal current or indirectly, as in the case of the photoconductivity technique (see Appendix G for a discussion of the photoconductivity technique). A P - N junction can be used to inject carriers, but the resulting current flow also includes electron injection from the N -type emitter structure into the P -type region. In a diode, direct separation of the two minority carrier currents is not possible, so the contribution from electron injection into the substrate must be accounted for by other means. While this may be feasible for

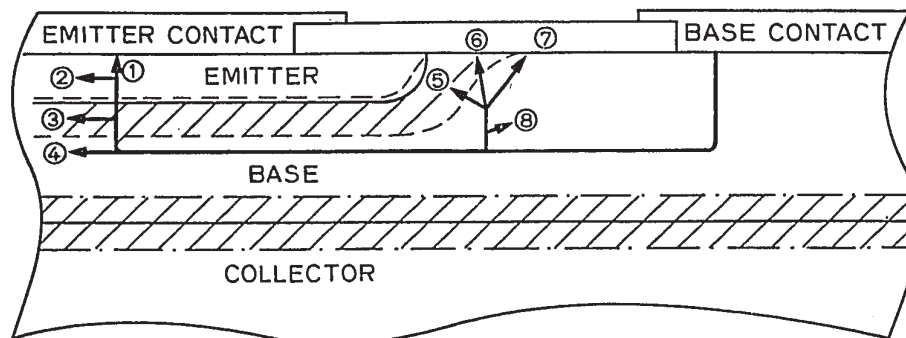


Figure 6.1: components of the base current

conventional emitters with their relatively large J_{oe} , the selective contact structures display J_{oe} values of $1 \times 10^{-14} \text{ A/cm}^2$ and below. Since this is much smaller than the J_o of the substrate region, it is not possible to subtract the substrate contribution without introducing large uncertainties in the deduced value for J_{oe} .

The use of a bipolar transistor allows separate measurement of the hole current. The presence of the reverse biased base-collector junction blocks hole transport through the collector electrode. All hole current then is restricted to flow through the base-emitter circuit. Therefore measurement of the base current is essentially a measurement of the hole current. Since hole currents must flow to supply recombination in various regions of the device in addition to the reverse injection into the emitter (the component we are interested in), these spurious components must be subtracted before a determination of J_{oe} can be made.

6.2. The Base Current

The base current in a BJT consists of several components as illustrated in Fig. 6.1. These can be broadly grouped into “area” and “perimeter” related components, depending upon whether they scale with the area or perimeter of the

emitter. The base current can be written as

$$I_b = A \cdot J_{area} + P \cdot J_{peri} \quad (6.1)$$

where A and P denote the area and perimeter of the emitter respectively. J_{area} includes J_{oe} (1,2), space charge recombination (3), and neutral base recombination (4). J_{area} also includes a small negative component due to generation in the B-C space charge layer which can usually be neglected. J_{peri} includes sidewall space charge recombination (5), perimeter surface recombination both within (6) and without (7) the depletion layer and recombination in the neutral base due to the “fringing” of current streamlines near the edge of the emitter (8); leakage current generated in the B-C depletion region is usually small enough to be neglected at room temperature. By using adjacent devices with different area/perimeter ratios we can solve the set of equations

$$\begin{aligned} I_{b,large} &= A_{large} \cdot J_{area} + P_{large} \cdot J_{peri} \\ I_{b,small} &= A_{small} \cdot J_{area} + P_{small} \cdot J_{peri} \end{aligned} \quad (6.2)$$

for J_{area} and J_{peri} . The only assumption made is that conditions, especially at the surface, are identical for both devices. This is a reasonable assumption if they are not widely separated.

Having obtained J_{area} , it remains to subtract the parasitics. These are primarily the depletion region (3) and neutral base region recombination (4). The depletion region component usually displays an ideality factor less than unity; by maintaining adherence to clean processing guidelines, especially in regard to metallic contaminants, it is possible to obtain base currents over four or more orders of magnitude for which this component is negligible, even for the lowest J_{oe} structures. Although it is possible to subtract this component using a curve fitting method [1], it was not found to be necessary for the bias range used in this study. Neutral base recombination displays the same bias dependence as J_{oe} , the desired component, so another technique, base width modulation, must be used to account for its presence.

By modulating the base-collector bias, the width of the quasi-neutral base region can be controlled. For the case of a uniformly doped base (as is the case for these test structures which use an epi layer for the base), the neutral base region recombination can be written as

$$J_{neut. base} = \frac{qn_{io}^2}{N_b} \frac{W_b}{2\tau_n} (e^{V_{bc}/V_T} - 1); \quad V_T = \frac{kT}{q} \quad (6.3)$$

where τ_n is the base region minority carrier lifetime and W_b is the neutral base width. If the widths corresponding to base-collector biases V_1 and V_2 are given by W_{b1} and W_{b2} respectively, then the accompanying change in base current is

$$\begin{aligned} \Delta J_{neut. base} &= \frac{qn_{io}^2}{2\tau_n N_b} (W_{b1} - W_{b2}) (e^{V_{bc}/V_T} - 1) \\ &= \frac{qn_{io}^2}{2\tau_n N_b} (\Delta W_b) (e^{V_{bc}/V_T} - 1) \end{aligned} \quad (6.4)$$

In charge control notation where Q_b is the integrated hole charge in the neutral base, this becomes

$$\Delta J_{neut. base} = \frac{\Delta Q_b}{\tau_n} \quad (6.5)$$

Since the collector current is given by

$$J_c = \frac{qn_{io}^2 D_{n,b}}{N_b J_c} (e^{V_{bc}/V_T} - 1) \quad (6.6)$$

The analog of eqn. (6.5) in charge control notation becomes

$$\Delta J_c = \frac{\Delta Q_b 2D_{n,b}}{W_b^2} \quad (6.7)$$

or recalling the definition for the base transit time

$$\tau_f = \frac{W_b^2}{2D_{n,b}}$$

we get

$$\Delta J_c = \frac{\Delta Q_b}{\tau_f} \quad (6.8)$$

Combining eqns (6.5,8) and solving for τ_n gives

$$\tau_n = \frac{\Delta J_e}{\Delta J_{neut. base}} \tau_f \quad (6.9)$$

$J_{neut. base}$ must be related to ΔI_b the measured change in the base terminal current. Since changing the base-collector bias also changes the B-C (base-collector) junction and other parasitic leakages, care must be taken to ensure that these components are either negligible or accounted for. Choosing V_{be} sufficiently large yet not so large that either current crowding or high level injection occurs will minimize the sensitivity to this leakage current. It is also important to ensure that base punchthrough is not approached. In the near-punchthrough region, it is possible for more lightly doped regions to punchthrough first effectively isolating regions of the base. In addition, collector induced barrier lowering causes the collector current to rise much faster than the simple inverse dependence on base width predicted by this simple model. The other components of I_b should remain constant since V_{be} is kept fixed during the measurement. Therefore

$$\Delta J_{neut. base} \leq \frac{\Delta I_b}{A}$$

where A is the active injecting area (see Appendix H). Therefore, a lower bound on the base lifetime can be established by attributing all of the change in ΔI_b to neutral base recombination. The equivalent saturation current due to this mechanism then establishes a lower limit for J_{oe} below which this mechanism must be accounted for if an overestimation of J_{oe} is to be avoided. The emitter J_{oe} that would produce an equivalent amount of recombination is given by

$$J_{oe} = \frac{qn_{io}^2 W_b}{2N_b \tau_n} \quad (6.10)$$

This technique as applied to a SIPOS contacted device with a 1.5 micron neutral base width yields the curves shown in Fig. 6.2. The base-emitter bias is kept fixed at 0.64 volts. Using equation (6.9) and the values from Fig. 6.2

$$\Delta I_b = 4.3 \text{ nA}, \Delta I_e = .076 \text{ mA}$$

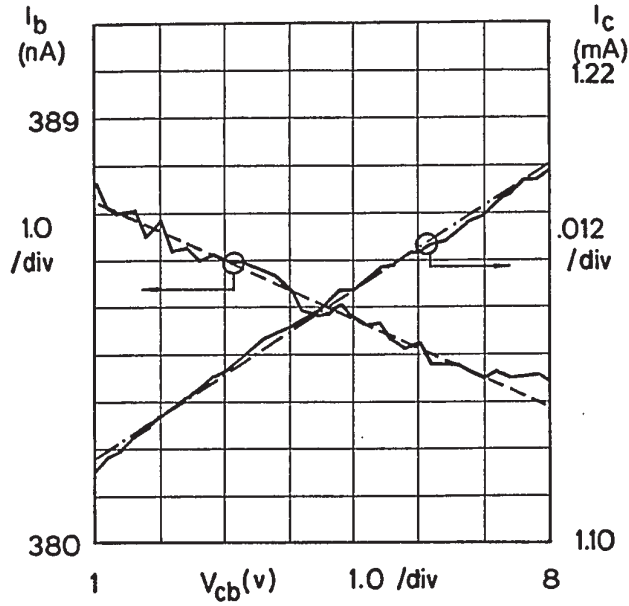


Figure 6.2: determination of τ_p in the base from BWM

for points measured at $V_{cb1} = 1 \text{ volt}$ $V_{cb2} = 8 \text{ volt}$ yield a value of $19 \mu\text{sec}$ for the neutral base lifetime which implies a saturation current contribution to the base current of $6 \times 10^{-15} \text{ A/cm}^2$. For a J_{oe} of $1 \times 10^{-14} \text{ A/cm}^2$, this represents a significant factor; however, determination of the base lifetime using this technique suffers from several uncertainties, mostly due to sensitivity to a myriad of parasitics that can not be adequately characterized or controlled at these current levels. Therefore, it is felt that J_{oe} below $2 \times 10^{-14} \text{ A/cm}^2$ can not be accurately determined using this technique with the available test structure; the deduced values from base current then serve as upper bounds on the true J_{oe} . To obtain greater precision requires the use of more sophisticated structures and/or techniques. This is not a problem for most of the devices studied which possess J_{oe} over $5 \times 10^{-14} \text{ A/cm}^2$; only for the lowest J_{oe} does this become a limitation.

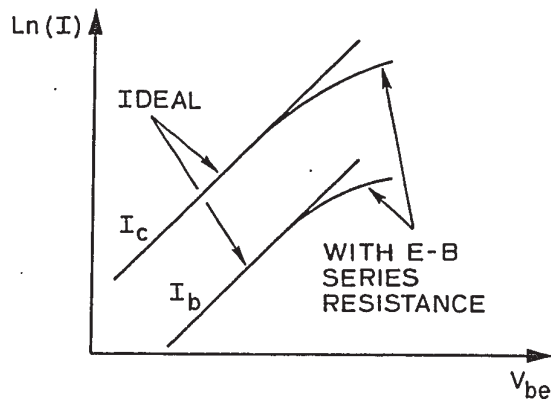


Figure 6.3: effect of series resistance on BJT characteristics

6.3. Series Resistance

The useful range of measurements can be extended by another decade in some cases if series resistance in the base emitter circuit can be accounted for. As can be seen from Fig. 6.3, the effect of base and emitter contact resistance and series resistance attributable to the extrinsic base region is to displace the the base and collector characteristics simultaneously. As the intrinsic beta is not affected, it is possible to extrapolate the collector characteristic from the low current regime, where the resistance effect is negligible, to higher currents. The voltage deviation of the actual collector characteristic from this ideal behavior can then be used to “correct” the base characteristic. An example of this procedure is presented in the next chapter. Collector series resistance can result in turn-on of the collector base junction. This results in a rapid rise in the base current and a dramatic falloff in beta. This effect can be postponed by application of a base-collector bias sufficient to keep the base-collector junction reverse biased. Base spreading resistance can lead to current crowding at sufficiently high base currents. Since injection becomes nonuniform once current crowding sets in, it is not possible to assign a proper “voltage offset” with which to correct the base characteristic. It is best to assure that this is not a limitation for the current range of interest by minimizing the base

sheet resistance consistent with other constraints such as minority carrier lifetime. A quasi-two-dimensional modeling of this effect is presented in Appendix I.

6.4. Temperature Measurement

Once J_{oe} has been extracted from the area related component of the base current, it remains to normalize it to $300^\circ K$. This is a standard procedure that permits comparison of J_{oe} (or equivalently G_e) values amongst various data sets. Due to the strong dependence of J_{oe} on temperature, care must be taken to ensure that the correct device temperature is used.

A common procedure is to deduce the device temperature from the slope of the collector characteristic from a Gummel plot. This is not always a reliable technique due to base width modulation induced by changing base-emitter bias. At larger biases, the neutral base width increases due to the reduction of the base-emitter depletion layer. This results in a smaller collector slope which can be misinterpreted as a higher device temperature. An illustration of this effect is presented in Appendix J. The errors involved can seem quite small; however, a change in slope of only 1% translates into an effective error of three to four degrees in the device temperature which can lead to an error of almost 2 in J_{oe} .

To avoid potential ambiguities, a copper-constantin thermocouple was mounted on the rear of a silicon wafer whose frontside metallization matched that of the device wafer. This assembly was in turn mounted on a teflon base which provided electrical insulation from the measurement chuck. This was deemed necessary because the chuck heater was injecting noise even when deenergized especially for measurements in the $1 \times 10^{-10} A$ range. All measurements were made in an electrostatically shielded light-tight box with all intrabox accessories such as the chuck and illumination lamp shorted to the enclosure ground.

6.5. Measurement Procedure

To avoid junction heating, most measurements were made with either 0 or 1 *volt* base-collector bias. The total current is limited to 10 *mA* by instrument limitations. This limits the peak dissipation during the last few seconds of the measurement cycle to less than 20 *mW*. Since the overall measurement duty cycle is less than 10% the temperature rise should be negligible. This is supported by actual device measurements where any sustained increase in the wafer temperature would have manifested itself as a change in the collector temperature (the temperature deduced from the slope of the temperature characteristic). The variance from device to device was random and of the order of 0.2° *K* indicating that no significant wafer heating was taking place. A device characteristic taken by ramping the base-emitter bias up then down displays no hysteresis. This indicates that all measurement points are taken under essentially quasi-static conditions and that effects such as device heating and trapping are either not significant or have time constants much shorter than the dwell time for each measurement point (i.e. less than a second).

The temperature as determined by the thermocouple was then used to “correct” the measured J_{oe} to $J_{oe}(300^\circ K)$ by using the formula

$$J_{oe}(300^\circ K) = J_{oe}(T_{device}) \cdot \frac{n_{io}^2(300^\circ K)}{n_{io}^2(T_{device})}$$

This neglects temperature variations in $D_{p,e}$, since these are small for a highly doped emitter and are dominated by the exponentially activated temperature dependence of the n_{io}^2 term. Use of the n_{io}^2 term implies that BGN effects were neglected, but since

$$|T_{meas} - 300^\circ K| < 8^\circ K$$

for all measurements, even a BGN value of 100 *mV* would only alter the J_{oe} by approximately 10%. The temperature dependence of beta measured for these devices indicated that in fact BGN is probably less than this. The value of J_{oe} corrected to 300° *K* can then serve as a value to use in device simulation codes, and as a standard

for comparison between various emitter structures. An equivalent figure of merit, the emitter Gummel number as introduced in Chapter 2 is defined by

$$G_e = \frac{qn_{io}^2(300^\circ K)}{J_{oe}(300^\circ K)} = \frac{33.6}{J_{oe}}$$

These quantities are preferable to use since they characterize the emitter region. Although comparison between beta of different devices is often alluded to when attempting to evaluate the relative merits of different emitter structures, it reflects the ratio G_e/G_b and variances in G_b associated with the differing structures makes a comparison based on beta alone of limited value.

6.6. Design Considerations for the Test Device

At the inception of this project, the emitter profiles that would yield the lowest J_{oe} were not known apriori. The structure had to be very flexible in terms of accomodating emitters with varying junction depths. Since determination of base lifetime was facilitated by use of a constant doping profile, an epi layer was used for the base region. Use of the epi structure also avoided the necessity of changing the base schedule every time the emitter structure was modified. The choice of epi thickness and doping reflect the tradeoffs between the desire to maintain base transparency which calls for thin lightly doped layers and the need to accomodate deeper emitter profiles and to minimize the base spreading resistance in these large area devices. A value of $2 - 3 \times 10^{16} \text{ cm}^{-3}$ and an epi thickness of 2.5μ were chosen as an optimal target, one which allows significant base width modulation, is reasonably transparent, and does not suffer from excessive current crowding effects. First order analysis indicates that these high current effects should not become significant until emitter- base biases of almost 0.7 volts are approached which is sufficient for the purposes of this study. High level injection in the base and base pushout (the Kirk effect [2]) are also negligible for the low current densities used in these measurements ($< 200 \text{ A/cm}^2$).

Processing constraints required that wafers be cleaved for process splits since

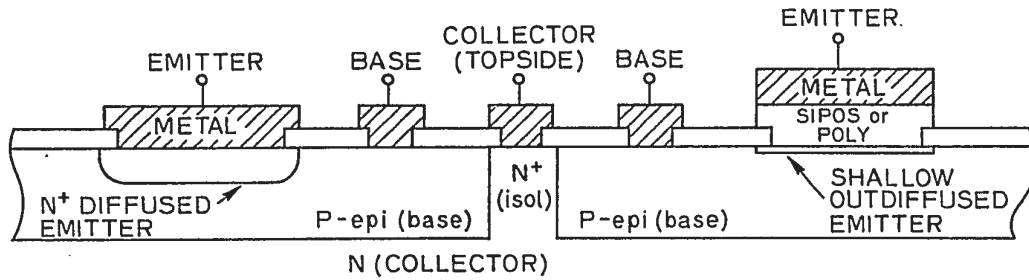


Figure 6.4: cross-section of BJT test structure

sawing introduced unacceptable contamination. This required the use of $< 100 >$ substrates which restricted the availability of wafers with the desired doping concentrations. Although the use of a buried layer was precluded by the thin lightly doped epi layer, collector series resistance was not a serious limitation despite the $1\text{ ohm} - \text{cm}$ substrate resistivity, and a one volt reverse bias on the base-collector junction proved an adequate countermeasure. Cross-sections for devices with a pre-diffused and an outdiffused emitter are shown in Fig. 6.4.

6.7. Summary

A BJT test structure was designed and a measurement procedure devised to evaluate the J_{oe} of various emitter+contact structures from measurements of the base terminal current. Precautions were taken to independently estimate device temperature and to ensure that the measurement regime was restricted to the region where the simplified one-dimensional models used in the analysis apply. The resolution of the technique is limited by uncertainties in the technique used to determine the neutral base recombination lifetime. Although this precluded accurate determination of J_{oe} below $1 \times 10^{14} \text{ A/cm}^2$, the dynamic range proved adequate for the purposes of this study.

Chapter 7

Experimental Values of J_{oe} for POLY and SIPOS Emitters

7.1. Introduction

Chapter 3 described in principle how emitters with low static J_{oe} values could be attained. It remains the task of this chapter to demonstrate that this can be achieved in practice. Using the procedures outlined in Chapter 6, J_{oe} for emitter structures using a (Si) homojunction and (SIPOS/Si, POLY/Si) heterojunction and SIS structures is determined from measurements of the base current as a function of several key processing parameters. The models developed in Chapter 3 are then utilized to explain these dependencies. This chapter concludes with a discussion of the relative merits of each structure.

An overview of the matrix of device structures used is presented in Fig. 7.1. Proceeding from the leftmost column, we can see that there are three different profiles in the monocrystalline region are present on a given die. Emitters A and B are standard diffused emitters approximately Gaussian in shape. Their profiles and process parameters are presented in the next section. Emitter C has no prediffused region, rather, it relies on the outdiffusion of impurities from the overlying film which can be either POLY or SIPOS. For each emitter type, there are different geometry devices. These include a large ($200\mu \times 200\mu$) and small ($76\mu \times 76\mu$) device

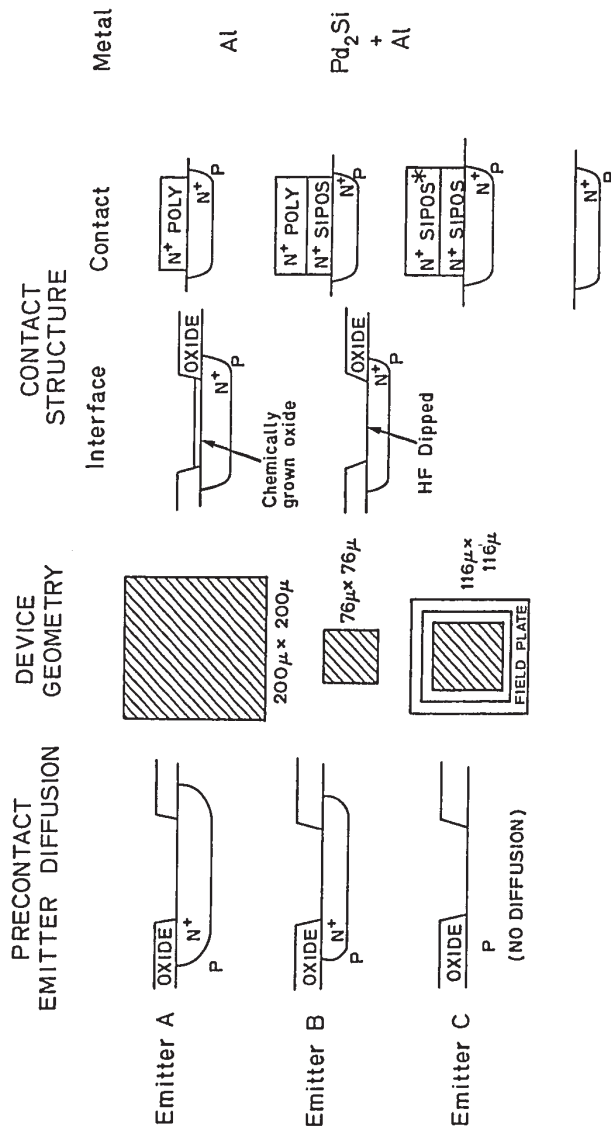


Figure 7.1: selective contact test structures

for determination of the area and perimeter components of the base current (see Chapter 6) and a device equipped with a field plate over the base-emitter junction whose potential can be independently set (Appendix K). For a given die, there is a choice of either leaving the chemical oxide formed by the cleaning procedure (see Appendix L for details of the chemical oxide) or stripping in 50:1 HF prior to SIPOS or POLY deposition. Stripping does not guarantee an oxide free interface as will be noted later. The deposited films are either N^+ POLY, a composite SIPOS sandwich or no deposited film. The metallization introduced another process split, some devices receiving a Pd evaporation and sinter prior to Al evaporation.

7.2. Format of Presentation

This study encompasses three main classes of emitters - “conventional”, SIPOS and POLY emitters. Conventional emitters consist of an emitter diffusion (type A or B) contacted by metal. They are present on every wafer and are intended for use as process monitors and to serve as benchmarks against which the more elaborate structures could be compared. SIPOS emitters (types A,B,C) were fabricated alongside the “conventional” devices. A separate set of wafers contained POLY and “conventional” emitters. Process splits were structured so that wherever possible, die originated from the same parent wafer. For example, the process split comparing devices with Pd+Al to devices with only Al was accomplished by shadowing the Al only devices with a mask during the Pd evaporation. Adherence to such a policy reduces uncertainties that would otherwise be introduced by process induced variations. All starting wafers were processed in the same batch up to the opening of the emitter window; therefore the diffusions forming the monocrystalline regions of type A and B devices were all performed using the same predeposition and drive-in.

For each class of device, experimental measurements of J_{oe} are presented followed by a brief discussion of salient points. This is followed by a modeling of the emitter structure using the numerical algorithm of Chapter 3 to see if the J_{oe} can be

explained in terms of the postulated mechanisms.

7.3. Conventional Emitters

The conventional emitter structures are present on every device wafer and act as process monitors. Parameters for the previously mentioned profiles are follows:

	Predep (T/t)	Drive-in (T/t)	Sheet Resistance	Junction Depth
EM-A	800°C/30min	1025°C/90min	60 $\Omega/\square$	1.0 μ
EM-B	800°C/30min	1025°C/30min	70 $\Omega/\square$	.66 μ

Phosphine (1100 ppm) and oxygen (63,000 ppm) in a nitrogen carrier constituted the predeposition ambient. Wafers were pushed in N_2/O_2 and allowed to equilibrate 10 minutes prior to introduction of the phosphine. The resulting P-glass was stripped prior to the drive-in which took place in a dry oxygen ambient with a 2% TCA flow to act as a getter for metalics. These emitters were left covered by a thermal oxide, but underwent the thermal cycles associated with POLY/SIPOS depositon and anneals which constituted the last high temperature steps of the process. Opening of the emitter contact windows was followed by evaporation of Al and then a forming gas anneal in 5% H_2 in an argon carrier.

Gummel plots for a large and small device are shown in Fig. 7.2a,b respectively. The collector current exhibits ideal behavior up to approximately .1mA where deviations are introduced due to series resistance in the base-emitter circuit. Using the procedure outlined in Chapter 6, the voltage drop induced by this series resistance is estimated by taking a tangent to the collector characteristic at low currents where it behaves ideally. In practice, this is done by choosing two points in the low current regime (typically $< 1 \times 10^{-5} A$) and taking a least squares fit to all (V_{be}, I_c) points bracketed by these points. This extrapolated ideal characteristic serves as a reference line from which to determine the voltage drop ΔV across the series resistance. This voltage is then subtracted from the measured (V_{be}, I_b) curve at the same V_{be} to yield the corrected base characteristic as shown in Fig. 7.2a,b. If the assumption of

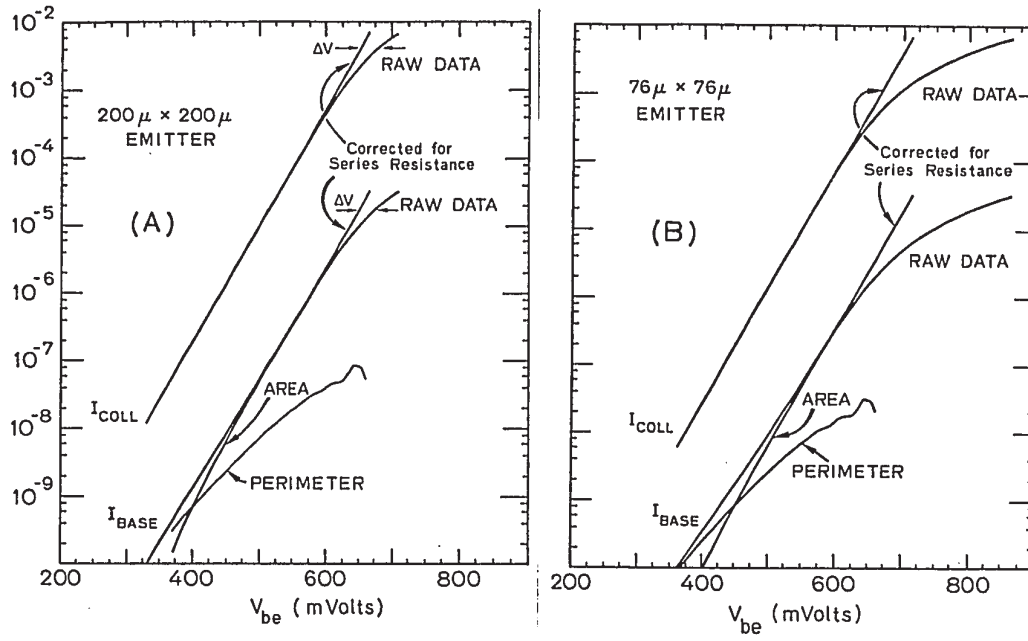


Figure 7.2: effect of series resistance on BJT characteristics

an extrinsic series resistance is valid (i.e. the deviation from ideal characteristic is not due, for example, to current crowding or high level injection) the corrected base characteristic will normally parallel the ideal collector characteristic; equivalently, the beta remains relatively constant.

From the corrected I_c and I_b vs V_{be} curves for the small and large devices, the area and perimeter components of the base current can be determined using equation (6.2). For the small device with its larger perimeter to area ratio, the perimeter component is more significant than for the larger device. This is evident from a comparison of Fig. 7.2 (a) and (b). The perimeter component also displays an ideality factor > 1 . This indicates that mechanisms such as space charge recombination are dominant contributors. It was shown in Chapter 6 that surface recombination where the emitter-base depletion region intersects the $Si-SiO_2$ in-

V_{be} (volts)	J_{oe} ($10^{-12} A/cm^2$)	V_{be} (volts)	J_{oe} ($10^{-12} A/cm^2$)
.392	1.18	.545	1.18
.409	1.21	.562	1.17
.426	1.21	.579	1.17
.443	1.20	.596	1.16
.460	1.19	.613	1.16
.477	1.19	.630	1.15
.494	1.19	.647	1.13
.511	1.19	.664	1.13
.528	1.18		

Table 7.1: measured J_{oe} vs. V_{be} for BJT with conventional emitter

terface is the dominant source of this perimeter component. Once the area related component is determined, J_{oe} can be computed by assuming an effective contact area. As discussed in Appendix H, this area is bounded by the emitter window and emitter diffusion areas. For the high contact recombination velocities characteristic of metallic contacts, this area is effectively given by the window area. A sample listing of J_{oe} vs V_{be} is presented in Table 7.1. The constancy of J_{oe} in this case is quite remarkable, the deviations being only a few percent over four orders of magnitude of base current. This indicates that the area related component is dominated by recombination in the neutral emitter and contact regions. Neutral base recombination is negligible for these high values of J_{oe} as was shown in Chapter 6. The constancy of J_{oe} also indicates that the junction temperature was correctly evaluated since an error in the temperature would introduce a spurious bias dependence into J_{oe} (strictly speaking, this only shows that the product mT where m is the ideality factor is correctly evaluated, but it seems unlikely that an error in T would be compensated exactly by the ideality factor, so it seems more reasonable to assume that m is identically unity) .

Anneal	$J_{oc} (A/cm^2)$	$G_b (s/cm^4)$
From Evap.	9.4×10^{-13}	1.86×10^{11}
300°C	9.4×10^{-13}	1.82×10^{11}
350°C	1.1×10^{-12}	1.75×10^{11}

Table 7.2: J_{oc}, G_b vs. FGA temperature ($T_A = 850^\circ C$) EM-B device

7.3.1. Influence of Aluminum Anneal Temperature

The tendency for *Si* to dissolve in *Al* can modify the morphology at the contact interface even at temperatures below the *Al-Si* eutectic. This can be accompanied by an increase in J_{oc} and reverse leakage and for very shallow junctions can result in junction spiking [1]. Although for the low ($< 200 A/cm^2$) current densities used in this study, relatively high specific contact resistances can be tolerated (up to $\approx 1 \times 10^{-3} \text{ ohm-cm}^2$), accuracy is enhanced if this series resistance can be minimized. Furthermore, annealing, especially with the presence of *Al* drastically, reduces the interface state density at the *Si-SiO₂* interface [2]; this helps to reduce the contribution of the perimeter component which in turn results in more accurate values of J_{oc} . The first set of tests was conducted to establish the highest annealing temperature that can be used without a significant increase in the J_{oc} . In these transparent emitters, this is an indication that the interfacial morphology is being significantly altered. The results for a device that had undergone a final high temperature anneal (T_A) of $850^\circ C$ in Table 7.2 indicate that for a $300^\circ C/15 \text{ min}$ anneal no change in J_{oc} is observed. At $350^\circ C$, a slight increase is apparent indicating that this temperature marks an upper limit that can be used without a more elaborate metallization scheme. The impact of anneal temperature on the SIPOS and POLY contacts will be discussed in the relevant sections. The base Gummel numbers display an apparent decrease with anneal temperature, although this decrease is slight amounting to only 6% and could be a coincidental result. It runs counter to the trend expected since the decrease in series resistance should produce an increase in the extrapolated collector saturation current and hence an increase in the effective

	EM-A		EM-B	
T_A	J_{oe} (A/cm^2)	G_b (s/cm^4)	J_{oe} (A/cm^2)	G_b (s/cm^4)
800°C	8.5×10^{-13} (2%)	8.9×10^{10} (2%)	1.1×10^{-12} (3%)	1.94×10^{11} (6%)
850°C	—	—	1.1×10^{-12} (6%)	1.75×10^{11}
900°C	8.3×10^{-13} (2%)	1.05×10^{11} (4%)	1.2×10^{-12} (4%)	1.96×10^{11} (2%)
1000°C	8.1×10^{-13} (2%)	9.16×10^{10} (1%)	1.1×10^{-12} (9%)	1.66×10^{11} (1%)

Table 7.3: J_{oe} , G_b vs. T_A , final high temperature anneal

Gummel number (see Chapter 6) . In any case, it is not likely at these temperatures that the base parameters are being modified. The J_{oe} for the remainder of the conventional emitter structures refer to devices that have been annealed at 350°C.

7.3.2. Influence of Final High Temperature Anneal

Since these emitters were driven in at 1025°C for 90 and 30 minutes for EM-A and B respectively, the short (< 15 min) anneals at 800°C, 850°C and 900°C used to anneal the SIPOS and POLY films should not perceptibly alter the emitter profile. The 1000°C anneal may have some impact on the junction depth for EM-B, but due to variations in epi thickness, it is not possible to directly compare base Gummel numbers between different wafers which would otherwise provide a direct indication of junction migration. The base Gummel numbers for the conventional structures are useful, however, for evaluating the extent of outdiffusion for adjacent emitter structures such as the SIPOS and POLY emitters.

Referring to Table 7.3, the J_{oe} for both EM-A and EM-B remain essentially unchanged by the high temperature anneals. The quantities in parenthesis are the standard deviations of the measured quantities, expressed as a percentage of the mean values.

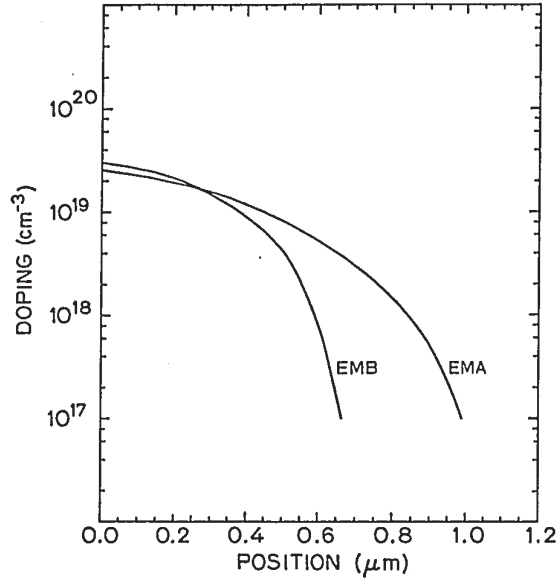


Figure 7.3: dopant profiles used in simulations of EM-A,B

7.3.3. Modeling the Conventional Emitter

The profiles for device types A and B relevant for the 800°C , 850°C and perhaps 900°C determined by spreading resistance are shown in Fig. 7.3 . The raw profiles for EM-A displayed a nonphysical dip in the profile near the surface. The junction depth agreed with that obtained by beveling and junction staining, but the sheet resistivity was approximately 16% too high. The erroneous surface points were replaced by an “analytic continuation” of the profile to the surface; better agreement with the sheet resistivity determined by independent four point probe measurements was obtained with the “corrected” profile. The profile for EM-B did not require this correction as agreement with both junction depth and sheet resistivity was satisfactory.

The dependence of J_{oc} for EM-A and EM-B using the profiles in Fig. 7.3 are shown in Fig. 7.4a,b respectively. Looking at the metallic contact for which S_c is in

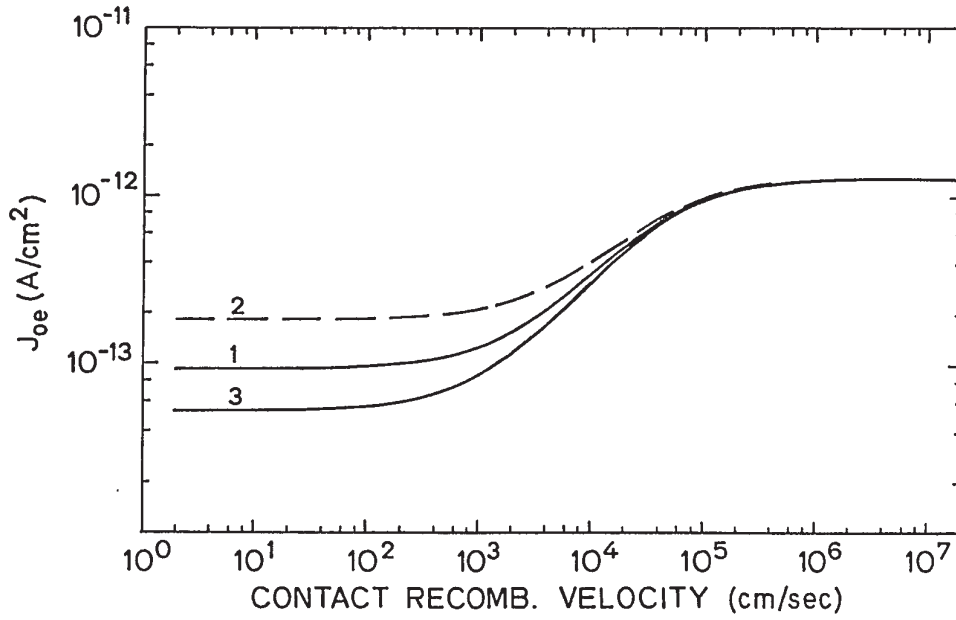
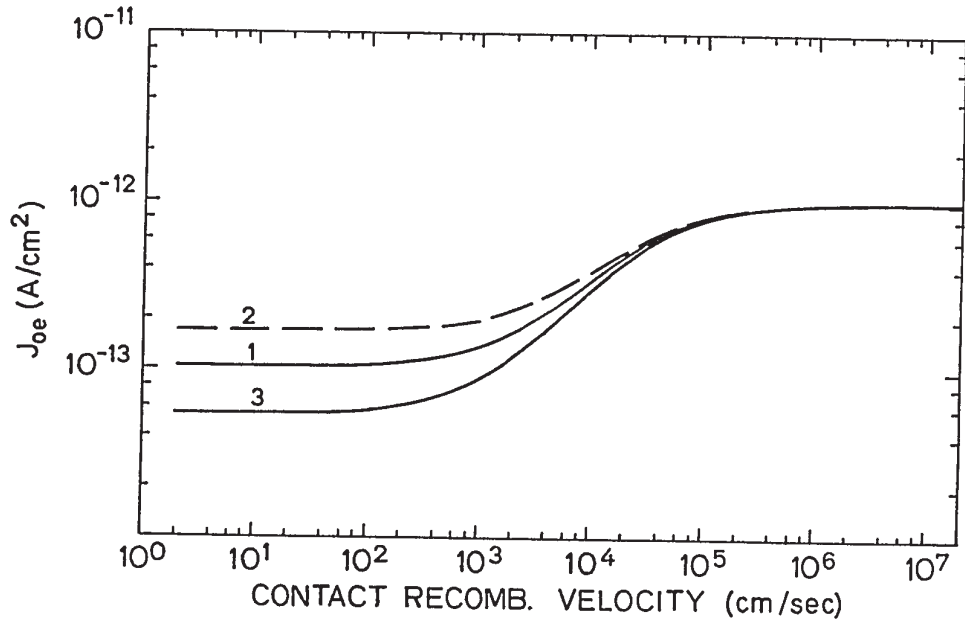


Figure 7.4: J_{oe} vs. S_c for EM-A (A) and EM-B (B)

	C_{srh}	C_{auger}
curve 1:	1.67×10^{-12}	5×10^{-32}
curve 2:	0	2.8×10^{-31}
curve 3:	5×10^{-13}	5×10^{-32}

Table 7.4: lifetime parameters used in J_{oe} simulations

the 1×10^6 to 1×10^7 cm/sec range, J_{oe} using the hole transport parameters in ref. [3] is predicted to be 1×10^{-12} and $1.25 \times 10^{-12} A/cm^2$ for EM-A and EM-B respectively. This compares with the experimental values 8.3×10^{-13} and $1.1 \times 10^{-12} A/cm^2$. The agreement is fair, the calculated J_{oe} is approximately 15% high in both cases. This is an indication that either the actual band gap narrowing is somewhat less than that predicted by the model used in these calculations, or that the assumed diffusion coefficient is too high, but the error is not sufficient to warrant a change at this point in view of the numerical uncertainty in the spreading resistance profiles. For each emitter, there is a family of three curves, each reflecting the use of a different set of lifetime parameters as listed in Table 7.4. in the lifetime model

$$\frac{1}{\tau_p} = C_{srh} \cdot N + C_{auger} \cdot N^2$$

where N is the doping density. The J_{oe} for large surface recombination velocities is insensitive to the lifetime parameters used, since the emitter remains transparent and the impact of bulk recombination on J_{oe} is negligible. However, the disparities in the low S_c regime can be resolved by the experimental values of J_{oe} measured on structures with low surface recombination velocity.

7.4. SIPOS Emitters

The heterojunction model of Matsushita discussed in Chapter 5 relies on the presence of a barrier in the valence band introduced by a postulated larger bandgap of SIPOS. Hence, the bulk properties of SIPOS especially as controlled by the oxygen

content, play a key role in that model. The first set of experiments with SIPOS emitters were designed to test this hypothesis.

7.4.1. Role of Oxygen in Determining J_{oe} :

A two layer SIPOS composite layer was deposited on structures EM-A and EM-B (EM-C was not available for this set of devices). As usual, adjacent structures with standard aluminum contacts were available to serve as benchmarks. The two layer structure consists of a thin SIPOS layer with varying oxygen content capped by a second layer where the nitrous oxide was turned off during the deposition. This layer, which is ostensibly silicon, serves to act as a buffer layer between the metallization and the thin “active” SIPOS layer underneath. It also provides good ohmic contact which is difficult to obtain if high oxygen content SIPOS is directly contacted by the metallization. If it is desired to use Pd as part of the metallization, this buffer layer also allows the formation of Pd_2Si which again is not possible with SIPOS that contains moderate amounts of oxygen. Structural and process parameters for this set of devices is provided in Table 7.5.

J_{oe} as a function of γ , the ratio of partial pressures of nitrous oxide to silane in the deposition ambient used during the deposition of layer “A”, are shown in Fig. 7.5. The value for $\gamma = 1.5$ should produce films comparable in oxygen content to those of Matsuhita [4] from extrapolation of oxygen content as determined by electron microprobe measurements for similar films. Wafers B2UL,F1, and F1L were annealed at $900^\circ C$ for a total time of 13 minutes (this implies a net time of approximately 11 min after discounting equilibration time). The rationale behind the use of such short anneal times, despite the problems imposed on process repeatability, can be attributed to balancing the need to restructure the SIPOS film to obtain good electronic properties (activation of dopant etc.) against the excessive outdiffusion of impurities into the bulk silicon which would degrade J_{oe} as shown in Chapter 3. Since the restructuring appears to take place in two steps - a rapid initial phase lasting less than 5 minutes followed by a phase characterized by gradual

Device	$N_2O(ppm)$	γ	$PH_3(ppm)$	thickness (Å)
Layer A				
F1	—	—	—	—
B2UL	300	.7	33	1000
B2L	300	.7	33	1600
F1L	600	1.5	33	1600
Layer B				
F1	off	0	33	3000
B2UL	off	0	8	2000
B2L	off	0	33	2000
F1L	off	0	33	2000

Table 7.5: process data for J_{oe} vs. γ device set

growth of the silicon grains [5] - it was felt that the use of anneals significantly in excess of 10 minutes would provide only marginal restructuring of the SIPOS while still contributing to phosphorous outdiffusion. Taken to the limit, this strategy would suggest the use of rapid thermal annealing techniques such as flash lamps; if the apparent “activation energy” for restructuring of the SIPOS film is greater than that for dopant diffusion, then the use of high temperature short duration anneals would be favored. For these experiments, such apparatus was not readily available, so short furnace anneals were used instead.

J_{oe} for $900^\circ C/13min$ anneals shown in Fig. 7.5, seem to be insensitive to the oxygen content of the films as reflected by the value of γ used during the deposition. If the model of Matsushita is correct, then the effective surface recombination velocity seen at the *Si*-SIPOS interface in the simplest model should be given by

$$S = S_{int} + v_{transp} \cdot e^{-\Delta E_v/(kT)}$$

where v_{transp} is the effective transport velocity in the SIPOS and ΔE_v is the barrier in the hole valence band which is postulated to be controlled by the oxygen content

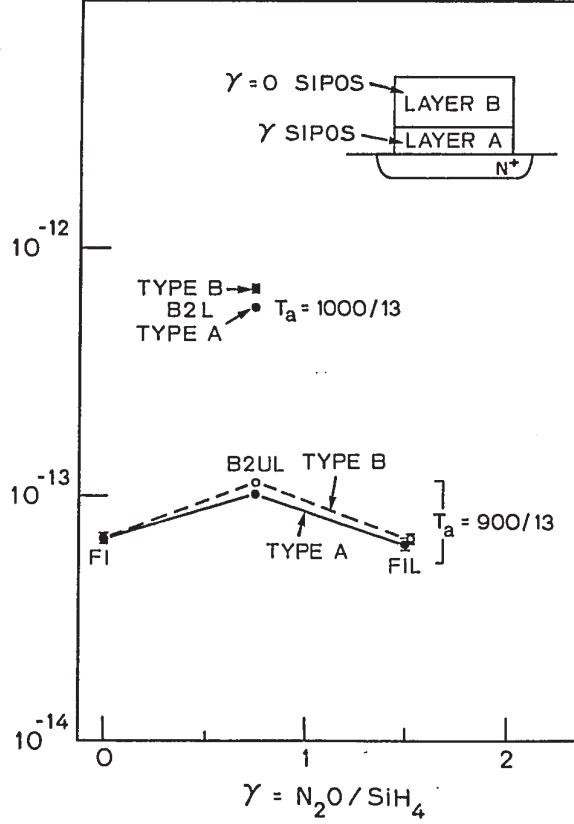


Figure 7.5: influence of oxygen content of SIPOS on J_{oe}

in the SIPOS. Based on this picture, one might have expected to see a monotonic decrease in J_{oe} with γ , the rate of decrease eventually becoming controlled by interfacial recombination when ΔE_v becomes large enough that

$$v_{transp} e^{-\Delta E_v / (kT)} < S_{int}.$$

This trend is not seen. Furthermore, the $\gamma = 0$ film, which is ostensibly POLY, performs as well as the high oxygen content films. Even assuming an extreme value of $D_{p,POLY} = D_{p,Si}/10$, the results for a composite 3000 Å POLY + EM-B structure yields a J_{oe} of $6 \times 10^{-13} \text{ A/cm}^2$, clearly far in excess of the observed J_{oe} .

7.4.2. The $\gamma = 0$ Film

In fact, turning off the nitrous oxide does not result in the deposition of “pure” POLY; rather, electron microprobe measurements indicate the presence of 1 – 2at% oxygen. Since the deposition ambient contains only 400 *ppm* of silane, it is conceivable that trace amounts (a few *ppm*) of water vapor and oxygen could account for this residual level of oxygen. These $\gamma = 0$ layers were also susceptible to attack by HF while POLY films deposited in an epi reactor using a hydrogen ambient remain inert. The role of small quantities of oxygen in SIPOS has not been investigated. It is possible that oxygen segregates to the *Si*-SIPOS interface during the reconstruction process as has been proposed for higher oxygen content films [6]. The fact that the films appear to disintegrate in HF into a fine dispersion of momentarily visible particles also suggests that Ni’s shell model may be appropriate even in this low oxygen content range. It seems unlikely from the small amount of oxygen present and the low resistivity of these $\gamma = 0$ SIPOS films that the oxygen rich shells are very effective barriers for either type of carrier. In that case, the J_{oe} reduction must be attributed to interfacial mechanisms since the oxygen content is too small for a wider gap to form in the SIPOS as required by the model of Matsushita [4].

7.4.2.1. J_{oe} vs. Junction Depth

J_{oe} for EM-A and EM-B are quite similar. This is not surprising in view of the simulations performed for these two profiles where for $S_e < 1 \times 10^4 \text{ cm/sec}$ the predicted J_{oe} was the same for both profiles.

7.4.2.2. J_{oe} vs. Anneal Temperature

Devices B2L and B2UL had similar structures (see Table 7.5) . The difference in thickness of the active SIPOS layer and difference in the phosphine flow rates for layer B should have marginal impact. Yet, Fig. 7.5 indicates that increasing the annealing temperature to 1000°C/15min increases J_{oe} by almost an order of magnitude. This dramatic increase in J_{oe} indicates that for this structure, annealing

T_A ($^{\circ}C$)	$J_{oc}(10^{-14} A/cm^2)$	$G_b(10^{11} s/cm^4)$		
Type B Device				
	no chem ox	with chem ox	no chem ox	with chem ox
800	5.4(6%)	—	1.62(5%)	—
850	5.62(1%)	5.8(5%)	1.74(2%)	1.78(3%)
900	9.5(8%)	6.4(10%)	1.93(2%)	1.97(2%)
1000	50(10%)	29.6(9%)	1.45(2%)	1.49(1%)
Type C device				
800	9.6(7%)	—	3.37(5%)	—
850	1.67(6%)	1.2	3.43(1%)	3.55
900	5.75(5%)	1.2(10%)	3.62(1%)	3.76(2%)
1000	53.4(3%)	21.2(6%)	2.54(1%)	2.82(1%)

Table 7.6: J_{oe}, G_b vs. SIPOS anneal temperature

temperature is more important than oxygen content from the viewpoint of attaining low J_{oe} structures.

At this point, questions still remained as to the optimal anneal temperature to use as well as the role of a possible interfacial oxide layer. The next set of experiments were designed to address these issues.

7.4.3. Influence of Anneal Temperature and Interfacial Oxide

After growing a chemical oxide (see Appendix L) on this set of four wafers, one side of the wafers was dipped in 10:1 HF to strip the chem ox (chemical oxide) from the emitters windows prior to deposition of the 3300 Å, $\gamma = 0$ SIPOS layer; phosphine and silane were maintained at 33 and 410 ppm respectively. The wafers, which were processed simultaneously to this point, were then given separate high temperature anneals of 800°C, 850°C, 900°C, and 1000°C for 15min (total anneal time). The J_{oe} results are summarized in Table 7.6.

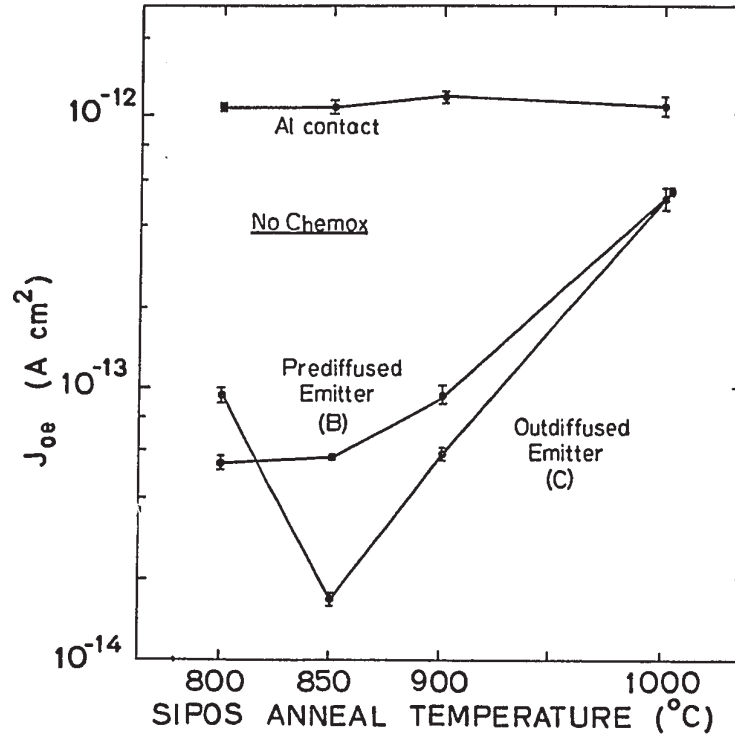


Figure 7.6: influence of anneal temperature on J_{oe} for SIPOS contacted emitters without chemox

7.4.4. Devices with No Chem Ox

For devices where the chem ox was stripped prior to SIPOS deposition, the J_{oe} values are plotted in Fig. 7.6. The only observation of significance regarding the “conventionally” contacted reference devices is that the J_{oe} is relatively constant, indicating that there were no irregularities in the processing or J_{oe} extraction procedures. Looking at the curves for the devices with the prediffused profile EM-B (device type “B”), J_{oe} remains fairly constant at 800°C and 850°C at a value of $5 - 6 \times 10^{-14} A/cm^2$. This is factor of 20 improvement over the metal contacted devices and well below the “fundamental limit” that can be achieved with conven-

tionally contacted Si homojunctions regardless of the profile (see chapter 3). At 900°C , J_{oe} increases to $9 \times 10^{-13} \text{A/cm}^2$ and at 1000°C to $5 \times 10^{-13} \text{A/cm}^2$, a ten-fold increase over the devices annealed at lower temperatures. For this structure, anneals at 800°C or 850°C produce the lowest J_{oe} . For the structure EM-C which has no pre-diffused emitter and relies solely on outdiffusion of phosphorous from the SIPOS layer to create the emitter-base junction (device type "C"), J_{oe} at 800°C is much larger than for the corresponding device type B. At 850°C the opposite is true and J_{oe} for device type C falls below $2 \times 10^{-14} \text{A/cm}^2$. At 900°C , it increases again to $5 - 6 \times 10^{-14} \text{A/cm}^2$ and at 1000°C is about the same as for the B type device. An interesting pattern emerges if a "baseline" J_{oe} is set at $5 \times 10^{-14} \text{A/cm}^2$ for the prediffused emitter. The incremental increase in J_{oe} above this baseline with increasing anneal temperature can be accounted for by the J_{oe} of the outdiffused emitter at the same temperature (except for the 800°C point). For example, at 850°C , for the type C device J_{oe} is $1.7 \times 10^{-14} \text{A/cm}^2$ and when added to the baseline of $5 \times 10^{-14} \text{A/cm}^2$ yields $6.7 \times 10^{-14} \text{A/cm}^2$ which is close to the actual measured value. At 1000°C , the increase for the type C device is so large that it dominates the baseline value and the resulting J_{oe} is roughly the same for both types of devices regardless of whether or not a prediffused emitter was present. Apparently, annealing at higher temperatures is adding a recombination component that is common to both device structures.

7.4.5. Modeling SIPOS Contacts Without a Chem Ox

For the low anneal temperatures (800°C , 850°C) outdiffusion is minimal and it is appropriate to use the EM-B profiles provided in Fig. 7.3. J_{oe} for this profile was computed as a function of S_c , the contact recombination velocity at the Si-SIPOS interface in Fig. 7.4b. Using the model parameters in [3] (curve 1), the J_{oe} at large S_c , which corresponds to the metallic contact, exhibit good agreement with the measured values, but in the low S_c regime, J_{oe} is at least a factor of two too large. Since the small Auger coefficient used leads to the dominance at

these doping levels of the SRH term, reducing the SRH coefficient by a factor of 3 (curve 3) provides good agreement at both extremes. As a test case, use of the commonly used value of $2.8 \times 10^{-31} A/cm^2$ for the Auger coefficient (curve 2) leads to a predicted value for J_{oe} that is far too high, even when the SRH component has been set to zero. This indicates that the smaller value for C_{auger} is more consistent with these experimental results. This triple set of lifetime coefficients will be used in the remaining simulations to evaluate the relative merits of each set.

The outdiffused layer at $800^\circ C$ is difficult to profile due to its shallow depth. The exact details of the profile are irrelevant since such thin emitters become transparent even for small values of S_c , and transport is limited by the surface recombination velocity, i.e.

$$J_{oe} = q \cdot S_c \cdot \frac{n_{io}^2}{N_{s,eff}} \quad (7.1)$$

where $N_{s,eff}$ is the effective surface dopant concentration. This suggests several mechanisms for the J_{oe} increase noted for the $800^\circ C$ anneal. Since the type B devices require surface recombination velocities well below $1000 cm/sec$ in order to explain their J_{oe} , and since to first order it seems reasonable to assume similarly low values for S_c for both device types since they were processed identically, it seems that the best explanation relies on the $N_{s,eff}$ term. At $800^\circ C$, the outdiffusion is likely to be minimal, and $N_{s,eff}$ probably has not yet attained its optimal value. This effect is also manifested as an increase in the majority carrier series resistance as will be discussed later.

The SRP (Spreading Resistance Profile) for an outdiffused layer after an anneal of $850^\circ C/15min$ indicates an outdiffusion of approximately $600 \text{ \AA}$ with a surface concentration of $8 \times 10^{18} cm^{-3}$. Admittedly, such shallow regions are difficult to profile with SRP, but this outdiffused regions lies in the shoulder region of the SIPOS profile and the distortion of shallow profiles induced by the presence of a $P-N$ junction is minimized. As indicated, the key parameter is the surface concentration, and at $8 \times 10^{18} cm^{-3}$, the value of $N_{s,eff}$ is almost saturated so upward variations in this parameter will not strongly impact the results. Simulations using the three lifetime

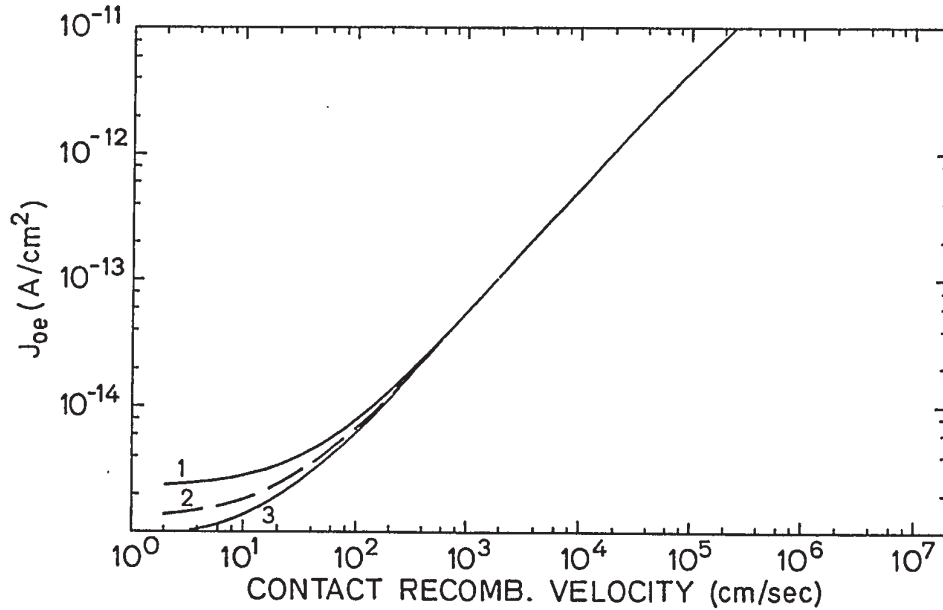


Figure 7.7: simulated J_{oe} vs. S_c for ERFC profile assumed for SIPOS contact without chemox

parameter sets for an ERFC profile with $8 \times 10^{18} \text{cm}^{-3}$ surface concentration are plotted in Fig. 7.7. As expected, the three curves merge for $S_c > 100 \text{ cm/sec}$ as the emitter becomes transparent and transport is limited by the surface recombination velocity. For the experimentally measured value of J_{oe} at 850°C which has an upper bound of $1.5 \times 10^{-14} \text{A/cm}^2$, the surface recombination velocity must be less than 200 cm/sec regardless of the lifetime parameter set used. This range for S_c is consistent with the values required to match the J_{oe} results for device type B.

Increasing the anneal temperature to 900°C marks the onset of J_{oe} degradation. Since the starting J_{oe} for the EM-B profile is only $5 \times 10^{-14} \text{A/cm}^2$, any recombination in the heavily doped surface regions arising from the outdiffusion of dopant will essentially add “linearly”; this follows from the observation that the hole imref is relatively flat across the lightly doped region closer to the emitter-base junction. In

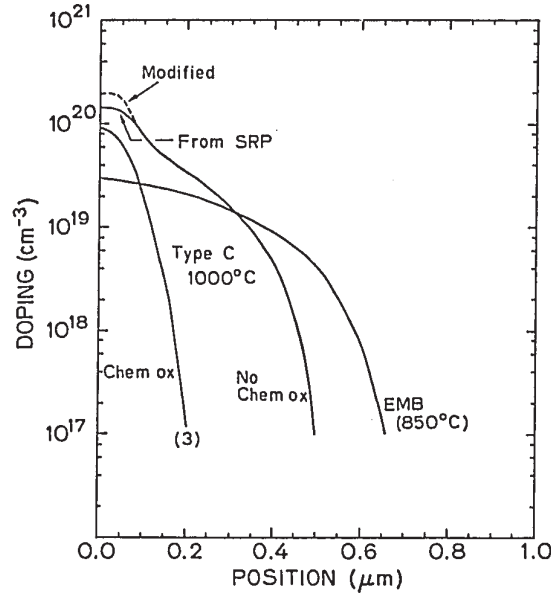


Figure 7.8: dopant profiles used to simulate SIPOS contacted devices

modelling the 1000°C results, the SRP relevant to device type C shown in Fig. 7.8 is used. At this high annealing temperature, a heavily doped region exhibiting the characteristic “kink” typical of phosphorous profiles in this doping range dominates the recombination picture. Looking at the J_{oe} vs S_e curves in Fig. 7.9, the dominant role of Auger recombination in the near surface region is evident. Using the standard parameter set (curve 1) gives a J_{oe} of $3.5 \times 10^{-13} \text{ A/cm}^2$ for $S_e = 200 \text{ cm/sec}$. Reducing the SRH coefficient by 3 only slightly lowers J_{oe} to $2.9 \times 10^{-13} \text{ A/cm}^2$. The discrepancy with the experimentally measured values of $5 \times 10^{-13} \text{ A/cm}^2$ may be attributed to either a large increase in the surface recombination velocity or to an error in the SRP.

Since the heterojunction model includes no detailed model for S_{int} , there is no mechanism available to explain the increase in S_e to $1 \times 10^4 \text{ cm/sec}$ needed to explain the discrepancy if in fact the SRP is correct. In an SIS model of SIPOS contacts,

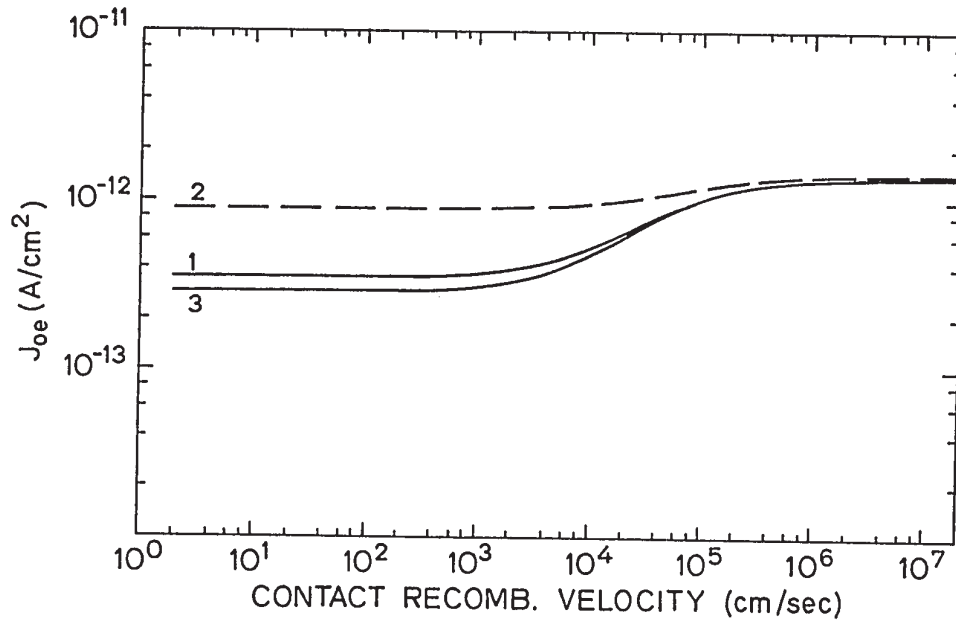


Figure 7.9: simulated J_{oe} for $T_A = 1000^\circ C$ SIPOS without chemox

where the presence of a thin interfacial oxide formed by the segregation of oxygen to the interface limits the hole transport, it is possible to envision that this layer may lose structural integrity at elevated temperatures. The large stresses associated with the tendency of the SIPOS to form crystalline regions coupled with the observation that silicon dioxide can flow viscously at temperatures above $950^\circ C$ indicates that disruption of this layer is possible. For low oxygen content films, it is likely that the oxide shells if present are incomplete; the low resistivity ($1 \times 10^{-3} \text{ ohm-cm}$) of SIPOS indicates that these oxide shells do not greatly hinder electron transport; if hole transport is similarly unimpeded, then values for the effective transport velocity in excess of $1 \times 10^4 \text{ cm/sec}$ are likely, in view of the fact that even monocrystalline silicon possesses transport velocities in excess of $8 \times 10^3 \text{ cm/sec}$. Therefore, disruption of the interfacial oxide layer can result in a large increase in S_e to $1 \times 10^4 \text{ cm/sec}$ and beyond.

The other mechanism relates to errors in the SRP profile. Since the error margin for resistivity is 15% and the volume correction factor adds 35% [8], there is room for considerable deviation. If the profile shown in Fig. 7.8 is modified so that the concentration in the flat topped region near the surface is $2 \times 10^{20} \text{ cm}^{-3}$ instead of $1.4 \times 10^{20} \text{ cm}^{-3}$, then using the smaller Auger and SRH coefficients yields a J_{oe} of $4.3 \times 10^{-13} \text{ A/cm}^2$, in fair agreement with the measured value of $5 \times 10^{-13} \text{ A/cm}^2$. Since the overlying SIPOS could not be selectively removed, it was not possible through independent measurement of the sheet resistivity to test the validity of this hypothesis. It is also possible to obtain better agreement by further adjustment of the Auger coefficient, but this is a largely unproductive exercise in view of the uncertainty in the profiling. At this point, since the low temperature profiles of EM-A and B on which lifetime parameter set three was established are felt to be more reliable, it remains the preferred set. Note that the use of the "standard" value for the Auger coefficient of 2.8×10^{-31} (curve 2) again leads to an excessive J_{oe} assuming that the profile is correct.

For this group of SIPOS contacted devices with no chemical oxide, optimal J_{oe} was found to result for an anneal temperature of 850°C . Anneals below and above this temperature increase J_{oe} most likely because of insufficient and excessive dopant outdiffusion respectively, although the role of S_{int} may not be negligible. Modeling results yield satisfactory agreement if the lifetime parameters are set at

$$C_{auger} = 5 \times 10^{-32} \text{ cm}^6 \text{ sec}^{-1}$$

$$C_{srh} = 5 \times 10^{-13} \text{ cm}^3 \text{ sec}^{-1}$$

7.4.6. SIPOS Devices with Chem Ox

Processed in parallel with the SIPOS contacted devices was another set of devices in which the chemical oxide was not stripped prior to the SIPOS deposition. These devices, which definitely possessed an SIS structure were intended to investigate the role of the interface in J_{oe} reduction. To minimize process variations, these

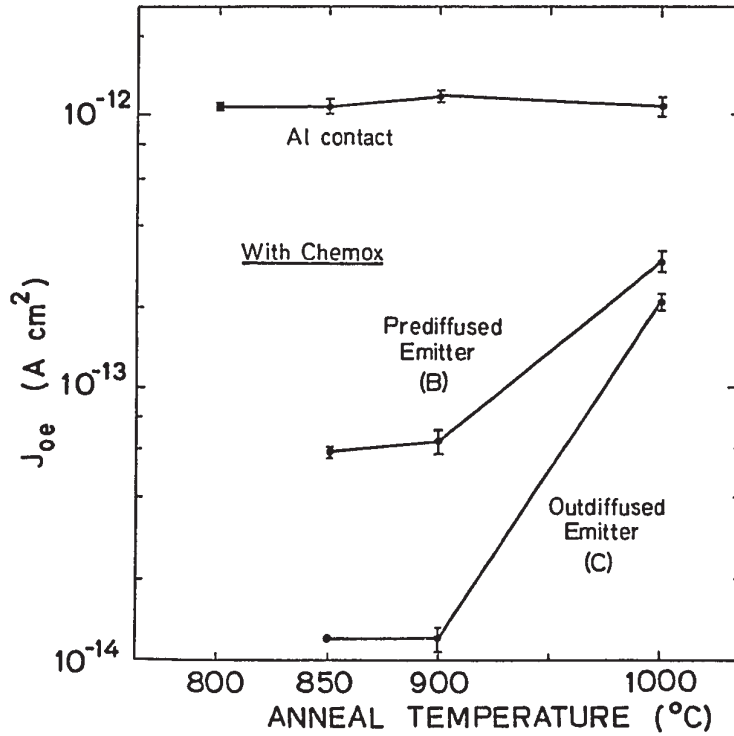


Figure 7.10: J_{oe} vs. T_A for SIPOS contacts with chemox

devices were located on a different region of the same wafer as the previous set of devices and therefore received the same SIPOS deposition and annealing treatments. The J_{oe} for 850°C, 900°C, and 1000°C are displayed in Fig. 7.10. Results for the 800°C were not available due to a problem with excessive series resistance which precluded determination of J_{oe} . However, this problem can be turned to advantage since investigation of the Gummel plots and beta vs I_c can yield insights into an interesting failure mode for the SIS structure. This will be discussed later.

Focusing on the type B device with the prediffused emitter, the J_{oe} measured for the 850°C anneal is the same as for the device where the chemical oxide was stripped. If J_{oe} reduction had relied solely on the bulk properties of the SIPOS, it

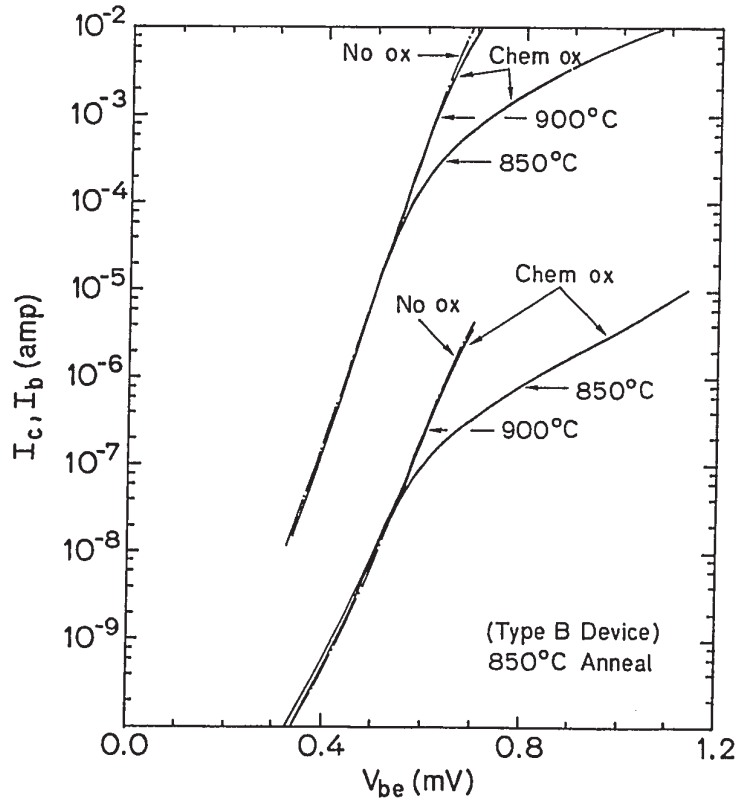


Figure 7.11: effect of chemox on series resistance of SIPOS contacts

seems unlikely that the introduction of the chemical oxide would have left the J_{oe} unchanged unless the interfaces were already the same to begin with. It seems more likely that J_{oe} is dominated by the same mechanism in both cases, namely interfacial recombination at the $Si-SiO_2$ interface. That a significant chem oxide layer was introduced is clear from the abrupt increase in series resistance evident in Fig. 7.11, where the raw data for two type B devices annealed at $850^\circ C$ is presented. Although the extracted value for J_{oe} is about the same for both devices, the presence of the chem oxide clearly introduces an additional voltage drop. The current gain for the chem oxide device is also seen to drop slightly at higher currents. This behavior is

accentuated in the 800C devices which will be discussed later.

As the anneal temperature is increased, J_{oe} increases as noted for the devices without the chem oxide, albeit at a slower rate. Looking at device type C at 850°C, J_{oe} of $1 \times 10^{-14} A/cm^2$ is attained which indicates that for devices annealed at this temperature, the introduction of the chemical oxide does not significantly alter J_{oe} . Again, the increase of J_{oe} with increasing annealing temperature can be explained by an incremental increase common to both B and C type devices. Phosphorous outdiffusion is likely the most significant mechanism, and the role of the chem oxide seems to be an indirect one, namely that of reducing the flux of outdiffusing phosphorous by acting as a diffusion barrier.

It was previously mentioned that the increase in J_{oe} with anneal temperature could also be explained in part by a disruption of the interfacial layer. In fact, as the anneal temperature is increased to 900°C, the excessive series resistance associated with the chemical oxide decreases dramatically as seen in Fig. 7.11. At first, this might seem to confirm that disruption of the layer has taken place, but there are other explanations. From equation (3.51), the contribution to the specific contact resistance made by the interfacial layer is given by

$$R_c \approx \frac{kT}{q^2} \frac{4}{v_{th}} \frac{1}{T_n N_s}$$

Disruption of the layer effectively increases the tunnel coefficient T_n to a value closer to unity in regions where the oxide loses integrity. However, the increased outdiffusion at elevated anneal temperatures also increases N_s , the surface dopant concentration which produces a similar lowering of R_c . From series resistance measurements alone, it is not possible to readily distinguish between these two mechanisms.

7.4.7. Modeling of a Chem Ox Device

The profile for a type C device with chem ox annealed at 1000°C as obtained by spreading resistance is shown as profile 3 in Fig. 7.8. Note that it is much shallower and has a much lower surface concentration than for the profile where the chem

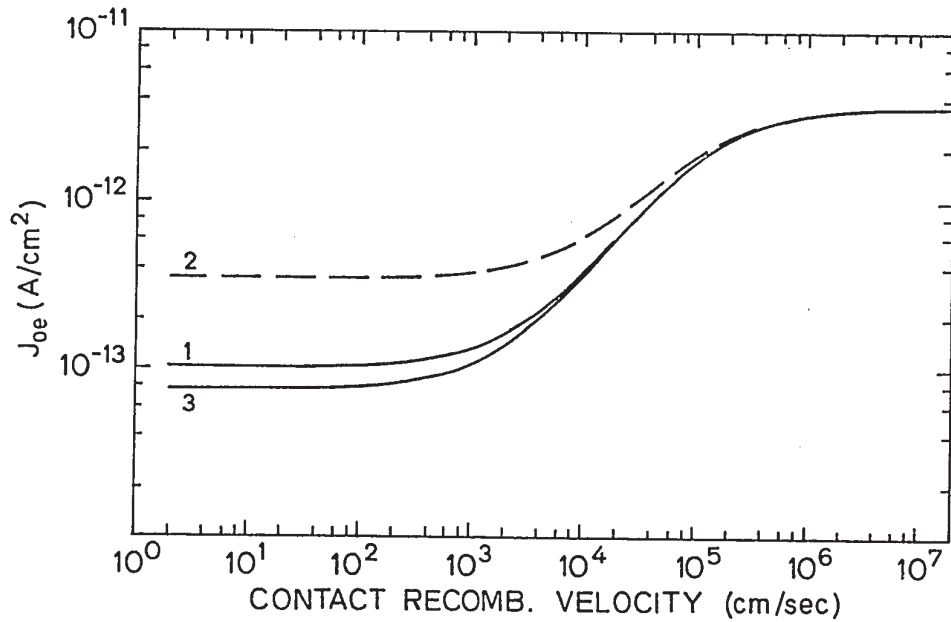


Figure 7.12: simulated J_{oe} vs. S_c for SIPOS contact with chemox

oxide is absent. The corresponding J_{oe} using the same three lifetime parameter sets is plotted in Fig 7.12. The predicted J_{oe} using the preferred lifetime parameters is $7.6 \times 10^{-14} \text{ A/cm}^2$. This compares with the experimental value of $2 \times 10^{-13} \text{ A/cm}^2$. The trend is correctly predicted, but again the predicted J_{oe} is too small. The same mechanisms invoked to explain the discrepancy for the devices without chem oxide also apply in this case. There is uncertainty as to the exact value of S_c ; if it has increased for the reasons mentioned previously, it is possible to resolve the differences in J_{oe} . The uncertainties in the SRP should be even more relevant to this profile which is steeper and shallower. Again, note that the use of the standard value for the Auger coefficient seems to predict a J_{oe} in excess of that experimentally observed.

7.4.8. 800°C SIPOS Devices With Chem Ox

The devices annealed at 800°C displayed peculiar characteristics due to the ex-

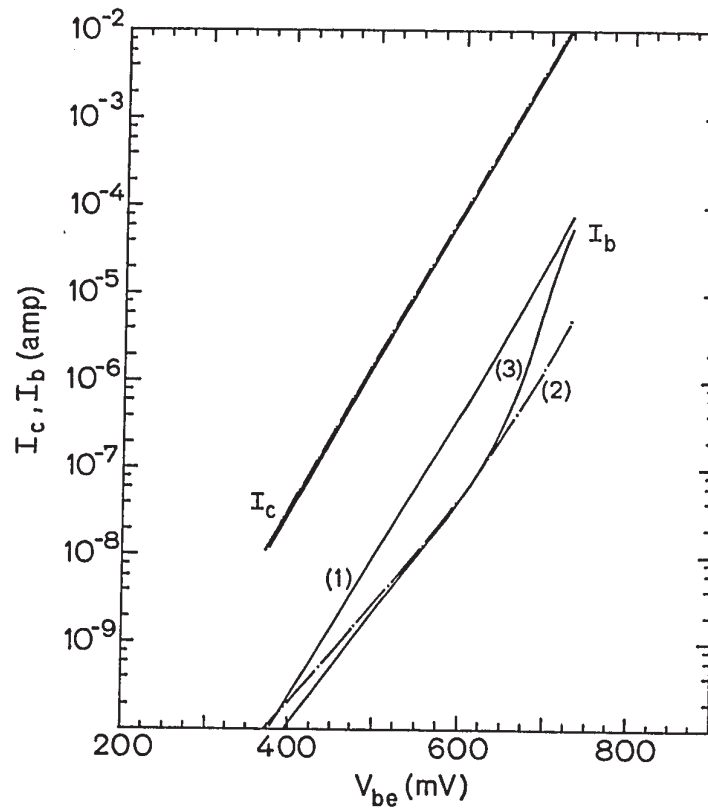


Figure 7.13: Gummel plots for emitters contacted by metal(1) SIPOS without(2) and with (3) chemox, $T_A = 800^\circ C$

cessive series resistance associated with the presence of the chemical oxide. Fig. 7.13 shows the Gummel plots corrected for series resistance for three devices on the same wafer. The first device (curve 1) was contacted with *Al*, the second (curve 2) with SIPOS only, and the third (curve 3) with chem ox and SIPOS. The behavior of the third device resembles that of device 2 at low currents, but at higher currents, “switches” to that of device 1, the *Al* contacted device. The excessive voltage drop which occurs across the chem oxide for device three (since the metallization and SIPOS layer are otherwise identical with those of device 2) apparently causes the

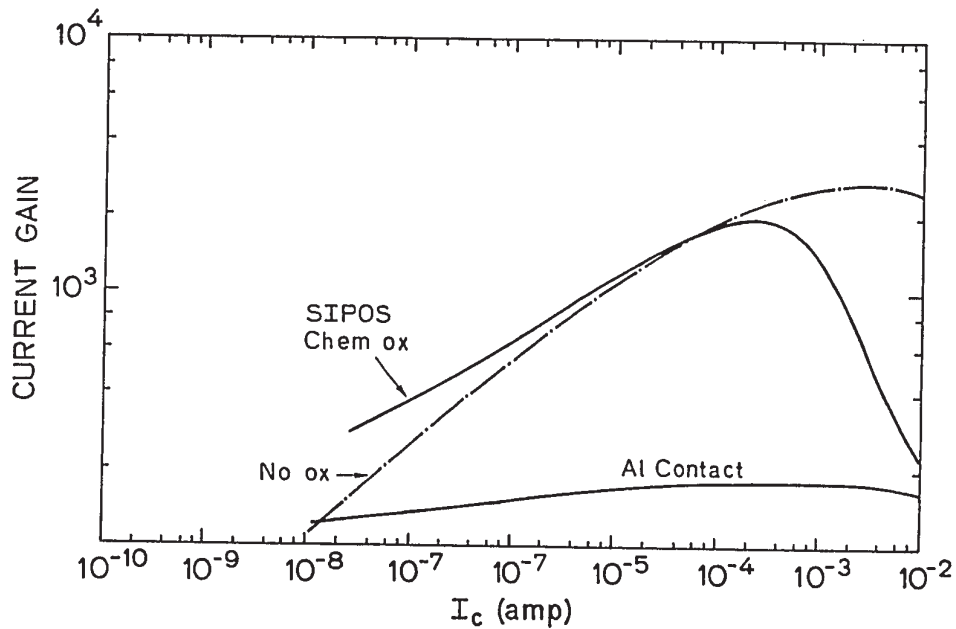


Figure 7.14: current gain vs. collector current for conventional and SIPOS (with and without chemox) devices; $T_A = 800^\circ\text{C}$

SIPOS contact in device 3 to behave as a “metallic” contact despite its lower J_{oc} at low currents. This behavior is apparent from an examination of the beta vs I_c characteristic in Fig. 7.14. The beta for the chem oxide device falls sharply and asymptotically approaches that of the metal contact. In Fig. 7.15 and 7.16, similar characteristics are displayed for a smaller ($76\mu \times 76\mu$) device where the change in behavior is even more striking. Recall that the β rolloff at low collector currents is due to the perimeter component (see Appendix K).

These observations can be qualitatively explained by referring to the SIS model structure depicted in Fig. 7.17a for thermal equilibrium. For the case where the insulating layer presents a low series resistance, i.e. the drop across the oxide layer ΔV_{ox} is much less than the junction voltage V_j , the barrier remains essentially unperturbed from its equilibrium shape (Fig. 7.17b). Therefore, the tunneling prob-

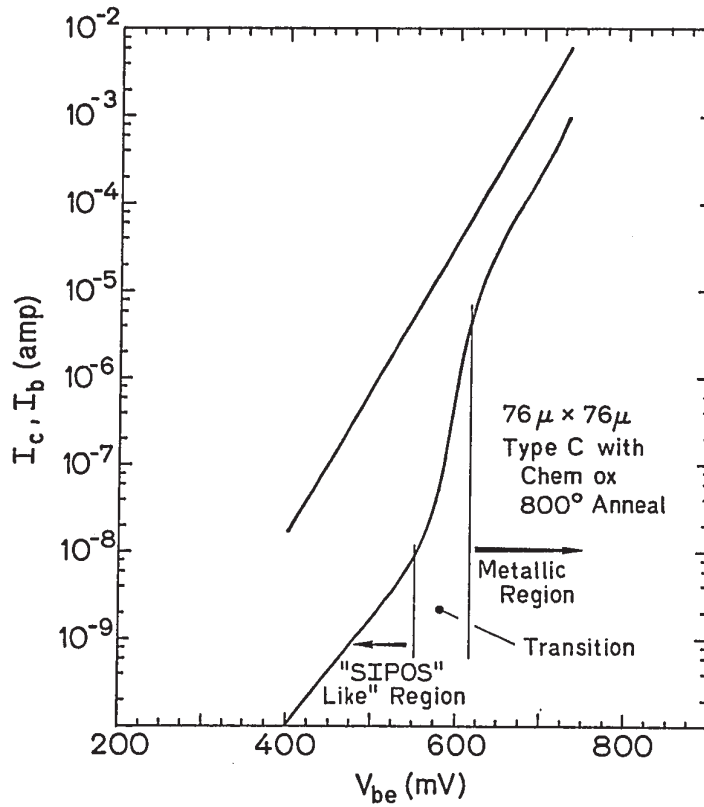


Figure 7.15: corrected Gummel plot for small SIPOS device w/ chemox

abilities T_e and T_h can be taken as constant which results in a bias independent J_{oe} . For the case depicted in Fig. 7.17c, the voltage drop across the oxide is substantial, and comparable to the junction voltage. This can result from either an excessively small value of T_n due to a thicker than optimal oxide, or due to a surface dopant density that is too low or a combination of the two (see eqn. 3.51). This voltage drop across the oxide can produce significant barrier modulation as indicated by the change in barrier shape from a rectangular to a trapezoidal shape as shown in Fig. 7.17(c). This has the effect of increasing both T_n and T_p ; the former results in a decrease in the effective series resistance (as has been experimentally

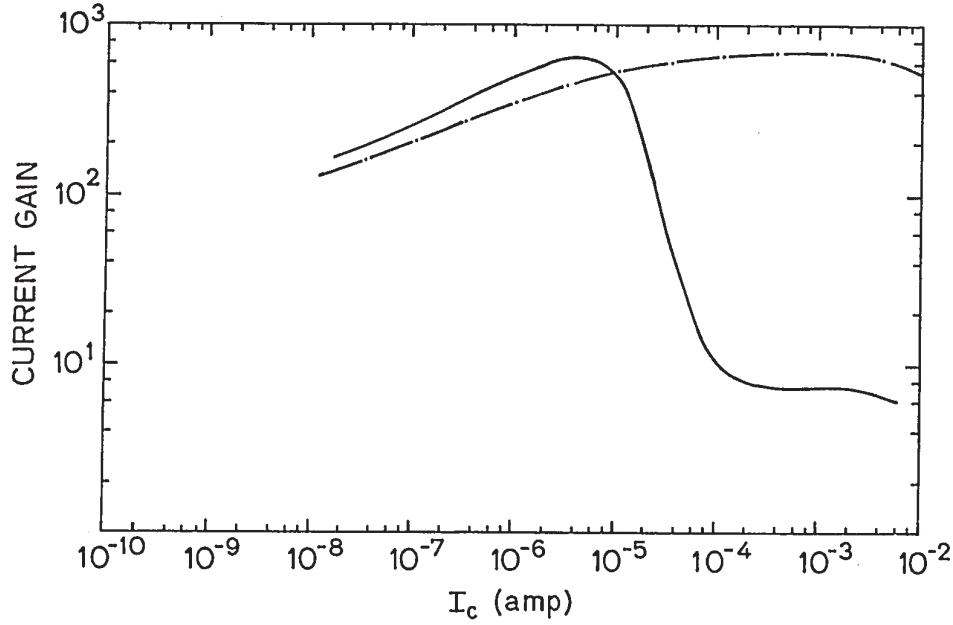


Figure 7.16: β vs. I_c for small SIPOS device w/ chemox

observed) but the latter has the undesired effect of making the interfacial oxide more transparent to holes. This compromises its role as a tunnel barrier. Holes that do tunnel are essentially hot holes in the SIPOS because of the discontinuity in the valence band introduced by the voltage drop across the oxide. These hot holes relax quickly and are not available to set up a countercurrent resulting from tunneling from the SIPOS into the silicon. The oxide/SIPOS combination then present a surface recombination velocity of

$$S = \frac{v_{th}T_h}{4}$$

and as $T_h \rightarrow 1$, S_c approaches the value for a metallic contact. Therefore, once the SIS structure develops significant barrier modulation, the hole barrier is attenuated accompanied by an increase in J_{oe} which can approach the metallic contact value if the effect is severe. This stresses the need for tight control over the oxide thickness as thicker oxides can lead not only to excessive series resistance, but also to compromise

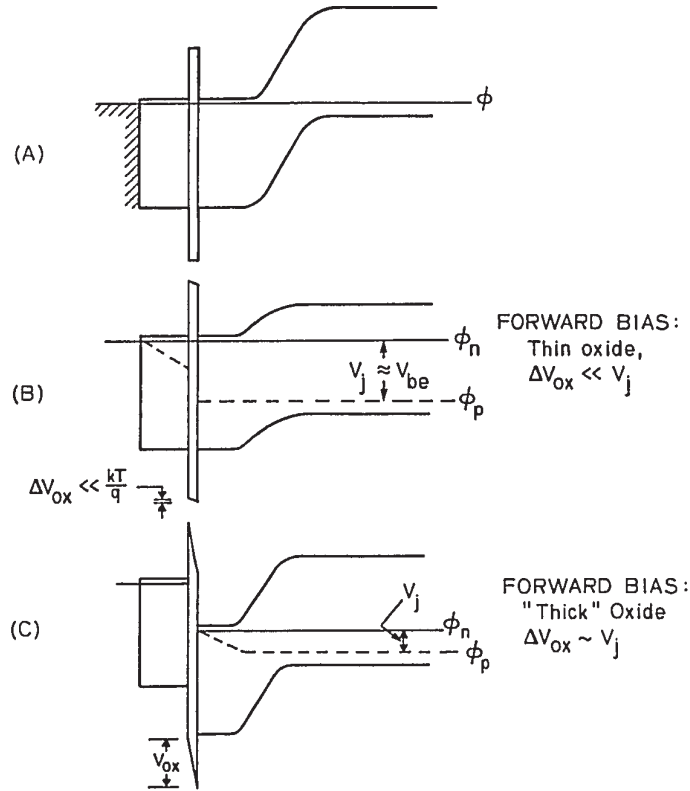


Figure 7.17: SIS structure in equilibrium (A), under bias with minimal oxide drop (B), effect of excessive oxide voltage drop (C)

of the oxide barrier to hole transport.

7.5. POLY Contacted Emitters

Since turning off the nitrous oxide during the SIPOS deposition did not result in the deposition of polysilicon, another set of devices was fabricated using the same process sequence but depositing POLY in an RF heated cold wall APCVD epi reactor. The gas flow rates used were $H_2 = 78 \text{ slm}$, $SiH_4 = 10^3 \text{ ppm}$, $PH_3/SiH_4 \approx$

T_A ($^{\circ}C$)	metal	$J_{oc}(10^{-14} A/cm^2)$		$G_b(10^{11} s/cm^4)$	
		no chem ox	with chem ox	no chem ox	with chem ox
Type B Device					
850 $^{\circ}C$	<i>Al</i>	56(2%)	5.4(5%)	1.8(1%)	1.75(1%)
	<i>Al + Pd</i>	—	—	—	—
900 $^{\circ}C$	<i>Al</i>	57(5%)	10.5	1.72(1%)	1.64
1000 $^{\circ}C$	<i>Al</i>	55	36(2%)	1.57	1.61(1%)
	<i>Al + Pd</i>	57	—	1.56	—
Type C Device					
850 $^{\circ}C$	<i>Al</i>	69(2%)	1.1	3.53(1%)	3.5
	<i>Al + Pd</i>	93(2%)	1.9(21%)	3.52(1%)	3.43(1%)
900 $^{\circ}C$	<i>Al</i>	82(9%)	3.9	3.46(1%)	3.36
1000 $^{\circ}C$	<i>Al</i>	54	37(3%)	2.84	2.84(2%)
	<i>Al + Pd</i>	65(1%)	51(3%)	2.81(1%)	2.82(1%)

Table 7.7: J_{oc} vs. T_A for POLY devices types B and C

2; the resulting thicknesses were 3100 Å, 3000 Å, 3600 Å for the devices annealed at 850 $^{\circ}C$, 900 $^{\circ}C$, and 1000 $^{\circ}C$ respectively. A summary of the J_{oc} results appears in Table 7.7. Yield problems were encountered with some of the *Al + Pd* devices; these are indicated by the missing entries. Note that the carrier gas used during the deposition is hydrogen which is purified of water vapor and oxygen by first passing through a hydrogen purifier. POLY films deposited under these conditions and annealed at 900 $^{\circ}C/15min$ yielded films of $1 \times 10^{-3} \Omega\text{-cm}$ resistivity, values quite similar to those of the $\gamma = 0$ SIPOS film. These films behaved as expected. They were inert in HF and when deposited on bare silicon, did not display the subtle off-color characteristic of the $\gamma = 0$ SIPOS films. This occurs for the SIPOS films presumably because of a slight mismatch in the refractive index between the

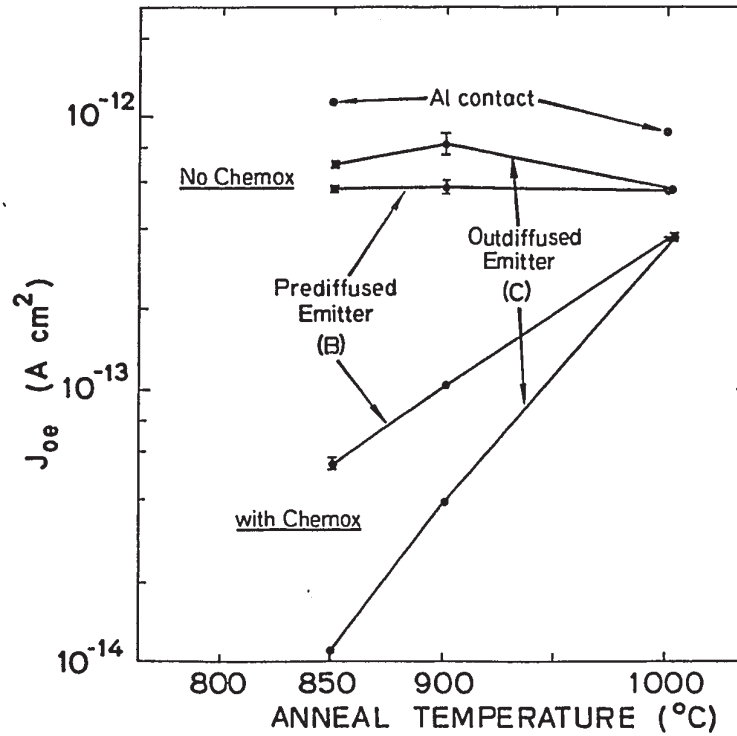


Figure 7.18: experimental J_{oe} vs. T_A for POLY contacted emitters

SIPOS and the underlying silicon, reflecting perhaps the presence of a small amount of oxygen in the films.

For this set of devices, half of each wafer received a Pd evaporation, nominally 350-400 Å in thickness; the other half was covered by a shadow mask. Sintering at 350°C/15min produced a Pd_2Si layer, thereby consuming a thickness of the POLY layer approximately equal to the Pd layer. This was followed by a standard Al flash evaporation and anneal at 350°C/15min. J_{oe} values for this set of devices is shown in Fig. 7.18.

7.5.1. Devices without a Chem Ox

These are the structures that were referred to in Chapter 3 as “intimate” POLY contacts. The native oxide seems to be reduced by exposure to hydrogen during the equilibration period before initiation of the POLY deposition (see Appendix L); whether this is simply an artifact of the ellipsometry technique is not clear. The J_{oe} for the 850°C type B device (EM-B + POLY composite) without the chem oxide and contacted by Al is $\approx 6 \times 10^{-13} A/cm^2$. This is a much higher value than for the corresponding SIPOS device and is consistent with a structure where the interface plays a negligible role in determining hole transport. Unfortunately, SRP’s are not available for these devices, so quantitative simulations will be limited to establishing the credibility of various scenarios. Knowledge of the profiles would be of limited value in any case as the minority carrier transport coefficients in heavily doped POLY are not known; however, a simple picture of the underlying physics is suggested by the J_{oe} results. Looking at the type C emitter at this anneal temperature, the J_{oe} for the Al and Al+Pd metallizations are $6.9 \times 10^{-13} A/cm^2$ and $9.3 \times 10^{-13} A/cm^2$ respectively. Since there is little outdiffusion at this temperature, the emitter is essentially entirely in the POLY; it is not surprising that the loss of approximately 400Å leads to an increase in J_{oe} . For the 900°C anneal, the Pd silicide devices experienced yield problems so only the Al contact data are available. J_{oe} for device type B remained unchanged. For the outdiffused emitter, J_{oe} is slightly higher than for the 850°C device. The increase is not large and is most likely attributable to variations in the thickness of the POLY layer due to poor uniformity control associated with the use of a small (7 cm × 23 cm) susceptor in the epi reactor. As the temperature is increased to 1000°C, the outdiffused dopant profile should dominate the recombination picture and device types should display similar J_{oe} . For device type C, J_{oe} decreases because the emitter has essentially widened and moved into the monocrystalline region. For this anneal temperature, the formation of silicide did not affect J_{oe} for device type B, but for the C type device J_{oe} it increased from $5.7 \times 10^{-13} A/cm^2$ to $6.5 \times 10^{-13} A/cm^2$ indicating that the emitter

is still somewhat transparent though not as much so as the 850°C device where the emitter width is much smaller.

7.5.2. Devices with Chem Ox

Leaving the chem oxide intact produces a drastic reduction in J_{oe} . This is quite at variance from the SIPOS case where J_{oe} at 850°C was left essentially unchanged by the introduction of the interfacial oxide. For the type B device, J_{oe} is again at the $5 \times 10^{-13} A/cm^2$ level indicating that, as in the SIPOS case, the surface recombination velocity presented by the chem ox/POLY layer is sufficiently low that recombination in the neutral emitter bulk dominates. The outdiffused device exhibits J_{oe} comparable to the lowest measured for the SIPOS case, although this lies in the regime where contributions from bulk base recombination can be significant. Changing the contact from Al to Al + Pd does not alter J_{oe} . This is consistent with the SIS model where the hole transport impedance is totally dominated by tunneling through the interfacial layer and although the reduction in overlying POLY thickness reduces the transport impedance of this region, it has negligible impact on the overall structure.

For the 900°C anneal, the Al contacted devices produced J_{oe} that were bracketed by the SIPOS devices with and without the chem oxide. The same observations made previously regarding the additivity of the increase in J_{oe} to both device types applies in this case as well. At 1000°C, the J_{oe} are equal for both profiles indicating that outdiffusion of dopant has dominated the recombination picture for both types of devices. J_{oe} increases slightly if the Al+Pd contact is used instead, from 3.7×10^{-13} to $5.0 \times 10^{-13} A/cm^2$. This is a particularly interesting observation as it indicates that J_{oe} is influenced by the loss of POLY thickness. If a valid trend, it indicates that the interfacial oxide at this point has lost much of its integrity, since it would otherwise tend to buffer J_{oe} from any such changes in the POLY layer.

7.6. Comparison of SIPOS and POLY Structures

The similarities in J_{oe} between SIPOS and POLY (with chem ox) contacted devices indicates that the same mechanism is dominating J_{oe} in both cases. For the low $800^{\circ}\text{C} - 850^{\circ}\text{C}$ anneals, the actual J_{oe} may be somewhat lower if the base contribution is discounted but it seems unlikely that the actual values could be much lower considering the constraints imposed by BGN and minority carrier lifetime. From a practical point of view, both structures are capable of producing J_{oe} of $1 \times 10^{-14} \text{ A/cm}^2$ and their relative merits must be established on other considerations. Increasing the anneal temperature above 900°C seriously erodes the performance advantage of these structures. It may be possible to bypass this restriction by using slower diffusing impurities such as As, but no such work has been reported in the literature. Both POLY and SIPOS devices are impacted, and to a similar extent, by this outdiffusion.

The presence of the chem oxide has been shown to be essential for the performance of the POLY SIS contact. Without this layer, the J_{oe} is essentially what would have been predicted for a thicker monocrystalline emitter with a comparable doping profile and slightly modified hole diffusion coefficient. Such a simulation for the 850°C type B device with a doping of $1 \times 10^{20} \text{ cm}^{-3}$ in the $3000 \text{ \AA}$ POLY layer yields a J_{oe} of $4.7 \times 10^{-13} \text{ A/cm}^2$ if a diffusion coefficient in the POLY is assumed to be one third that of comparably doped monocrystalline silicon. For the type C device, the predicted value is $6.6 \times 10^{-13} \text{ A/cm}^2$; reducing the POLY layer to $2500 \text{ \AA}$ to simulate the $\text{Al}+\text{Pd}$ device increases it slightly to $7.4 \times 10^{-13} \text{ A/cm}^2$. These results can be compared with the experimental values in Table 7.7. That fair agreement is obtained is somewhat fortuitous considering the rather crude nature of the assumptions made for the hole transport coefficients in the POLY layer.

For the SIPOS devices, the presence of the chem oxide does not appear to lower the J_{oe} except indirectly at higher temperatures where it serves as an additional diffusion barrier to limit the flux of outdiffusing phosphorous. It is possible that its introduction can lower J_{oe} below $1 \times 10^{-14} \text{ A/cm}^2$, and that may not be detected if

the “noise floor” established by bulk base recombination is of this order, but this is of interest only if other parasitic components of the base current can be reduced so as to make the value of J_{oe} at $1 \times 10^{-14} A/cm^2$ the dominant component of the base current.

7.7. Proposed Mechanism for SIPOS Emitter Operation

From these observations, the following model of SIPOS action is proposed. Referring to Fig. 7.19a, SIPOS in the as deposited form consists of amorphous silicon islands embedded in a silicon rich matrix. In a band diagram, the amorphous islands are shown to possess a larger gap than crystalline silicon. It is likely that this is a highly defective structure with a high density of gap states. Annealing promotes the crystallization of the $\alpha - Si$ islands with accompanying rejection of oxygen from the crystals to the surrounding matrix, thereby forming a multiphase structure of the shell morphology proposed by Ni [9]. This model has also been proposed by Goodman [10] based on optical and Raman backscattering experiments. The band gap of the small crystallites is shown as a fluctuating quantity depending upon the size of the grain. Possible mechanisms for this dependence include lattice dilation, quantum size effects, and failure to crystallize at a given annealing temperature for that particular grain size. The expulsion of oxygen outside the grains also causes the formation of an oxide layer at the Si -SIPOS interface. This would be apparent as an interfacial peak in the oxygen profile. This layer, formed as a byproduct of the SIPOS reconstruction, acts much like the intentionally introduced chem oxide in blocking hole injection into the SIPOS. There is also the possibility that a thin oxide layer is growing as a result of exposure to oxygen in the nitrogen ambient in several ppm during the equilibration period prior to SIPOS deposition. That this is at best only a partial factor is indicated by the fact that POLY contact devices where a thin oxide was “thermally” grown by deliberate exposure of the wafer at elevated temperatures ($25^\circ C \ll T \ll 650^\circ C$) to room air produced J_{oe} an order of magnitude greater than comparable SIPOS contacts. The exact magnitude of the

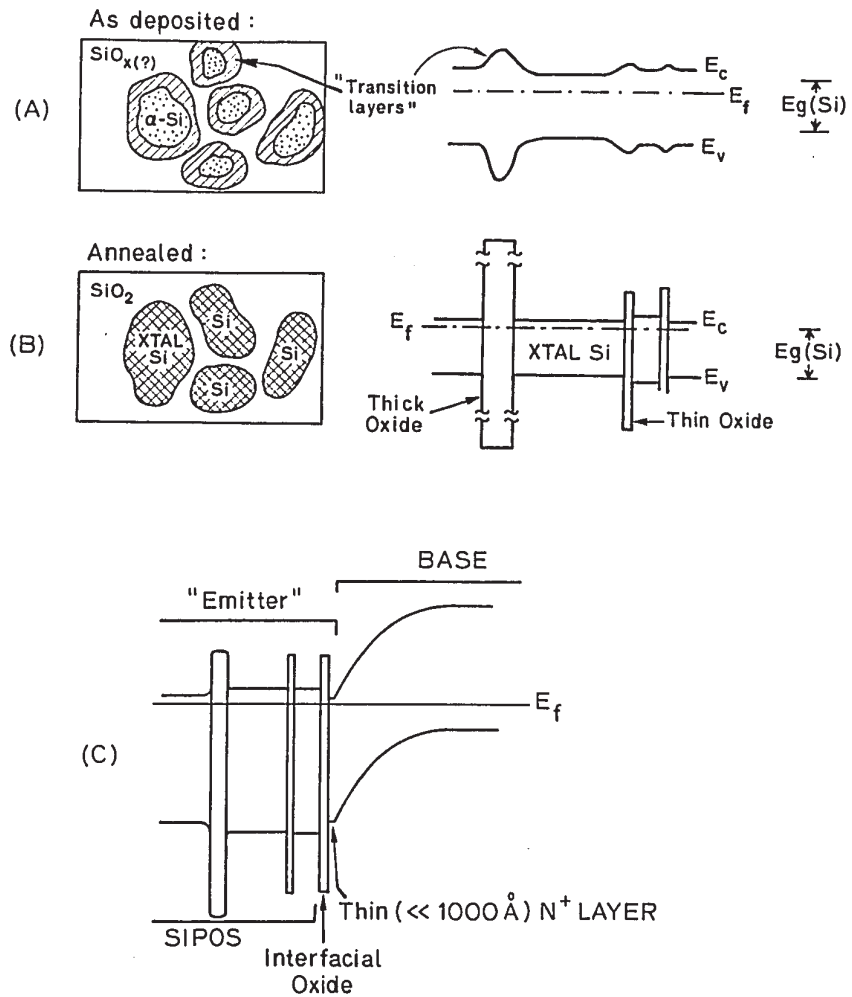


Figure 7.19: influence of annealing on SIPOS microstructure

wafer temperature was unknown during this oxide growth procedure which served primarily as a “worst case” exposure of wafers to ambient oxygen during the loading and equilibration process due to backstreaming of oxygen into the deposition ambient etc. In this scenario adding the chem oxide layer to the SIPOS contact is therefore seen to result in two layers of interfacial oxide, one deliberately grown, the other, a by-product of the SIPOS annealing process. This thicker composite oxide can explain the observed increase in series resistance (thicker tunnel barrier and less outdiffused impurity which lowers the surface concentration). The J_{oe} should not be impacted if the Si -chem ox and Si -(SIPOS formed)oxide are similar in terms of interfacial recombination activity.

The band diagram for a SIPOS contacted device should then resemble Fig. 7.19(c) where the key components include

- (1) thin tunnelling oxide at the Si-SIPOS interface
- (2) thin outdiffused layer, whose surface concentration should be in the $3 \times 10^{19} - 1 \times 10^{20}$ range and ideally only a few Debye lengths thick

The transport properties are then dictated by the interfacial oxide and to some extent by the near interfacial properties of the SIPOS. The bulk properties of SIPOS are felt to have little direct impact except in determining bulk majority carrier series resistance.

The POLY contacts can be understood as simply an extension of the monocrystalline emitter with slightly modified hole transport properties in the case of intimate POLY contacts and on the basis of the SIS structure if the chem oxide is left intact. In both the POLY SIS and SIPOS contacts, the effective surface recombination velocity is of the order of 200 cm/sec for devices annealed below 900°C . Above this temperature, there are indications that this layer may lose continuity with an attendant increase in the effective surface recombination velocity.

Since the J_{oe} is similar for both SIPOS and POLY SIS contacts, relative merits must be based on other considerations. Contact resistance, a key figure of merit for contact structures, was measured using a TLM structure [11] for a $\gamma = 0$ and $\gamma = 1.5$

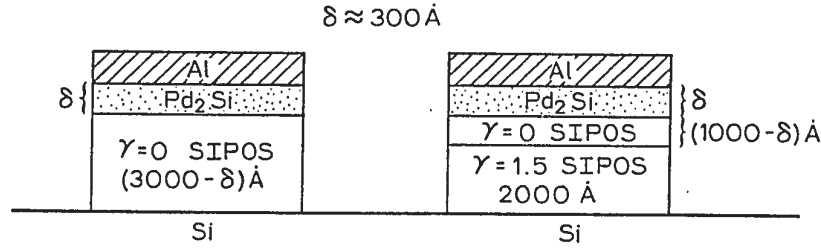


Figure 7.20: cross-section of SIPOS layers for contact resistance

SIPOS films with the structures shown in Fig. 7.20. The value of $R_c = 1 \times 10^{-4} \text{ ohm-cm}^2$ for the composite $\gamma = 1.5$ structure is clearly excessive for the large majority of high current density device applications. Even for the $\gamma = 0$ device, the value of $1.8 \times 10^{-6} \text{ ohm-cm}^2$ may need to be reduced by increasing the doping concentration in the SIPOS at the metal interface. Similar results are not available for the POLY contacts but an upper limit of $5 \times 10^{-4} \text{ ohm-cm}^2$ can be inferred from the device measurements although the actual value is probably much less. This paucity of data regarding the majority carrier series resistance is a major shortcoming that will need to be rectified before the relative merits of these contact structures can be established.

A major issue for the POLY SIS contact structure is the reproducibility and long term stability of the interfacial oxide. Control, especially over the thickness, is crucial to achieve a proper balance between the conflicting requirements between low majority carrier series resistance and low hole tunnelling probabilities. For SIPOS contacts, control over oxygen content is necessary to avoid excessive bulk resistivity; however, assuming that the proposed SIS structure for the SIPOS contact (without the chem ox) is correct, the role of bulk oxygen content in determining the thickness of the interfacial oxide layer is not known. In such a case, the kinetics governing the formation of this layer would certainly be a key area of investigation. If it proves to be a more stable and/or reproducible effect than the deliberate introduction of a

thin chemical or thermal interfacial oxide as in the case of the POLY SIS structure, SIPOS may possess some technological advantages.

7.8. Determining the Bandgap of SIPOS Using a C-V Technique

It is apparent that there are significant areas that have not been adequately addressed by the set of experiments described in this chapter. In closing, this section will describe a fairly straightforward technique that can be used to obtain a rough estimate of the bandgap in SIPOS. Determination of this parameter can be used as a plausibility check for the model of Matsushita [4] which relies on the SIPOS to present a barrier in the valence band to reverse injection of holes because of the wider bandgap of SIPOS. By establishing the equilibrium Fermi level separation between heavily doped *P* and *N* type SIPOS layers (and assuming that doping does not appreciably alter the density of states), an estimate of the bandgap can be obtained. This follows from the observation that in order to account for the low resistivity of heavily doped SIPOS, a relatively large and mobile carrier density is required. This in turn requires that the Fermi levels be close to the mobility edge. From elementary C-V theory [12] we have

$$V_{fb} = -\frac{1}{C_{ox}} (Q_f + Q_{it}(\phi_s = 0)) + \phi_{MS} \quad (7.2)$$

where V_{fb} is the flat band voltage, Q_f is the fixed charge and $Q_{it}(0)$ is the contribution to interfacial charge from fast states when the silicon is under flatband conditions, and ϕ_{MS} is the gate-substrate work function difference, or equivalently the separation between the Fermi levels in the gate and substrate materials. By fabricating *N* and *P* type SIPOS gates, and taking the difference of the V_{fb} , the difference in the Fermi levels between *N* and *P* type SIPOS can be determined if the quantities Q_f and $Q_{it}(0)$ can be characterized or assumed constant. Since it is usually more convenient to fabricate these on separate wafers, a reference standard is needed to keep track of the differences in Q_f and $Q_{it}(0)$ between wafers arising from variations in materials parameters as well as from different annealing steps.

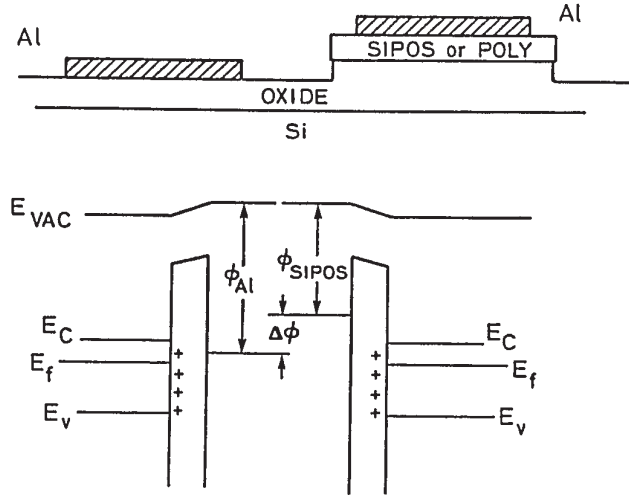


Figure 7.21: $C - V$ technique to determine SIPOS workfunction

By etching the SIPOS from selected regions, an *Al* gate capacitor can be placed adjacent to each SIPOS capacitor (Fig. 7.21) . Then we can define the quantity

$$\begin{aligned} \Delta V_{fb} &= V_{fb}(Al) - V_{fb}(SIPOS) \\ &\approx \left(\frac{1}{C_{ox,SIPOS}} - \frac{1}{C_{ox,Al}} \right) (Q_f + Q_{it}(\phi_s = 0)) + \Delta\phi_{MS} \end{aligned} \quad (7.3)$$

Since there is some controversy as to the procedure used to extract Q_f and $Q_{it}(0)$ from C-V measurements reflecting to some extent the uncertainty in the contributions of such factors as an orientation dependence of ϕ_{MS} , it is best to minimize the quantity

$$\left(\frac{1}{C_{ox,SIPOS}} - \frac{1}{C_{ox,Al}} \right)$$

by minimizing the extent of oxide undercut on the adjacent control *Al* gate capacitor. Taking the difference then between V_{fb} measured on a *P*-SIPOS gate wafer and an *N*-SIPOS gate wafer yields the desired quantity. If this Fermi level separation is not much greater than that obtained from analogous measurements of *P* and *N* POLY gate wafers, it seems unlikely that the wide gap picture of SIPOS is valid, which would in turn make the SIS picture a more likely alternative.

7.9. summary

This chapter has presented experimentally measured values of J_{oe} for SIPOS and POLY contact structures. The low values of J_{oe} achieved for both structures is in the neighborhood of $1 \times 10^{-14} \text{ A/cm}^2$ suggesting that similar limiting factors may be operative in both cases. SIPOS contacts appear to be able to attain these values without the deliberate introduction of an interfacial oxide. For POLY contacts, however, this layer is essential to attain low J_{oe} as in its absence, the POLY acts essentially as an extension of the monocrystalline region.

These experiments indicate that the SIS model is more relevant to SIPOS action than the heterojunction model proposed by Matsushita [4]. If this scenario is correct, then any technological advantages that SIPOS possesses over POLY SIS structures for selective contacts will rely on the mechanism by which this interfacial layer is formed. Further work needs to be done before the relative merits of these structures can be established.

Chapter 8

Conclusions

Potential enhancements in bipolar device performance extend beyond the straightforward applications illustrated in Chapter 2. The relaxation of design constraints that result from the reduction J_{oe} provides greater flexibility and may suggest new guidelines for device optimization. It has been demonstrated that homojunctions made in silicon are limited by materials parameters associated with high level doping effects to J_{oe} values in excess of $1 \times 10^{-13} A/cm^2$. Examination of the parameters on a phenomenological level indicate that several alternate structures can reduce J_{oe} below this “fundamental” limit; however, the SIS structure incorporating a thin interfacial silicon dioxide that both passivates the interface and imposes a transport barrier to holes appears to offer a solution that is compatible with standard process technology. Modeling of both static and dynamic behavior of this structure produced guidelines regarding the extent of outdiffusion that could be tolerated. It was seen that reduction of J_{oe} involved attaining a profile with sufficiently high surface concentration to achieve a low equilibrium surface hole density, yet shallow enough to minimize the impact of Auger recombination. High frequency performance which relies on the minimization of hole storage, was shown to favor the shallow structures as well. Reduction of the extent of the outdiffused layer consistent with the surface concentration requirement was shown to favor both static and dynamic performance.

Experimental BJT's incorporating POLY and SIPOS contacts supported the conclusions drawn in Chapter 3 regarding general process guidelines to achieve low

static values of J_{oe} . It was shown that for POLY emitters, that the presence of the interfacial oxide is mandatory for good performance. "Intimate" POLY contacts behaving essentially as extensions of the monocrystalline regions, do not seem to be able to offer much in the way of achieving low absolute values of J_{oe} although relative enhancements are readily attained. Although no attempt was made to understand hole transport mechanisms in heavily doped POLY, it is felt that they are understood albeit on an unsophisticated phenomenological level. The situation for SIPOS contacts is different. Low oxygen content films behaved surprisingly well. Although all indications point to an interfacial barrier action for this type of contact as well, the exact nature and even origin of this interfacial layer is poorly understood. In part, this is attributable to a lack of adequate control over the SIPOS deposition ambient.

Further investigation may serve to clarify some of these issues. Since the most interesting physics appears to lie in the low oxygen content SIPOS films, tighter control over residuals in the deposition ambient is necessary. A more sophisticated wafer loading system is needed to limit the growth of unintentional oxides, and special care must also be taken to avoid contamination by residual organics. The CV experiment suggested in Chapter 7 should shed some light over whether bulk or interfacial barrier action is responsible for the blocking of hole transport by N^+ -SIPOS contacts. If the interfacial oxide is found to result from the annealing of the SIPOS film itself, then the kinetics of this reaction need to be studied. The long term stability of these interfacial layers has not been established; work on possible failure mechanisms is needed. Characterization of majority carrier series resistance was only cursorily done, and only for the SIPOS contacts. This is a major drawback, as it makes evaluation of the relative merits of SIPOS and POLY contact structures difficult. This is certainly an area that needs extensive characterization before application of these structures to high speed or high current density applications can be considered practical.

The dynamic performance of these contact structures was not experimentally

investigated. Actual device geometries optimized to take advantage of the low J_{sc} need to be fabricated to see if the potential gains that were suggested can be realized in practice. For application to the point contact solar cell, the performance of the *P*-type counterpart must be established. Comparison to the *N*-type may also shed some light on the physics of these contacts.

Appendix A

Mechanical Analogy for the Solar Cell

A simple mechanical analogy of the point contact cell may be useful in elucidating the nature of the tradeoffs that are involved in the optimization process. Consider a pressure vessel (e.g. a steam boiler) as shown in Fig. A.1. The pressure of the working fluid inside the vessel is the analog of the Fermi level separation inside the generation region of the cell. To the vessel are attached several openings

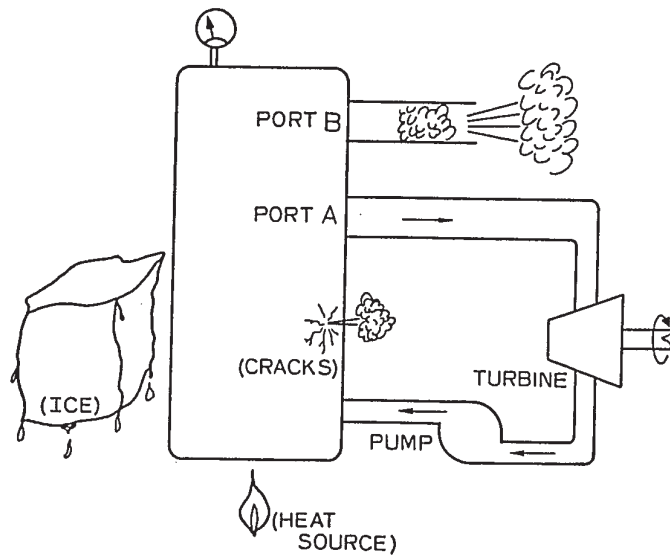


Figure A.1: mechanical analogy for a solar cell

or ports. Port A represents the means by which useful work can be extracted (to drive a turbine for example); the contacts would be the homologous structure in the cell. The cross sectional area of the mechanical port reflects the contact coverage fraction in the cell. For example, the small coverage fraction would be reflected by a small diameter port. In parallel with port A is port B which has the same cross sectional area; however this port empties to the atmosphere without doing any useful work. Unlike port A, this port is obstructed by a viscous plug which limits the “leak rate”. Port B represents the loss of energy to contact recombination. The “viscosity” of the plug is inversely related to the J_o of the contact. A zero value of J_o would be equivalent to closing the port altogether. For finite values, a larger value of J_o would imply that the plug would offer less resistance to fluid flow. The other opening is not really a port, but is drawn as cracks in the wall of the vessel. Loss of pressure due to lack of wall integrity is the analog of loss of carriers in the cell to “avoidable” recombination mechanisms such as bulk defect related recombination. As in the manufacture of pressure vessels where care is needed to minimize the presence of these leaks, so is care required in the fabrication of the cell to ensure that its properties such as bulk lifetime are not similarly degraded.

The driving force behind the mechanical engine is the constant influx of heat which is represented by the flame under the vessel. As is the case with the solar cell where the energy influx is fixed by the solar flux, the heat source is similarly constrained (i.e. there is no means for storage). This point illustrates that the figure of merit is not energy conversion efficiency, but rather power conversion efficiency. This is an important point for by operating both machines closer to equilibrium (i.e. a smaller energy influx) losses due to irreversible processes are reduced. However the energy influx is fixed and what is not converted is wasted.

The last feature of this mechanical analogy is represented by the large block of ice placed next to the vessel though not in direct contact. The vessel may transfer heat to it via several mechanisms such as radiation and convection. However, since it is not closely coupled to the ice, the vessel temperature is not held constant.

The major function for this low temperature heat reservoir is to prevent the vessel pressure (via the temperature) from becoming infinite. In the solar cell, there are similar restraining forces that keep the photogenerated carrier density finite. An example of such a dissipative process would be energy transfer from hot carriers that result from Auger recombination to the lattice which is effectively at a lower temperature. Radiative recombination exerts a similar restraining influence.

If it were possible to block all the ports and to seal the leaks in the vessel, the pressure and (for this fluid) temperature of the vessel would continue to rise until the transfer of heat to the ice block matched the energy influx from the flame. If the ice were not present, the temperature would rise until the flame temperature were reached and thermal equilibrium were attained. For the solar cell, the carrier concentration increases when the terminals are disconnected from the external circuit until the loss of carriers to dissipative mechanisms matches the photoproduction rate. The cell would eventually attain equilibrium when it reached the sun temperature (in principle, assuming they are the only objects in the universe), but since it is heat sinked, the lattice temperature is kept at a much lower value and an energy flux continues to flow through the cell. For a practical solar cell, it is not possible to close port B due to the presence of the contact diffusions, nor it is possible to seal all the leaks. The steady state pressure that results in the vessel with these losses is then a reflection of the open circuit voltage of the solar cell.

In order to obtain useful work from the vessel, port A must be opened. The working fluid is assumed to be a viscous fluid with the property that its viscosity decreases with pressure. A similar situation holds for the intrinsic base region of the cell which relies on conductivity modulation to minimize ohmic drops. The greater the carrier concentration, the greater the conductivity and the lesser the internal losses for a given current. As port A is opened, the fluid in the vessel begins to exit (but is replenished at the bottom) through this port. The driving force is a pressure gradient, so there will be some pressure drop and attendant loss of available energy. Note that the parasitic leakage through port B siphons off part of the available

fluid. Increasing the flow rate also increases the pressure drop. Since the available power is the product of the pressure and flow rate, it is clear that there is some intermediate flow rate at which maximum power will be developed. Flow rates above this optimum reduce the pressure excessively; flow rates less than this preserve the pressure, but the power is reduced because of the decreased flow rate. For the cell, exactly analogous statements can be made about the maximum power point (mpp). Currents above the mpp result in excessive depletion of carrier density which is reflected in the reduced terminal voltage. Currents below this value maintain a large cell voltage, but power is again reduced because of the smaller current.

Reducing the losses through contact recombination by reducing the coverage fraction would be analogous to reducing port B losses by reducing the cross sectional area (this also means a reduction of port A's cross sectional area since A and B model the same contact structure). The same problem arises in both cases because reducing the port area results in a constriction of flow. In order to maintain the same flow rate, the pressure drop must be increased; the net result is a loss of available power. A similar type of spreading resistance due to current crowding in the cell prevents the use of arbitrarily small contacts. Reducing pressure drops by using larger contacts (or ports) increases the parasitic losses. This tends to reduce the carrier concentration (or vessel pressure) and increases base region ohmic drops (internal viscous losses in the vessel since viscosity increases with decreasing pressure).

The resolution of this conflict lies in noting that J_o is the key parameter since it relates parasitic losses to pressure drops through the contact area. If J_o were very small, port B would be essentially plugged, and the port area could be made larger to minimize pressure drops without incurring a penalty in excessive parasitic losses. In the solar cell, development of contacts that block minority carrier transport, i.e. the development of low J_o contact structures would then permit the use of larger contacts to reduce current crowding effects without the penalty of losing too many carriers to contact recombination.

In this discussion, it has been inherently assumed that the loss through port B (i.e. contact recombination) are dominant. In the current generation of cells, this is the case; however, if contact recombination is significantly reduced, then the other loss mechanisms (i.e. the “cracks”) may become dominant. In order to achieve the highest cell efficiencies, it will be necessary to take all of these into account.

Appendix B

Bipolar Process Using Selective Emitter Contacts

What follows is a cross-sectional schematic representation of an eight mask level bipolar process using SIPOS (or POLY) emitters. There are two separate base implants; these provide both normal and superbeta transistors as well as two resistor ρ_s values. Only the vertical NPN's have been illustrated to simplify the diagram; however, lateral PNP, JFET, MOS capacitor, and fully isolated SIPOS (POLY) resistor structures are also available.

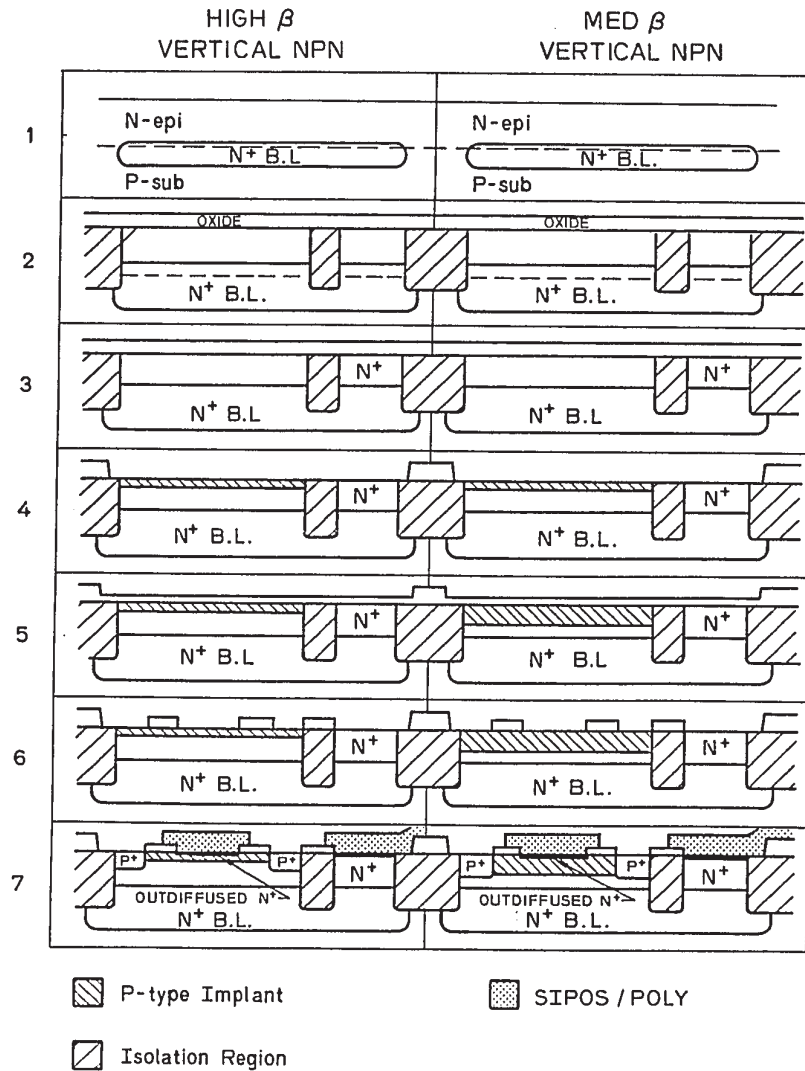


Figure B.1: bipolar process using selective emitter contacts

Appendix C

Proof of Ebers-Moll Reciprocity

C.1. Plausibility Argument

Before proceeding to a formal proof of reciprocity, a plausibility argument will be presented. Looking at the hole transport equations (3.13,14)

$$\vec{J}_p = \left(\frac{qn_{io}^2}{N_{eff}D_p} \right) \nabla u \quad (C.1)$$

$$\nabla \cdot \vec{J}_p = \frac{-qn_{io}^2}{N_{eff}\tau_p} u \quad (C.2)$$

If we generalize to finite frequencies, then

$$\nabla \cdot \vec{J}_p = -\frac{qn_{io}^2}{N_{eff}\tau_p} u - \frac{qn_{io}^2 j\omega}{N_{eff}} u \quad (C.3)$$

we can construct a one-dimensional finite element equivalent circuit.

$$J_p = \frac{qn_{io}^2}{N_{eff}D_p} \frac{1}{\Delta x} \Delta u \quad (C.4)$$

$$\Delta J_p = -\left(\frac{qn_{io}^2}{N_{eff}\tau_p} \Delta x \right) u - \left(\frac{qn_{io}^2}{N_{eff}\tau_p} j\omega \Delta x \right) u \quad (C.5)$$

By making u the analog of voltage and J the analog of current, these can be modeled by the elements shown in Fig. C.1 where

$$G_1 = \frac{qn_{io}^2 D_p}{N_{eff} \Delta x} \quad (C.6)$$

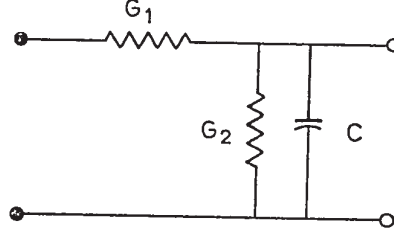


Figure C.1: lumped parameter equivalent circuit for transport equations

$$G_2 = \frac{qn_{i0}^2 \Delta x}{N_{eff} \tau_p} \quad (C.7)$$

$$C = \frac{qn_{i0}^2 \Delta x}{N_{eff}} \quad (C.8)$$

The diffused region can be modeled as a sequential chain of such elements and as x becomes small compared to the characteristic dimensions the approximation should accurately model the region. As long as $x > 0$, the equivalent circuit can be seen to be constructed from a finite number of linear time invariant resistors and capacitors. From linear network theory [1], any such network must be reciprocal if we assume the system begins in a zero initial state. This result can be extended to three dimensions by constructing a three dimensional grid of equivalent conductances and capacitances.

C.2. Formal Proof

Consider a semiconductor region with contacts C_1 and C_2 positioned on the surface of the semiconductor as shown in Fig. C.2 ; C_1 and C_2 represent equipotential regions. Although not essential for the proof, we shall assume for convenience that the region is of finite extent. The two port equivalent circuit is shown in Fig. C.3 . In order to prove reciprocity, we must show that

$$\frac{I_a(C_1)}{u_a(C_2)} = \frac{I_b(C_2)}{u_b(C_1)} \quad (C.9)$$

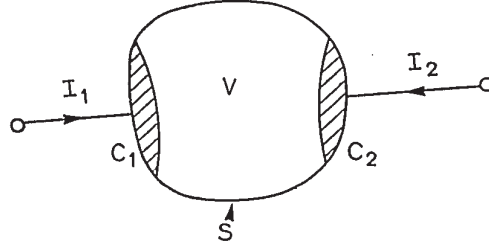


Figure C.2: control volume used for reciprocity proof

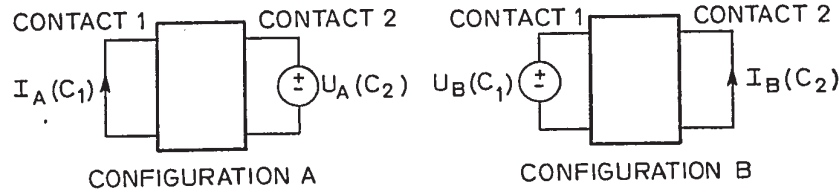


Figure C.3: equivalent circuit for control volume of Fig. C.2

where $u_a(C_1)$ represents the value of u at contact C_1 in configuration a, and $I_a(C_1)$ represents the resulting terminal current flowing into the contact C_1 in the same configuration. u_b and I_b refer to the analogous quantities in configuration b where the roles of contacts C_1 and C_2 are reversed.

Consider the quantity

$$\nabla \cdot (u_a \vec{J}_b) - \nabla \cdot (u_b \vec{J}_a)$$

where the a and b subscripts refer to the corresponding configurations and where $\vec{J}$ represents the hole current density. Note that u is a scalar function while $\vec{J}$ is a vector function; they are both complex functions of space and frequency. Integrating over the volume V shown in Fig. C.2, we get

$$\int_V dv \left[\nabla \cdot (u_a \vec{J}_b) - \nabla \cdot (u_b \vec{J}_a) \right] = \oint_S \hat{n} \cdot dS (u_a \vec{J}_b - u_b \vec{J}_a) \quad (C.10)$$

where the divergence theorem has been invoked. Breaking down the surface integral

into its constituent terms

$$\oint \hat{n} dS \rightarrow \int_{C_1} \hat{n} dS + \int_{C_2} \hat{n} dS + \int_{rest} \hat{n} dS$$

For the $\int_{C_1}$ term

$$\int_{C_1} \hat{n} \cdot (u_a \vec{J}_b - u_b \vec{J}_a) dS = u_a(C_1) \int_{C_1} \hat{n} \cdot \vec{J}_b dS - u_b(C_1) \int_{C_1} \hat{n} \cdot \vec{J}_a dS \quad (C.11)$$

In configuration A, $u_a(C_1) = 0$,

$$\int_{C_1} \hat{n} \cdot (u_a \vec{J}_b - u_b \vec{J}_a) dS = u_b(C_1) I_a(C_1) \quad (C.12)$$

Similarly for $\int_{C_2}$

$$\int_{C_2} \hat{n} \cdot (u_a \vec{J}_b - u_b \vec{J}_a) = -u_a(C_2) I_b(C_1) \quad (C.13)$$

If we draw the surface S outside the semiconductor, then $J_a = J_b = 0$. This presents no problem since we can always regard vacuum as a semiconductor with zero pn product. Therefore, the integral over the rest of the surface is

$$\int \hat{n} \cdot dS (u_a \vec{J}_b - u_b \vec{J}_a) = 0 \quad (C.14)$$

so that we can write for the surface integral

$$\oint \hat{n} \cdot dS (u_a \vec{J}_b - u_b \vec{J}_a) = u_b(C_1) I_a(C_1) - u_a(C_2) I_b(C_2) \quad (C.15)$$

If we now consider the volume integral,

$$\begin{aligned} \int dv [\nabla \cdot (u_a \vec{J}_b) - \nabla \cdot (u_b \vec{J}_a)] = \\ \int dv (u_a \nabla \cdot \vec{J}_b + \vec{J}_b \cdot \nabla u_a - \vec{J}_a \cdot \nabla u_b - u_b \nabla \cdot \vec{J}_a) \end{aligned} \quad (C.16)$$

and referring to eqns. (C.1,3), we can define scalar functions α and θ

$$\begin{aligned} \vec{J}_p &= \left(\frac{qn_{io}^2}{N_{eff}/D_p} \right) \nabla u \equiv \alpha \nabla u \\ \nabla \cdot \vec{J}_p &= \left(-\frac{qn_{io}^2}{N_{eff}\tau_p} - \frac{qn_{io}^2}{N_{eff}} j\omega \right) u \equiv \theta u \end{aligned}$$

and substituting for $\vec{J}$, we get

$$\begin{aligned}
& \int dv [\nabla \cdot (u_a \vec{J}_b) - \nabla \cdot (u_b \vec{J}_a)] = \\
& \int dv (u_a \nabla \cdot \vec{J}_b - u_b \nabla \cdot \vec{J}_a + \alpha \nabla u_b \cdot \nabla u_a - \alpha \nabla u_a \cdot \nabla u_b) \\
& = \int dv (u_a \nabla \cdot \vec{J}_b - u_b \nabla \cdot \vec{J}_a)
\end{aligned} \tag{C.17}$$

Further substituting for $\nabla \cdot \vec{J}$

$$= \int dv (u_a \theta u_b - u_b \theta u_a) \tag{C.18}$$

So we can write the surface integral as

$$\oint_S \hat{n} dS \cdot (u_a \vec{J}_b - u_b \vec{J}_a) = \int dv [\nabla \cdot (u_a \vec{J}_b) - \nabla \cdot (u_b \vec{J}_a)] = 0 \tag{C.19}$$

since θ and α are scalar functions. Therefore,

$$\oint_S \hat{n} dS \cdot (u_a \vec{J}_b - u_b \vec{J}_a) = \int dv [\nabla \cdot (u_a \vec{J}_b) - \nabla \cdot (u_b \vec{J}_a)] \tag{C.20}$$

reduces to

$$u_b(C_1)I_a(C_1) - u_a(C_2)I_b(C_2) = 0 \tag{C.21}$$

$$\frac{I_a(C_1)}{u_a(C_2)} = \frac{I_b(C_2)}{u_b(C_1)}$$

Q.E.D.

This result implies for example that Ebers Moll reciprocity holds for the base region of a BJT for an arbitrary doping profile and for arbitrary frequencies. It also implies that the representation of the diffused region by a “T” equivalent network is valid.

Note: This theorem has been proved by W. Shockley in [2].

Appendix D

Model Parameters Used in Simulations

The following plots are graphical representations of the model parameters used in the numerical simulations of Chapters 3,4, and 7. The minority carrier hole diffusion coefficient is assumed to be the same as when holes are majority carriers. The formula, derived from [1] and plotted in Fig. D.1, is

$$D_p = \frac{kT}{q} \cdot \left(54.3 + \frac{406.97}{1 + 3.7447 \times 10^{-18} N} \right) \quad (D.1)$$

The band gap narrowing model is taken from del Alamo [2] for $N > 5 \times 10^{18} \text{ cm}^{-3}$; for N below this value, a smooth fit is made to Slotboom's model [3] since the model

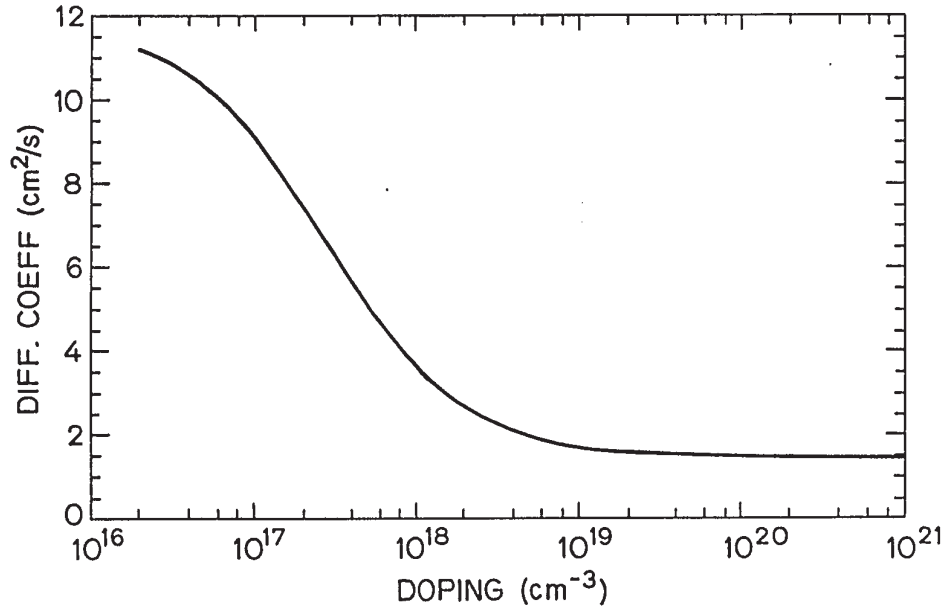


Figure D.1: model used for hole diff. coefficient vs. dopant density

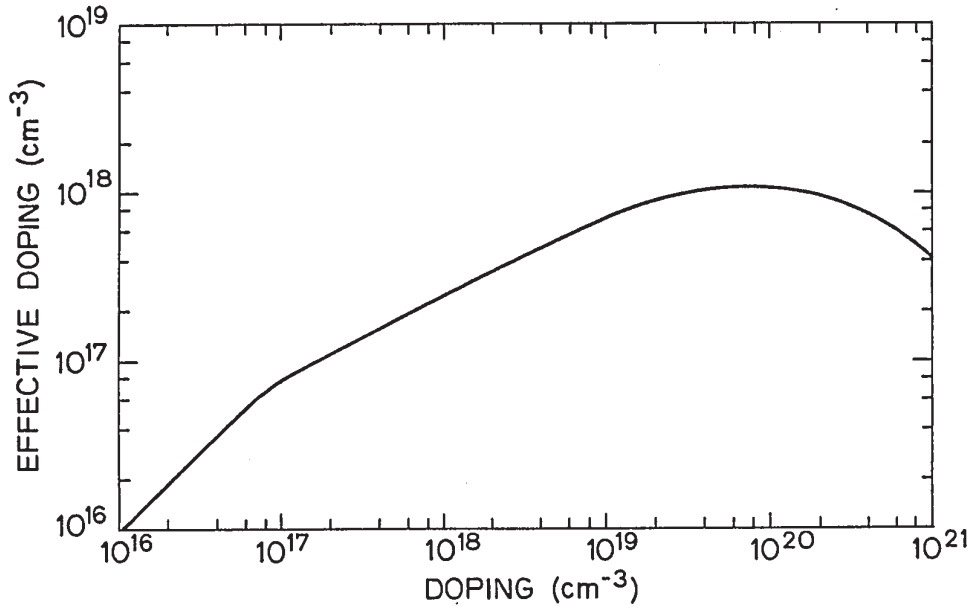


Figure D.2: band gap narrowing model used in simulations

given in [2] is not intended for use at the lower doping levels. The resulting value for the “effective doping density” N_{eff} given by

$$N_{eff} = N \cdot e^{-\Delta E_g/kT} \quad (D.2)$$

, where ΔE_g is the effective band gap shrinkage obtained from the band gap narrowing model, is plotted in Fig. D.2 . N_{eff} saturates at a value $\approx 1 \times 10^{18}$ for $N \approx 8 \times 10^{19}$ corresponding to the maximum hole barrier that can be attained in monocrystalline *Si* through doping. The hole lifetime, modeled by the equation

$$\tau_p = [C_{auger} N^2 + C_{srh} N]^{-1} \quad (D.3)$$

is displayed in Fig. D.3 where the values of $C_{auger} = 5 \times 10^{-32}$ and $C_{srh} = 1.67 \times 10^{-12}$ found in [2] have been used. Below $N = 4 \times 10^{19}$, the SRH term dominates; above this value, Auger effects are dominant.

The saturation current for a uniformly doped semi-infinite monocrystalline *Si* region using these model parameters is displayed in Fig. D.4 . It indicates a mini-

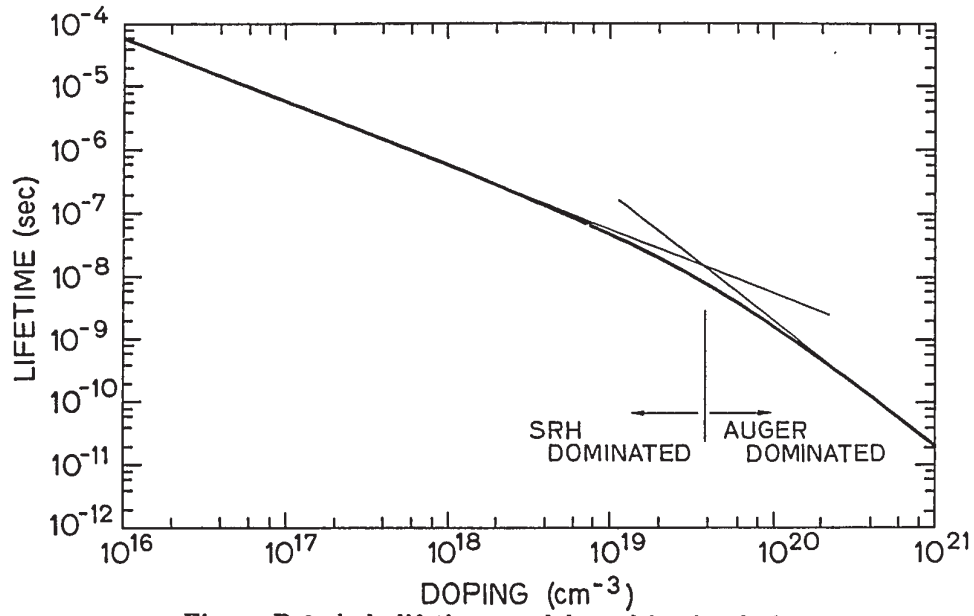


Figure D.3: hole lifetime model used in simulations

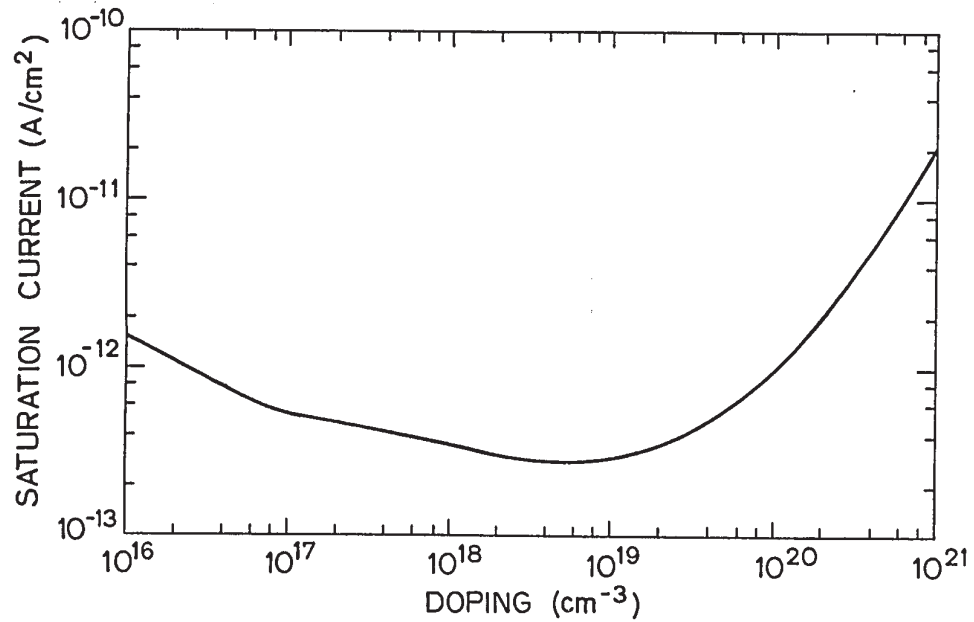


Figure D.4: J_{oe} vs. N for semi-infinite uniform region using model parameters

mum value for J_o of $3 \times 10^{-13} A/cm^2$ is achieved at $N \approx 1 \times 10^{19} cm^{-3}$. Below this value, the hole density increases because of the fall off in N_{eff} ; above this value the concerted effects of Auger recombination and the fall off in N_{eff} predicted by the band gap narrowing model produce a rapid increase in J_o .

Appendix E

Refinement of the Charge Control Result

In Chapter 4, it was shown that for the transparent case with $S_c = \infty$, that

$$J_p(0) = \frac{Q_p}{\tau_F} \left(1 + j\omega \frac{W^2}{2D_p} \right) \quad (E.1)$$

whereas the exact solution gives, when expanded to second order in the perturbation parameter W/L_p^* ,

$$J_p(0) = \frac{Q_p}{\tau_F} \left(1 + j\omega \frac{W^2}{3D_p} \right) \quad (E.2)$$

The seeds of this discrepancy lie in the cavalier assumptions made in the charge control derivation. In eqn. (4.20), it was assumed that the transport factor was identically unity and that the hole distribution is satisfactorily modeled by a linear distribution for the condition

$$\omega \frac{W^2}{D_p} \ll 1.$$

That both assumptions lack validity will become apparent in the following analysis.

Consider the region depicted in Fig. 4.1. We can write

$$J_p(W) = \alpha_T J_p(0) \quad (E.3)$$

where α_T is the hole transport factor. Charge control says

$$J_{p,net} = J_p(0) - J_p(W) = \frac{dQ_p}{dt} + \frac{Q_p}{\tau_p} \quad (E.4)$$

Using the exact solution for $J_p(0)$ from eqn. (4.34) and noting that

$$J_p(W) = \frac{2qD_p p_o}{L_p^o} \sqrt{1 + j\omega\tau_p} \left[\frac{e^{-\lambda}}{1 - e^{-2\lambda}} \right] \quad (E.5)$$

where $\lambda \equiv W/L_p^*$. Solving for

$$\begin{aligned}\alpha_T &\equiv J_p(W)/J_p(0) \\ &= \left(\frac{2e^{-\lambda}}{1 + e^{-2W/L_p^*}} \right)\end{aligned}\quad (E.6)$$

Substitution into eqn. (A.4) gives

$$\begin{aligned}J_p &= \frac{1}{1 - \alpha_T} \left[j\omega + \frac{1}{\tau_p} \right] Q_p \\ &= \frac{2}{\lambda^2} \left(1 + \frac{5}{12}\lambda^2 \right) \left[j\omega + \frac{1}{\tau_p} \right] Q_p\end{aligned}\quad (E.7)$$

The integrated hole charge is given by

$$Q_p = q \int_0^W \frac{p_o}{1 - e^{-2\lambda}} \left[e^{-x/L_p^*} - e^{-2\lambda} e^{x/L_p^*} \right] dx \quad (E.8)$$

To order λ^2 , this becomes

$$Q_p = \frac{qp_o W}{2} \left[1 - \frac{\lambda^2}{12} \right] \quad (E.9)$$

Incorporation into eqn. (A.7) gives

$$J_p(0) = \frac{2}{\lambda^2} \left(1 + \frac{5}{12}\lambda^2 \right) \left[\frac{qp_o W}{2} \left(1 - \frac{\lambda^2}{12} \right) \right] \left(j\omega + \frac{1}{\tau_p} \right) \quad (E.10)$$

Substituting in for λ and simplifying gives

$$J_p(0) = Q_p \left(\frac{2D_p}{W^2} \right) \left(1 + j\omega \frac{5}{12} \frac{W^2}{D_p} - j\omega \frac{1}{12} \frac{W^2}{D_p} \right) \quad (E.11)$$

The $5/12$ term comes from the nonunity transport factor. Compared to the original charge control result, this is a factor of $W^2/12D_p$ smaller. The use of a more refined hole distribution gives an additional correction of $W^2/12D_p$. The refined charge control result, to order λ^2 becomes

$$\begin{aligned}J_p(0) &= Q_p \left(\frac{2D_p}{W^2} \right) \left(1 + \frac{4}{12} j\omega \frac{W^2}{D} \right) \\ &= \frac{Q_p}{\tau_F} \left(1 + j\omega \frac{W^2}{3D_p} \right)\end{aligned}\quad (E.12)$$

which is the same result given by the exact solution when expanded to the same order in λ . Note that both “corrections” contribute equally to the result.

Appendix F

Preheater Calculations

Consider the control volume illustrated in Fig. F.1. If we assign an effective gas stream temperature $T_g(x)$ and a tube wall temperature $T_w(x)$, we can write the energy balance equation as:

$$\begin{aligned} dQ(x) &= p dx \hbar_c (T_w - T_g) \\ &= \dot{m} c_p (T_g) [T_g(x + dx) - T_g(x)] \end{aligned} \quad (F.1)$$

where $\hbar_c$ is an effective convective heat transfer coefficient, and p is the wetted perimeter of the tube. The mass flow rate can be written as

$$\dot{m} = \frac{dm}{dV} \frac{dV}{dt} = \rho(T_g(x)) \frac{dV(x)}{dt} = \text{constant} \quad (F.2)$$

where V is the gas volume and ρ is the mass density which depends upon the local gas stream temperature. For simplicity, assume

$$c_p = \text{constant} = c_p(\text{eff. avg. temp.})$$

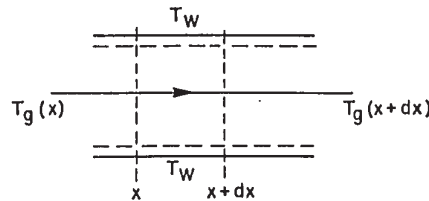


Figure F.1: control volume for heat transfer equation

Then (C.1) becomes

$$p\hbar_c(x) [T_w(x) - T_g(x)] = \dot{m}c_p \frac{dT_g}{dx} \quad (F.3)$$

or rearrranging

$$\frac{dT_g}{dx} + \frac{p\hbar_c(x)}{\dot{m}c_p} T_g = \frac{p\hbar_c(x)}{\dot{m}c_p} T_w(x) \quad (F.4)$$

Let

$$\alpha(x) \equiv \frac{p}{\dot{m}c_p} \hbar_c(x)$$

then

$$\frac{dT_g}{dx} + \alpha(x) T_g = \alpha(x) T_w \quad (F.5)$$

Since the assumption of $T_w = \text{constant}$ is probably untenable, let

$$T_w = \theta \cdot (T_m - T_g)$$

where T_m is the temperature of the heat source and θ depends only on the mass flow rate. For low flow rates, it should approach 1; for high flow rates, it may assume a value closer to 0. Substituting into (6) gives

$$\frac{dT_g}{dx} + \alpha(x)\theta T_g = \alpha(x)\theta T_m \quad (F.6)$$

Looking at $\hbar_c(x)$ in equation (5), recall that it can be written as [1]

$$\hbar_c = k_f / \delta$$

where k_f is the thermal conductivity and δ is the effective boundary layer thickness. The data for nitrogen from the gas tables [2] can be fitted by

$$k_f = k_{fo} \cdot (T/273); \quad k_{fo} = .013^\circ K^{-1}$$

where T is in Kelvin. This gives an error of approximately 20% at the higher temperatures (but heat transfer is diminished at the higher temperatures anyway). Since the Reynolds number at the anticipated flow rates indicate the flow should be

turbulent, we shall assume that the effective boundary layer remains constant once the flow is fully developed. Therefore

$$\begin{aligned}\hbar_c(x) &= \frac{k_{fo}}{\delta(\dot{m})} \left(\frac{T}{273} \right) \\ &= \hbar_{co}(\dot{m}) T_g\end{aligned}\tag{F.7}$$

where $\hbar_{co}$ will be an experimentally determined parameter. Since

$$\alpha(x) = \alpha_o(\dot{m}) T_g\tag{F.8}$$

where

$$\alpha_o(\dot{m}) = \frac{p}{\dot{m} c_p} \hbar_{co}(\dot{m})$$

we can write

$$\frac{dT_g}{dx} + \alpha_o(\dot{m}) \theta T_g^2 = \alpha_o(\dot{m}) \theta T_m T_g\tag{F.9}$$

Absorbing θ into α_o by defining

$$\alpha'_o \equiv \theta \alpha_o$$

eqn. (C.9) becomes

$$\frac{dT_g}{dx} + \alpha'_o(\dot{m}) T_g^2 = \alpha'_o(\dot{m}) T_m T_g\tag{F.10}$$

rearranging, we get

$$dT_g = \alpha'_o(\dot{m}) [T_m T_g - T_g^2] dx\tag{F.11}$$

Integrating both sides,

$$\int_{T_o}^T \frac{dT_g}{T_m T_g - T_g^2} = \int_0^x \alpha'_o(\dot{m}) dx\tag{F.12}$$

and solving, we get

$$T_g(x) = T_m \left[1 - \left(1 - \frac{T_m}{T_0} \right) e^{\alpha'_o(\dot{m}) T_m x} \right]^{-1}\tag{F.13}$$

where T_0 is the inlet temperature. At $x = L$, where L is the total active preheater length,

$$T(x = L) = T_m \left[1 - \left(1 - \frac{T_m}{T_0} \right) e^{-\alpha'_o(\dot{m}) T_m L} \right]^{-1}\tag{F.14}$$

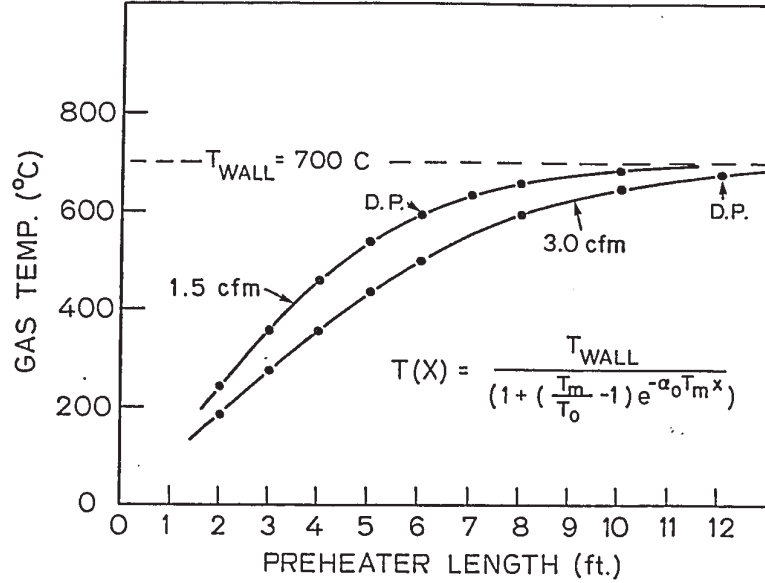


Figure F.2: gas temperature profile for N_2 preheater

We are then left with 2 parameters T_m and α'_0 to be experimentally determined. Using a 6 foot test length of straight 10mm ID quartz tubing for various flow rates yielded

$$\alpha'_0 = 5.1 \times 10^{-4}$$

for a mass flow rate of 1.5 scfm and

$$T_m = 710^\circ C.$$

Using this model, the preheater performance for a twin channel parallel configuration and a single channel serial geometry were compared. Although the exit temperature for the parallel geometry is 90% that of the serial, but the lower pressure drop makes this the preferred structure. A plot of average gas stream temperature vs. distance along the preheater tube for both configurations is shown in Fig. C.2. For the parallel geometry, the exit temperature at the design point gas

flow rate of 1.5 scfm is 600°C , a 100°C differential from the 700°C effective heat source temperature.

Appendix G

Conductivity Decay Technique for J_{oe}

Consider the apparatus depicted in Fig. G.1. By photogenerating carriers with a short duration flash lamp, it is possible to conductivity modulate the silicon wafer if the doping level is not too high. By monitoring the temporal decay of the photo-enhanced conductivity, it is possible to deduce the carrier lifetime. In this case, the change in wafer conductivity is sensed as an imbalance of the bridge which is monitored by an oscilloscope. This technique is regularly used by the Thermo-Photovoltaic Group at Stanford University as a process monitoring tool. If the additional requirements of sufficiently long minority carrier lifetime and low level injection are also met, a convenient method of determining J_{oe} suggests itself [1].

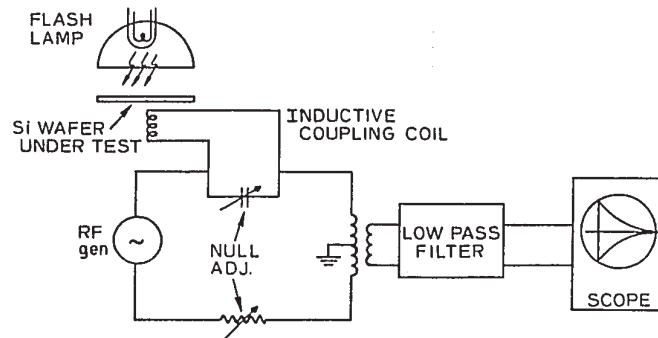


Figure G.1: apparatus used to determine J_{oe} from conductivity decay lifetime measurements

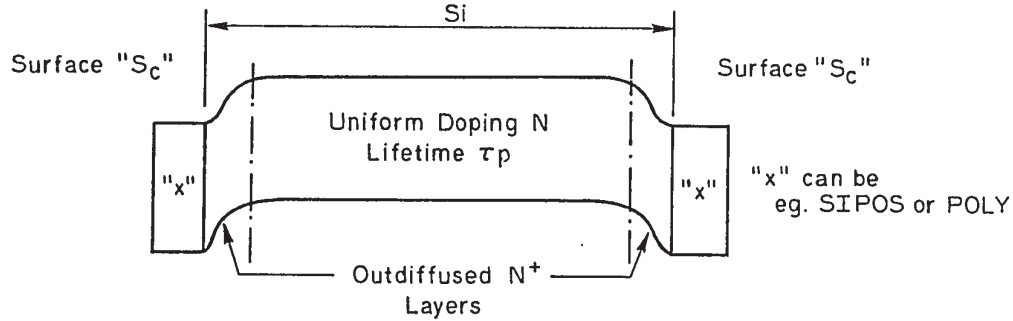


Figure G.2: band diagram of lifetime test wafer

Referring to the band diagram and wafer cross-section in Fig. G.2 where for concreteness we consider N type material, we can write the hole conservation equation as

$$\frac{1}{q} \oint_S \vec{J}_p \cdot \hat{n} dS = -\frac{\partial}{\partial t} \int_0^W p dx - \int_0^W \frac{p}{\tau_p} dx \quad (G.1)$$

where W is the wafer thickness and τ_p is the bulk hole lifetime. After the flash ends at time $t = 0+$, the photogenerated carriers rapidly diffuse across the wafer within a transit time given by

$$t_{transit} = \frac{W^2}{2D_p}$$

assuming of course that the hole lifetime satisfies $\tau_p > t_{transit}$. If this condition is met, then the hole distribution in the bulk can be assumed to have equilibrated and we can write equation (G.1) as

$$\frac{1}{q} \oint_S \vec{J}_p \cdot \hat{n} dS = W \frac{dp}{dt} - \frac{W}{\tau_p} p \quad (G.2)$$

Now, the surface terms can be written as

$$\frac{1}{q} \oint_S \vec{J}_p \cdot \hat{n} dS = 2sp \quad (G.3)$$

where s is the effective surface recombination velocity presented to holes in the bulk by the $N+$ surface regions. This can be related to J_{oe} , the hole saturation current

for the surface regions by

$$s = \frac{J_{oe}N}{qn_{io}^2} \quad (G.4)$$

where N is the bulk doping density. Then equation (G.2) becomes

$$-\frac{dp}{dt} = \frac{p}{\tau_p} + p \frac{2NJ_{oe}}{Wqn_{io}^2} \quad (G.5)$$

The conductivity decay remains exponential as expected since all recombination has been assumed linear in the hole density and the decay time constant is given by

$$\frac{1}{\tau_{decay}} = \frac{1}{\tau_p} + \frac{2NJ_{oe}}{Wqn_{io}^2} \quad (G.6)$$

Solving for J_{oe} , we have

$$J_{oe} = \frac{qn_{io}^2}{2N} W \left[\frac{1}{\tau_{decay}} - \frac{1}{\tau_p} \right] \quad (G.7)$$

This relation sets limitations on the sensitivity of this technique. Since wafer thicknesses are limited to $50 - 100\mu$ from mechanical considerations, in order to measure J_{oe} on the order of $1 \times 10^{-14} A/cm^2$, it is necessary to have

$$N\tau_p > 1 \times 10^{15} W$$

For a wafer with $3 \times 10^{16} cm^{-3}$ doping density, this requires that the bulk lifetime exceed $350\mu sec$ for a 100μ wafer. For low J_{oe} wafers, it is possible that space charge recombination constitutes a significant component. In this case, equation (G.5) becomes

$$-\frac{dp}{dt} = \frac{p}{\tau_p} + \left(\frac{2K}{qW} \sqrt{\frac{N}{n_{io}^2}} \right) \sqrt{p} \quad (G.8)$$

where K is a constant. If the space charge term dominates, this decays as

$$p(t) = \left[p(t=0) - \frac{K}{qW} \sqrt{\frac{N}{n_{io}^2}} t \right]^2 \quad (G.9)$$

This nonexponential decay signals that other recombination mechanisms are dominant.

This technique for J_{oe} extraction imposes very little processing overhead, but it presupposes the availability of both high lifetime starting material, (in this case float zone which is difficult to obtain in higher doping densities) and the availability of an ultraclean processing facility that is capable of maintaining lifetimes in the $500\mu sec$ range.

Appendix H

Perimeter Streamline Patterns

The one dimensional approximation breaks down for the perimeter regions in the BJT. There are two main cases of interest where an estimate of the lateral spread of current flow is necessary to extract device parameters from experimental measurements. In the epi base region, electrons are injected laterally from the emitter sidewall into the extrinsic base as well as vertically into the intrinsic base region. As can be seen from Fig. H.1a, this lateral injection increases the effective injecting area of the emitter. To see if this is a major factor, ΔW should be evaluated. If $\Delta W \ll W$, where W is the lateral dimension of the emitter diffusion, then this effect can be neglected.

In the emitter region itself, a similar effect occurs. Unlike outdiffused POLY and SIPOS contacts which are self-aligned, the diffusion and contact areas differ

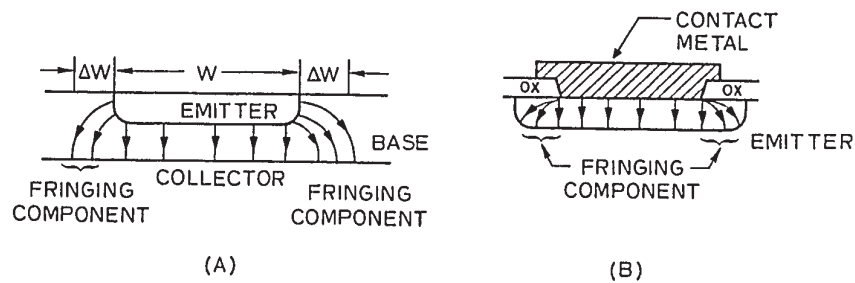


Figure H.1: fringing component of electrons injected into the base

in prediffused emitters because of the border imposed by alignment tolerances. If S_c , the contact recombination velocity in the contact window, is significantly greater than S_{ox} the surface recombination velocity under the passivating oxide, then minority carriers will tend to be drawn in by the contact from adjacent areas. At first it was thought that this collection region might extend as far as a diffusion length from the edge of the contact window; further study indicated that this was an excessively large upper bound. In any case, the effective collection area is increased, so that if the geometrical contact area is used, an overestimate for J_{oc} will result. To ensure that this effect is minimal again requires an estimate for the extent of peripheral collection. This ambiguity is not present for outdiffused emitters using SIPOS or POLY due to the shallow junction depths.

For the base and for the emitter with metallic contact, the assumption of minority carrier transparency is a valid approximation. Additionally, it will be assumed that the gradient of (N_{eff}/D) is small compared to the gradients in the hole density; this is equivalent to the assumption of a constant doping profile. These assumptions allow the use of Laplace's instead of Poisson's equation, and the powerful techniques of conformal mapping can be applied. Recall, for the base region

$$\nabla \cdot \vec{J}_n = 0 \quad (H.1)$$

$$\vec{J}_n = -qD_n \nabla n \quad (H.2)$$

Therefore

$$\nabla^2 n = 0 \quad (H.3)$$

i.e. LaPlace's equation.

The analogy to draw is between $n \iff \phi$ and $\vec{J} \iff \vec{E}$ where n is the electron density, ϕ the electrostatic potential, $\vec{J}$ the current density, and $\vec{E}$ the electric field. The problem can be formulated as shown in Fig. H.2 where ∇_n refers to the normal gradient. The hole density (electric potential) is fixed at the emitter base junction by V_{be} and at the reverse biased base collector junction, it is constrained to be

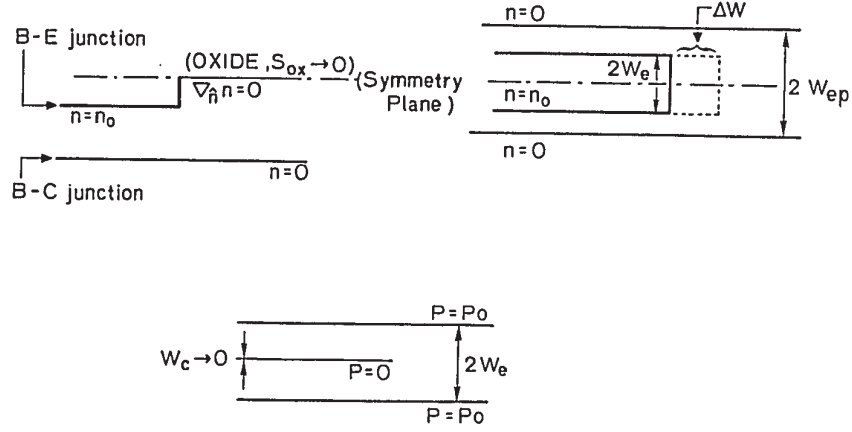


Figure H.2: electrostatic analogy with idealized geometry for the fringe region before and after reflection about symmetry plane

zero. The rounded sidewall has been approximated by a rectangle for simplicity. From the symmetry of the problem, reflection about the x -axis will automatically satisfy $\nabla_n n = 0$, as shown in Fig. H.2 . The problem has been reduced to that of a conducting slab placed between two equipotential plates. Using conformal mapping techniques, the effective end correction ΔW is given by [1]:

$$\Delta W = \frac{2}{\pi} \left[W_{epi} \ln \left(\frac{2W_{epi} - W_e}{W_{epi} - W_e} \right) - W_e \ln \left(\frac{\sqrt{W_e(2W_{epi} - W_e)}}{W_{epi} - W_e} \right) \right] \quad (H.4)$$

For a prediffused emitter where $W_e = 0.8\mu$ and $W_{epi} = 1.9\mu$

$$\Delta W = 0.8\mu$$

For the large emitter device which measures $200\mu \times 200\mu$ this represents only a 2% correction in effective area which is negligible.

For the emitter contact, the situation is analogous if the relevant dimensions and potentials are used as shown in Fig. H.3 . W now must be taken in the limit

as W_c the “contact depth” is reduced to zero. i.e.:

$$\Delta W \rightarrow \frac{2}{\pi}(W_c \ln 2) = .44 W_c$$

Therefore, the contact only sweeps out carriers from a short distance beyond the contact window. This has important implications for low surface recombination velocity contact structures which rely on the presence of a thin interfacial layer such as an oxide to act as a hole barrier. The detrimental effects of two dimensional defects such as cracks are localized (a point defect, being localized in three dimensions would be even more strongly localized). If the intercrack separation exceeds the fringing distance ΔW , significant J_{oe} reduction is still possible.

There remains the case of the prediffused emitters with SIPOS or POLY contacts which are presented in chapter 7. This represents a difficult analytical problem because the recombination under the passivating oxide and in the bulk of the emitter may no longer be negligible. As S_c approaches S_{ox} , the “fringe” spacing loses meaning and the actual collection area becomes ill defined. However, from a practical standpoint, a lower and upper bound can be established at the contact and diffusion areas respectively. For the geometries used, the J_{oe} values that result differ by about 20% which is a sufficiently small difference that taking an average is feasible.

Appendix I

Base Current Crowding

Current crowding was briefly mentioned in Chapter 2 in conjunction with other advantages offered by reducing the base sheet resistivity. In this appendix, a numerical calculation based on a one-dimensional approximation is presented to see at which bias point this effect imposes a limitation on the extraction of J_{oe} from measurement of base current in the test structures.

Consider the cylindrical region in Fig. H.1. By using a region with the same lateral area as the actual intrinsic base region which is rectangular, the voltage drops accompanying the base current can be computed using simple one-dimensional equations. Referring to the differential annular region bounded by r and $r + dr$ as

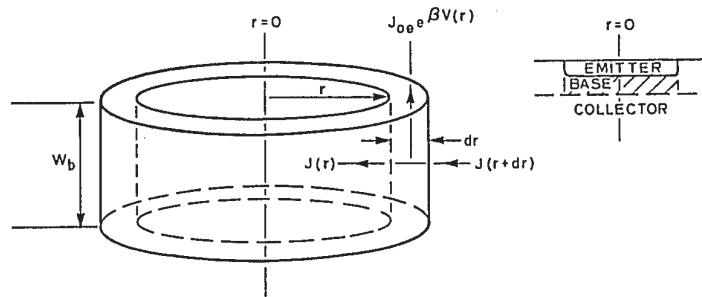


Figure I.1: intrinsic base region (inset), differential element used in quasi-two-dimensional analysis of base current crowding

shown in Fig. I.2 , we can express base current conservation by

$$W_b 2\pi(r + dr)J(r + dr) = W_b(2\pi r)J(r) + 2\pi r dr J_{oe} e^{V(r)/V_T} \quad (I.1)$$

Rearranging yields

$$J(r + dr) \left[1 + \frac{dr}{r} \right] = J(r) + \frac{dr}{W_b} J_{oe} e^{V(r)/V_T} \quad (I.2)$$

where $V(r)$ is the base-emitter bias at radius r . From Ohm's law, defining σ as the base region conductivity

$$J(r) = \sigma E(r) \quad (I.3)$$

$$\frac{dV(r)}{dr} = -\frac{1}{\sigma} J(r) = -\frac{1}{\sigma} J(r) \quad (I.4)$$

since the electric field is the negative gradient of the potential. Equations (H.1,2) when discretized become

$$J_{n+1} = \left(\frac{1}{1 + \Delta r/r} \right) \left(J_n + \frac{\Delta r}{W_b} J_{oe} e^{V_n/V_T} \right) \quad (I.5)$$

$$V_{n+1} = V_n + \frac{\Delta r}{\sigma} J_n \quad (I.6)$$

The boundary conditions are

$$J_{b(r=0)} = 0 \quad V_{(r=0)} = V_{set}$$

where V_{set} is some specified voltage. Equations can then be solved radially outward yielding both the potential distribution and base current densities. The potential and current distributions for a 1μ base with $.5 \text{ ohm-cm}$ resistivity and an emitter J_{oe} of $1 \times 10^{-12} \text{ A/cm}^2$ are displayed in Fig. H.2a,b respectively. The parameters were chosen to depict the worst case scenario (largest device, largest J_{oe}). Current crowding can be defined to begin at the point where the voltage drop from the center of the base to the periphery equals kT/q . For this particular device structure, current crowding begins for an applied potential of approximately 0.66 volts. The terminal voltage is actually somewhat higher as the resistive drop in the extrinsic

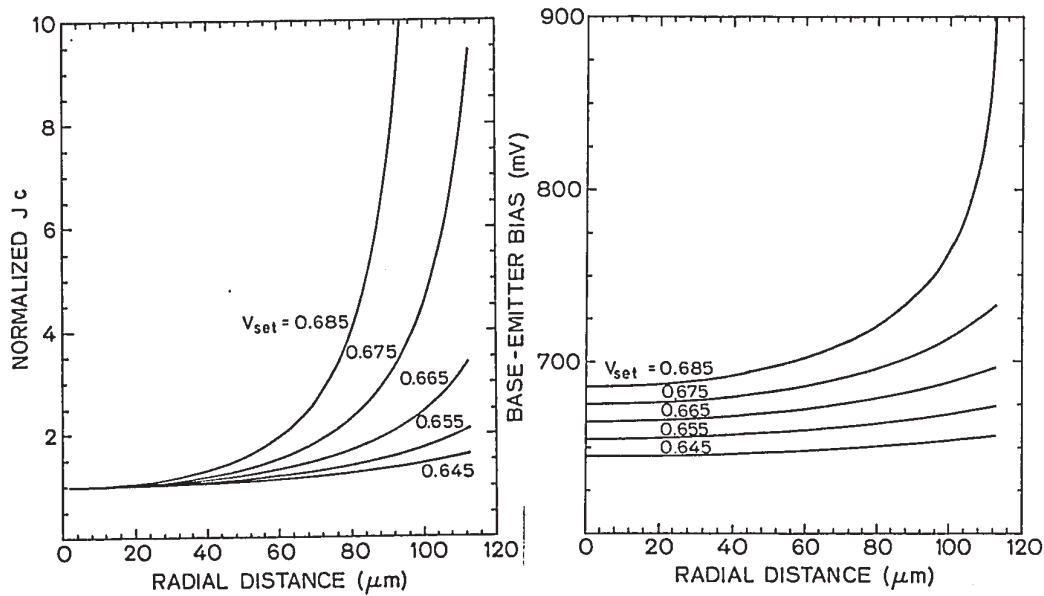


Figure I.2: current density and base emitter potential vs. radial distance

base region was not included. Also, the beta predicted by this simple model does not exhibit any rolloff. Apart from high level injection effects near the perimeter, injection from the emitter sidewall which becomes significant once current crowding is effective produces the observed rolloff in beta [1]; these effects were not included in this treatment.

Appendix J

Effect of BWM on the Collector Temperature

The effects of base width modulation on the collector characteristic associated with changing base-collector bias are readily apparent. A similar effect occurs for changing base-emitter bias, but because the magnitude of the voltage excursion is limited to less than the diffusion potential, the effect is quite small. For extracting device temperatures from the slope of the collector characteristic, however, this effect can introduce significant errors if not accounted for.

The collector current can be written as

$$J_c = J_{oc} e^{V_{bc}/V_T} \quad (J.1)$$

where for a fixed base-collector bias, J_{oc} is taken as a constant. Then taking the derivative with respect to V_{bc} and solving for T , we have

$$T = \frac{q}{k} \left(\frac{\partial J_c / J_c}{\partial V_{bc}} \right)^{-1} \quad (J.2)$$

This then is taken as a measure of the device temperature (or more precisely, the emitter-base junction temperature). In reality J_{oc} exhibits a dependence on V_{bc}

$$J_{oc} = \frac{q D_n n_{i0}^2}{N_b} \frac{1}{W_b} \quad (J.3)$$

$$W_b = W_{bb} - W_{dep}(V_{bc}) \quad (J.4)$$

where the quantities refer to Fig. J.1. This dependence on V_{bc} can be expressed in an alternate form by introducing a bias dependent “ideality factor”, m , such that

$$J_c = J'_{oc} e^{V_{bc}/(mV_T)} \quad (J.5)$$

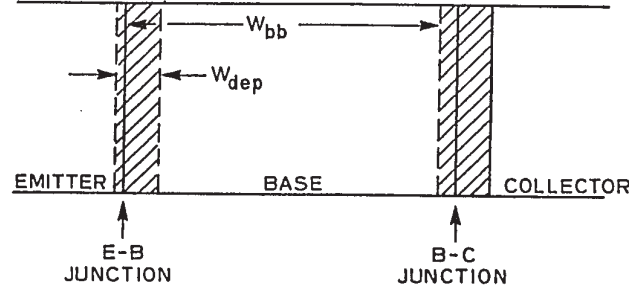


Figure J.1: coordinate definition for BWM analysis

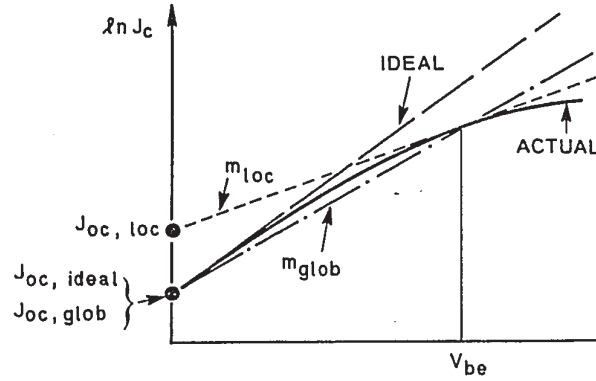


Figure J.2: graphical representation of effective slope m

where J'_{oc} is either J_{oc} or $J_{oc, loc}$ depending upon the definition of m (see Fig. J.2). We can then define two quantities, the “global ideality factor” and the “local ideality factor” by m_{glob} and m_{loc}

$$m_{glob}T \equiv \frac{q}{k} \frac{V_{be}}{\ln(J_c/J_{oc})} \quad (J.6)$$

$$m_{loc}T \equiv \frac{q}{k} \left(\frac{\partial J_c/J_c}{\partial V_{be}} \right)^{-1} \quad (J.7)$$

these can be related to the collector characteristic on a Gummel plot as shown in Fig. J.2. The local ideality factor is seen to be the tangent to the collector characteristic and is the quantity that is more conveniently used. Since it is the

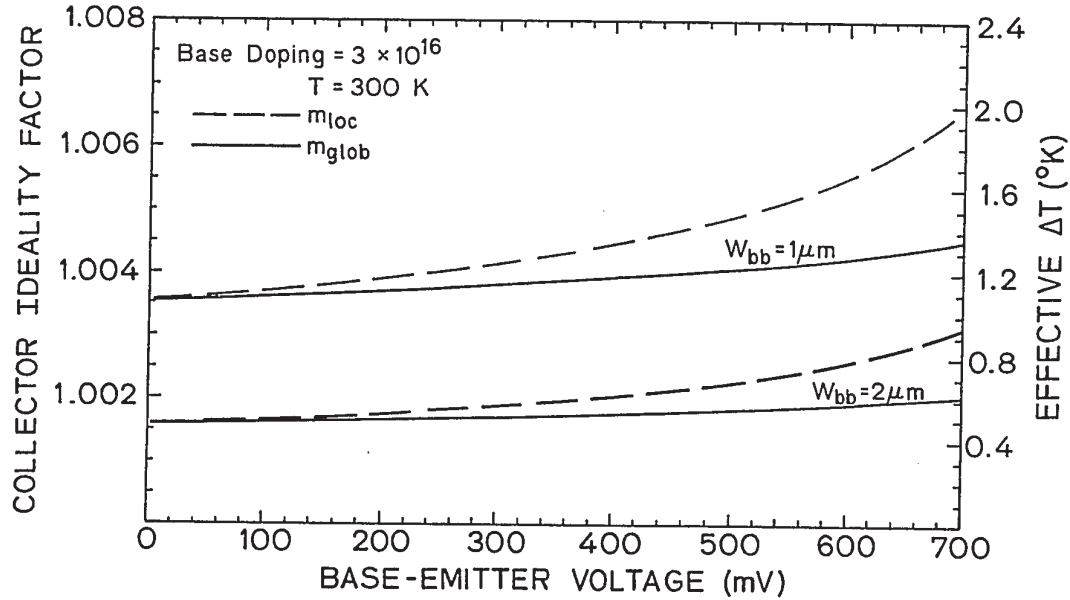


Figure J.3: collector ideality factor, m , vs. V_{be}

quantity mT that appears in (J.2) the error in assuming a unity ideality factor is equivalent to an error in the deduced temperature. The computed characteristics for a uniformly doped base of $3 \times 10^{16} \text{ cm}^{-3}$ and widths W_{bb} (see Fig. J.1) of 1μ and 2μ are displayed in Fig. J.3. These parameters are typical of the test devices. For the $W_{bb} = 1\mu$ device, the effective ideality factor averages 1.0045 for the bias range used to estimate the collector slope; this translates into an error in temperature of 1.4°K , which is quite significant. For bases with smaller Gummel numbers, the effect is more dramatic; collector induced barrier lowering for thin bases operating near punchthrough also aggravates this effect. BWM also affects the value of the base Gummel number since J_{oc} tends to be overestimated. This effect is small however since the fractional error in slope is much less than unity.

Appendix K

Source of the Base Perimeter Current

One of the device geometries in the test mask is a transistor with a field plate over the emitter-base junction as shown in Fig. K.1a. The potential of this plate with respect to the quasi-neutral base can be independently adjusted. For negative bias (plate negative with respect to the base), this results in the formation of an accumulation region at the surface of the relatively lightly doped base in the emitter perimeter region under the field plate as seen in Fig. K.1b. Changing the bias to a more positive voltage depletes and eventually forms an inversion layer in this region as shown in Fig. K.1c.

Gummel plots for a SIPOS contacted device equipped with this field plate are shown in Fig. K.2. For $V_f = -4$ to -10 volts, the portion of the base under the plate is accumulated and the base current is minimal and independent of bias. Increasing V_f to -2 volts results in the appearance of a shoulder in the base characteristic which becomes more prominent as V_f is increased to 4 volts. For this range of bias, the surface is depleted and the recombination is maximized when the surface potential is adjusted so that holes and electrons are present in roughly equal densities at the surface. Since electrons are only present in significant amounts within a lateral distance given approximately by the epi width (see Appendix G), this recombination is only significant for a region extending at most several microns from the emitter sidewall. Increasing the bias to 6 volts marks the onset of inversion and as the surface is now depleted of holes, the recombination current falls again. With the establishment of this inversion layer, the emitter area is effectively increased to

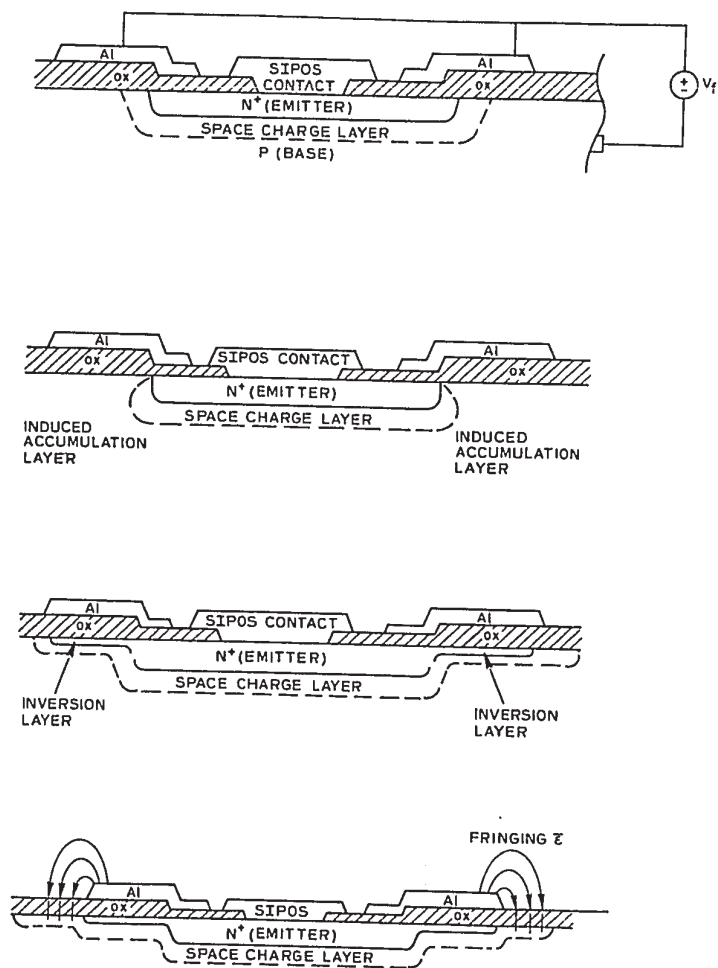


Figure K.1: cross-section of BJT with field plate for several bias conditions

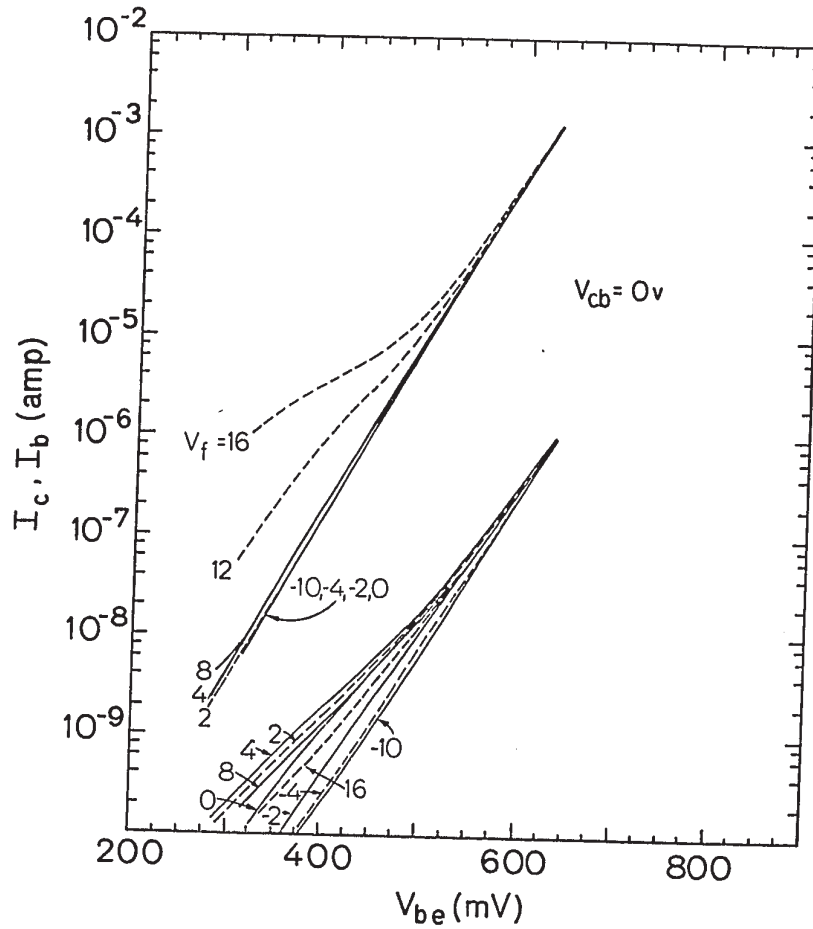


Figure K.2: Gummel plots of field plate equipped BJT

encompass the field induced emitter formed under the plate. As V_g is increased from 12 to 16 volts, the base current approaches a limiting value given by the new geometry as shown in Fig. K.1d. Due to the fringing components of the electric field, the perimeter surrounding the field induced emitter tends to be depleted which keeps the base current above what would be expected from the simple increase in the effective emitter area alone, but below the $V_g = 4$ volts case because the perimeter area is not as strongly depleted.

Examination of the collector characteristic also reveals some interesting trends.

For -10 to 0 volts, there is no apparent change in the collector current. At 4 volts, where the base current has peaked, there is a slight increase, and at 8 volts, the contribution of the field induced emitter has become significant. Note that the increase in collector current tapers off at high current levels. This is probably due to the high sheet resistivity of the induced emitter region, and since current must also flow laterally in this region, lateral voltage drops result in turn off of this perimeter emitter at higher current levels. For $V_g > 10$ volts, a large shoulder appears in the collector curve although the base characteristic displays only minor perturbations. From the layout of the device, this effect is believed to result from the turn-on of a parasitic FET which presents a shunt path from the emitter to collector. This parasitic FET action is no consequence for the geometries and bias points used in devices for J_{oe} extraction.

From the analysis of this structure, it is apparent that the perimeter component which is dominant at low bias and causes a drastic rolloff of beta for high current gain devices is primarily attributable to recombination at the surface of the base in the emitter base-perimeter region. This region includes the space charge layer and any contributions to this layer that result from the combined effects of a low surface doping in the base and positive oxide charge as well as any contributions from external fields. This low current beta degradation should not be as severe in practical devices in which surface inversion is difficult to attain because of the higher surface doping concentrations.

Appendix L

Experiments with Chemically Grown Oxide

In the process of wafer cleaning, $NH_4OH + H_2O_2$ and $HCl + H_2O_2$ solutions are used. These solutions, intended to remove organics and metallics respectively, leave a thin layer of SiO_2 where bare silicon is exposed to these solutions. This layer is used as the “insulator” in the POLY and SIPOS SIS structures. For the actual device fabrication, the wafers see two sequential cleaning cycles as listed in Table L.1; the first is intended as the cleaning step, the second is used to grow the actual oxide.

Since thickness is one the key parameters that must be controlled, a series of experiments was conducted to measure the thickness of these films using ellipsometry as a measurement technique. Although the absolute accuracy of these measurements is uncertain, the precision of the technique is quite good ($\approx 1\text{\AA}$) providing that suitable precautions are taken (such as stabilizing the laser before use). In Fig. L.1, the “ellipsometric” thicknesses are plotted as a function of several different test conditions. Oxide on the wafers A,B,C,D was grown at a $46^\circ C$ bath temperature. The difference between measurements taken 20 minutes (unprimed) and 13 hours (primed) after the final DI rinse are negligible indicating that oxide growth due to exposure to ambient oxygen either occurs in less than twenty minutes or is minimal for these wafers. It is not possible to resolve this contribution without resorting to inordinate measures. Wafer A was dipped in 50:1 HF to remove the chem oxide and was rinsed in DI water for 5 min prior to spin drying. It serves as a control wafer indicating the minimum thickness that can be achieved for practical

1. $H_2SO_4 + H_2O_2$ (1 : 1) 10 minutes
 2. DI H_2O_2 rinse 10 minutes
 3. $H_2O_2 + HF$ (50 : 1) strip
 4. DI rinse 5 min (ultrasonic)
 5. $NH_4OH + H_2O_2 + H_2O$ (1 : 1 : 4.5) (ultrasonic) 45°C – 60°C 10 minutes
 6. 50 : 1 HF strip
 7. DI rinse 5 min (ultrasonic)
 8. 8. $HCL + H_2O_2 + H_2O$ (1 : 1 : 4.5) (ultrasonic) 45°C – 60°C 10 minutes
 9. DI rinse 5 minutes (ultrasonic)
 10. 50 : 1 HF strip
 11. DI rinse 5 minutes (ultrasonic)
- to grow the chem oxide, repeat steps 5,7,8,9

Table L.1: chemox growth sequence

devices using the standard process sequence. Wafer B received only the $HCl+H_2O_2$ growth step in the second cycle while wafer C received both oxide growth steps. The ellipsometer gives the same reading for both wafers. Wafers D and E are repeats of B and C except that the bath temperature was raised to 59°C. Again there is no significant change in thickness indicating that growth may be a self-limiting process. An interesting change occurs however if the wafers are annealed in a H_2 ambient at 800°C for 12 minutes. This simulates the equilibration phase in the epi reactor prior to start of the POLY deposition. The control wafer without the chem oxide reads 13 Å before and after the anneal, but the chem ox wafer is reduced from 16 Å to 13 Å. Although the ellipsometer does not distinguish between these films, they result in quite different device behavior as was seen in chapter 7. The films may be densified, dried, or reduced by this anneal or possibly some combination thereof;

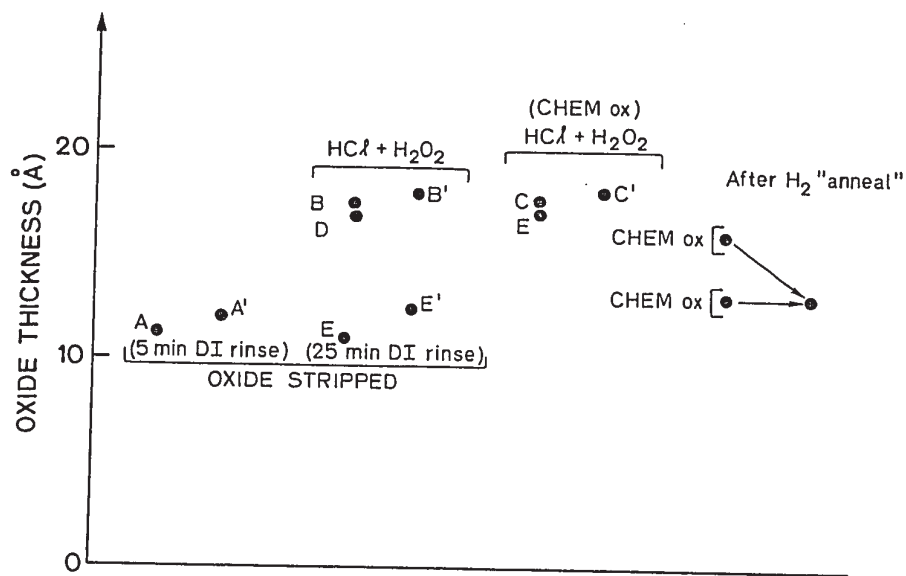


Figure L.1: ellipsometric chemox thickness for several process conditions

however, to quantify impact of this treatment on the tunnel barrier to holes is a difficult experiment.

These results should not be taken as an accurate measure of the interfacial oxide thickness. The influence of adsorbed water or organic films may not be inconsiderable for these thin films. The chemical reactivity of these films is quite different, a $NH_4OH + H_2O_2$ films being more resistant to HF than their $HCl + H_2O_2$ counterparts; however, the ellipsometer does not make any distinction between these films. Even knowledge of the true oxide thickness is of limited value since without independent information of the electronic structure of these films, it presents no new fundamental information beyond that available in the phenomenological tunneling probability.

Appendix M

Bibliography

Chapter 1

- [1] T. Matsushita, et al, "A SIPOS-Si Heterojunction Transistor," Jap. J. Applied Physics, vol. 20(suppl.), p. 75

Chapter 2

- [1] A. Fahrenbruch, Fundamentals of Solar Cells, Academic Press, 1983
- [2] H. de Graaf, et al, "The Emitter Efficiency of Bipolar Transistors," Solid State Electronics, vol. 20, 1977, p. 515
- [3] R. Smith, Semiconductors, Chapter 9, Cambridge University Press, 1979
- [4] S. Sze, Physics of Semiconductor Devices, Chapter 8, John Wiley and Sons, 1969
- [5] R. Swanson, "Silicon Photovoltaic Cells in TPV Conversion," EPRI report ER-1272, Research Project 790-2
"Point-Contact Silicon Solar Cells," EPRI report AP-2859, Research Project 790-2
- [6] R. Pritchard, Electrical Characteristics of Transistors, McGraw Hill, 1967
- [7] S. Ghandi, The Theory and Practice of Microelectronics, John Wiley and Sons, 1968
- [8] H. Gummel, "Measurement of the Number of Impurities in the Base Layer of a Transistor," Proc. IRE, vol. 49, 1961, p. 834
- [9] J. del Alamo, et al, "Physics and Modeling of Heavily Doped Emitters," submitted for publication to IEEE Trans. Electron Devices

- [10] N. Fletcher, "Some Aspects of the Design of Power Transistors," Proc. IRE, vol. 43, 1955, p. 551
- [11] A. Grebene, Analog Integrated Circuit Design, Krieger Pub., 1978
- [12] P. Gray, Analysis and Design of Analog Integrated Circuits, John Wiley and Sons, 1984

Chapter 3

- [1] J. del Alamo, et al, "Physics and Modelling of Heavily Doped Emitters," submitted for publication to IEEE Trans. Electron Devices
- [2] R. Stratton, "Volt-Current Characteristics for Tunneling Through Insulating Films," J. Phys. Chem. Solids, vol. 23, 1962, p. 1177
- [3] R. Smith, Semiconductors, Cambridge University Press, 1978
- [4] T. Ning, et al, "Effect of Emitter Contact on Current Gain of Silicon Bipolar Devices," Tech. Digest IEDM, 1979, p. 473
- [5] H. de Graaff, et al, "The SIS Tunnel Emitter: A Theory for Emitters with Thin Interface Layers," IEEE Trans. Electron Devices 26, 1979, p. 1771
- [6] B. Soerowirdjo, et al, "The Influence of Surface Treatments on the Electrical Characteristics of Polysilicon Emitter Bipolar Transistors," Tech. Digest IEDM, 1982, p. 668
- [7] J. Fossum, et al, "A Minority Carrier Transport Model for Polysilicon Contacts to Bipolar Devices, Including Solar Cells," Tech. Digest IEDM, 1980, p. 280
- [8] A. Eltoukhy, et al, "The Role of the Interfacial Layer in Polysilicon Emitter Bipolar Transistors," IEEE Trans. Electron Devices 29, 1982, p. 1862
- [9] T. Kamins "Hall Mobility in Chemically Deposited Polycrystalline Silicon," J. Appl. Phys., vol. 42, 1971, p. 4357
- [10] J. Snel, "The Effect of Donors or Acceptors on the $Si - SiO_2$ Interface," Inst. Phys. Ser. # 50, Chapter 2, 1980, p. 119
- [11] A. Milnes, Heterojunctions and Metal-Semiconductor Junctions, Academic Press, 1972

Chapter 4

- [1] J. del Alamo, et al, "Physics and Modelling of Heavily Doped Emitters," submitted for publication to IEEE Trans. Electron Devices
- [2] J. Snel, "The Effect of Donors and Acceptors on the $Si - SiO_2$ Interface," Inst. Phys. Ser. # 50, Chapter 2, 1980, p. 119
- [3] R. Warner, Integrated Circuits – Design Principles and Fabrication, Chapter 4, Mc Graw Hill, 1965
- [4] J. Lamming, "Microwave Transistors," in Microwave Devices - Device Circuit Interactions, Howes & Morgan ed., John Wiley and Sons, 1976
- [6] R. Fair, et al, "A Quantitative Model for the Diffusion of Phosphorous in Silicon and the Emitter Dip Effect," J. Electrochem. Soc., vol. 124, 1977, p. 1107
- [7] A. Phillips, Transistor Engineering and Introduction to Integrated underline Semiconductor Circuits, McGraw Hill, 1962

Chapter 5

- [1] T. Matsushita, et al, "Highly Reliable High Voltage Transistors by Use of the SIPOS Process," IEEE Trans. Electron Devices, vol. 23, 1976, p. 826
- [2] T. Matsushita, et al. "A Silicon Heterojunction Transistor," Appl. Phys. Lett., vol. 35, 1979, p. 549
- [3] T. Matsushita, et al, " A SIPOS-Si Heterojunction Transistor," Jap. J. Applied Physics, vol. 20 suppl., p. 75
- [4] H. Philipp, "Optical Properties of Non-Crystalline Si , SiO , SiO_x , and SiO_2 ," J. Phys. Chem. Solids 32, 1971, p. 1935
- [5] H. Yamoto, et al, "New Technology by the Use of the Doped SIPOS Films," Electrochem. Soc. 1979 meeting, abstract 168
- [6] N. Zommer, "Characteristics of CO_2 Deposited SIPOS Films," IEEE Trans. Electron Devices, vol. 27, 1980, p. 2056
- [7] W. Knolle, et al, "A Model of SIPOS Deposition Based on Infrared Spectroscopic Analysis," J. Electrochem. Soc., vol. 127, 1980, p. 2254

- [8] J. Thomas, et al, "AES and XPS Studies of Semi-Insulating Polycrystalline Silicon (SIPOS) Layers," J. Electrochem. Soc., vol. 126, 1979, p. 1766
- [9] T. Adachi, et al, "AES and PES Studies of Semi-Insulating Polycrystalline (SIPOS) Films," J. Electrochem. Soc., vol. 127, 1980, p. 1617
- [10] M. Hamasaki, et al, "Crystallographic Study of SIPOS Doped with Oxygen Atoms," J. Appl. Phys. vol. 49, 1978, p. 3987
- [11] A. Goodman, "Optical Properties of Amorphous and Recrystallized SiO_x Layers," Inst. Phys. Conf. # 43, 1979, p. 805
- [12] A. Hartstein, et al, "Observation of Amorphous Silicon Regions in Si -rich SiO_2 Films," Appl. Phys. Lett., vol. 36, 1980, p. 836
- [13] J. Ni, et al, "Electrical Conductivity of SIPOS and Its Dependence on Oxygen Content," Appl. Phys. Lett., vol. 39, 1981, p. 554
- [14] K. Ng, et al, "Assymetry in the SiO_2 Tunneling Barriers to Holes and Electrons," J. Appl. Phys., vol. 51, 1980, p. 2153
- [15] T. Yamaguchi, et al, "An Advanced MOS-IC Process Technology Using Local Oxidation of Oxygen-doped Polysilicon Films," IEEE J. Solid State Circuits, vol. 13, 1978, p. 472
- [16] T. Adachi, et al, "Electronic Properties of P-doped SIPOS," Electrochem. Soc. 1979 Meeting, abstract 167
- [17] B. Abeles, et al, "Structural and Electrical Properties of Granular Metal Films," Adv. Phys., vol. 24, 1975, p.407
- [18] M. Hamasaki, et al, "Electronic Properties of SIPOS Doped with Oxygen Atoms," Sol. St. Comm., vol. 21, 1977, p. 591
- [19] M. Lannoo, et al, "A Cluster Plus Effective Medium Tight Binding Study of SiO_x Systems," Sol. St. Comm., vol. 28, 1978, p. 733
- [20] K. Seaward, in Proceedings ECS Fall Meeting, vol. 77-2, p. 948, Oct. 1977

Chapter 6

- [1] F. Lindholm,, et al "A Methodology for Experimentally Based Determination of Gap Shrinkage and Effective Lifetimes in the Emitter and Base of

p-n Junction Solar Cells and Other p-n Junction Devices," IEEE Trans. Electron Devices, vol. 24, 1977, p. 402

- [2] C. Kirk, "A Theory of Transistor Cutoff Frequency Falloff at High Current Densities," IRE Trans. Electron Devices, vol. 9, 1962, p. 164

Chapter 7

- [1] M. Finetti, et al, "Aluminum-Silicon Ohmic Contact on "shallow" n^+/p Junctions," Sol. St. Electron., vol. 23, p. 255
- [2] T. Hall, " A Study of the Interface Properties of Deposited Silicon Dioxide on Silicon," Engineers Thesis, Stanford Electronics Laboratories, Stanford, CA.
- [3] J. del Alamo, et al, "The Physics and Modeling of Hevaily Doped Emitters," submitted for publication to IEEE Trans. Electron Devices
- [4] T. Matsushita, et al, " A SIPOS-Si Heterojunction Transistor," Jap. J. Applied Physics, vol. 20 suppl., p. 75
- [5] M. Hamasaki, et al, "Crystallographic Study of SIPOS Doped with Oxygen Atoms," J. Appl. Phys., vol. 49, 1978, p. 3987
- [6] T. Adachi, "Electronic Properties of P-doped SIPOS," Electrochem. Soc. 1979 Meeting, abstract 167
- [8] specification on Spreading Resistance Profile Data Sheet, SOLECON Labs
- [9] J. Ni, et al, "Electrical Conductivity of SIPOS and Its Dependence on Oxygen Content," Appl. Phys. Lett., vol. 39, 1981, p. 554
- [10] A. Goodman, "Optical Properties of Amorphous and Recrystallized SiO_2 Layers," Inst. Phys. Conf. # 43, 1979, p. 805
- [11] H. Berger, et al, "Models for Contacts to Planar Devices," Tech. Digest International Solid State Circuits Conf., 1969, p. 160
- [12] E. Nicollian, MOS Physics and Technology, John Wiley and Sons, 1982