AN ANALYTIC MODEL FOR HIGH ELECTRON MOBILITY TRANSISTORS

by

Adrian John Hill

Submitted in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy
in the
Department of Electronic Engineering,
University of Natal
1986

ABSTRACT

The last six years has seen the emergence and rapid development of a new type of field effect transistor, the High Electron Mobility Transistor (HEMT), which offers improved performance in both digital and analogue circuits compared with circuits incorporating either MEtal Semiconductor (MES) or Metal Oxide Semiconductor (MOS) FETs. A new physically-based analytic model for HEMTs, which predicts the DC and RF electrical performance from the material and structural parameters of the device, is presented. The efficacy of the model is demonstrated with comparisons between simulated and measured device characteristics, at DC and microwave frequencies.

The good agreement with experiment obtained with the model indicates that velocity overshoot effects are considerably less important in HEMTs than has been widely assumed, and that the electron transit velocity in submicron devices is approximately 10^7 cm/s, rather than around $2x10^7$ cm/s.

The Inverted HEMT, one of the major HEMT structural variants, is emphasized throughout this work because of its potential advantages over other variants, and practical results from 0.5 micron gate length Inverted HEMTs are presented.

PREFACE

The reasearch described in this thesis was carried out by the author in the Compound Semiconductor Laboratory at Hirst Research Centre, Wembley, England during the period from June 1984 to November 1986. During this period the author was on secondment from the Department of Electronic Engineering of the University of Natal at Durban, South Africa, where he was employed as an Assistant Research Fellow under the supervision of Professor H.L. Nattrass, Head of Department.

This thesis represents original work by the author and has not been submitted in any form to another University for any degree. Where use has been made of work carried out by others, it has been duly acknowledged in the text.

ACKNOWLEDGEMENTS

The author would like to thank all those who have contributed in any way to this work, and in particular:

Professor H.L.Nattrass for providing the opportunity to go to Hirst Research Centre, and for supervision and guidance.

Professor K.F. Poole for his support and encouragement

Dr. P.H. Ladbrooke for his technical supervision and infectious enthusiasm. His considerable experience in MESFET modelling provided the seeds of ideas for much of this work.

Drs. M.S. Frost, T.M. Kerr and C. Sansom, Mr. S. Ransome and Mr. D. Westwood for device fabrication.

Mr. C. Caddick and Mr. J.P. Bridge for RF assessment and useful discussions.

and

Mrs. F. Viljoen for typing.

The financial support of the University of Natal and Hirst Research Centre is also gratefully acknowledged.

CONTENTS

Pages are numbered by Chapter or Appendix and a more detailed contents page is located at the start of each chapter.

LIST OF PRINCIPAL SYMBOLS

d	Gate to 2DEG spacing
E_c,(E_v)	Conduction-, (valence-) band energy
ΔE_c	Conduction-band discontinuity
E_g	Energy band gap
$f(E)$	Fermi-Dirac disribution
F	Electric field
F_s	Critical field for electron velocity saturation
$g(E)$	Density of available states function
g_m,(g_{mo})	Extrinsic, (intrinsic) transconductance
h	Planck's constant
k	Boltzmann's constant
L_1,(L_3)	Length of region I, (III)
m_c	Density of states effective mass
N_D	Donor density in doped AlGaAs
n_{so}	Undepleted 2DEG density
n_{ss}	Occupied surface state density
q	Electronic charge
R	Ratio of substrate to 2DEG current
R_s,R_d,R_g	Source, drain and gate resistances
v_s	Saturated drift velocity of 2DEG electrons
V_1,(V_2)	2DEG potential at the end of region I (II)
V_{DS}	Applied drain-source voltage
V_{GS}	Applied gate-source voltage
V_T	Threshold voltage
Y_a	Doped AlGaAs layer thickness
Y_n	Cap layer thickness
Z	Gate width
ε	Permittivity
ρ	Volume charge density
τ_e	Energy relaxation time
μ_o	Low field electron mobility
ϕ_m	Schottky-barrier height

LIST OF FIGURES

(The Figure captions appearing in this list are shortened versions of the titles appearing in the text).

CHAPTER ONE

INTRODUCTION

INTRODUCTION

This thesis represents a contribution by the author to the field of gallium arsenide transistor modelling. A new analytic model for High Electron Mobility Transistors (HEMTs) is presented, which predicts DC and RF electrical characteristics from the material and structural parameters of the device. The model has been used as the basis of a software simulation package written by the author. Comparisons are made between simulated and measured HEMT electrical performance to demonstrate the efficacy of the model.

The last six years has seen the emergence and rapid development of HEMTs, a new class of compound semiconductor field effect transistor which exploit the superior transport properties of electrons moving adjacent to a heterojunction interface. These devices provide enhanced performance over conventional metal semiconductor field effect transistors (MESFETs) in both digital and analogue integrated circuits.

In the next chapter, the principles of operation of HEMTs are presented and the two main structural variants, conventional HEMT (CHEMT) and inverted HEMT (IHEMT), are discussed. To date, the IHEMT has received much less attention from researchers than the CHEMT because of the larger mobilities obtainable in the latter structure. Throughout this work, the IHEMT is emphasised because of its potential advantages over the CHEMT. Several of these advantages, which have not been reported elsewhere, are also discussed in Chapter 2. Practical results obtained on devices fabricated at Hirst Research Centre, to the author's layer specification, are also presented.

These results indicate that 0.5 micron gate length devices can provide useful amplifier gain to 30 GHz. The dependence of the conduction band energy discontinuity on aluminium mole fraction, in gallium-arsenide/aluminium gallium arsenide (GaAs/AlGaAs) heterojunctions, is also discussed in Chapter 2. The dependence of this important parameter was independently deduced by the author and has subsequently been shown to be correct by recently published data. The chapter is concluded with a discussion on superlattice-based HEMTs and practical results obtained on devices of this type, fabricated at Hirst Research Centre.

In chapter 3, a new analytic model for the IHEMT is presented. While several models for predicting the I_{DS}-V_{DS} characteristics of CHEMTs have been published, none for the IHEMT (apart from the author's own work) have appeared in the literature to date. In this chapter, the IHEMT model is expanded to analyse CHEMTs as well. The model presented here includes several physical effects not included in other models, such as the extension of the active region beyond the drain edge of the gate, and the effect of occupied surface charge. The increase in drain current arising from the real-space transfer of hot electrons into the layer under the primary conducting path is also considered for the first time. The model is physically-based (process independent) and the physical quantities used in the model are not required to assume non-physical values to obtain good agreement between experimental and simulation results. In particular, a physically realistic value of 10^7 cm/s is used for the saturated drift velocity in all the simulations, rather than treating this parameter as an adjustable parameter to obtain a good fit to measured data.

In Chapter 4, the model is extended to predict the RF performance of HEMTs. The RF performance is represented by an equivalent circuit, the elements of which (with the exception of some parasitics) are calculated using the DC model. The RF model is a large-signal (non-linear) one, because the bias dependence of the elements is implicit in their calculation from the DC model. Scattering parameters and several amplifier gains are calculated from the equivalent circuit.

The DC and RF models have formed the basis of a software simulation program written by the author. This program has facilitated the comparisons between measured and simulated device characteristics presented in this thesis, and finds increasing application in device diagnostic work carried out at Hirst Research Centre.

In Chapter 5, conclusions are drawn from the results of this work and suggestions for further model development are made.

The appendices contain additional information related to topics in the main thesis, and papers written by the author. Appendix A contains a review of HEMT development written for the GEC Journal of Research. Appendix B details some of the model equation derivations discussed in Chapter 3. Appendices C to E contain papers by the author presented at conferences in 1985, which report results of practical IHEMTs and early modelling work. Appendix F is a reprint of an Electronics Letter on the subject of conduction band discontinuities. Appendix G contains a paper written by Ladbrooke to which the thesis author contributed the calculations, using a modified form of the model described above, for the comparison between GaAs and Si FETs.

This work yields a different interpretation to that commonly given for the enhanced performance of GaAs FETs over their silicon counterparts.

References cited in the text are included at the end of each respective chapter.

CHAPTER TWO

CHEMTS and IHEMTs

2.1 INTRODUCTION

The last six years has seen the emergence and rapid development of field-effect transistors which exploit the superior transport properties of electrons moving along the heterojunction interface between two compound semiconductor materials. Several acronyms are used for these transistors including HEMT (High Electron Mobility Transistor), MODFET (Modulation-Doped Field Effect Transistor), TEGFET or 2DEGFET (Two-dimensional Electron Gas FET) and SDHT (Selectively Doped Heterojunction Transistor).

HEMT, the acronym used by the developers of the first such device [2.1], will be used here as the generic term for these transistors. The two main structural variants, namely the conventional HEMT and inverted HEMT, will be termed CHEMT and IHEMT respectively, when a distinction between them is required. The IHEMT has been largely neglected in published papers in favour of the CHEMT and consequently most papers use the term HEMT to mean specifically CHEMT.

In the next section, the device structure and principles of operation are outlined. Potential advantages of IHEMTs compared with CHMTs are discussed. In section 2.3, practical results obtained on IHEMTs fabricated at GEC (Hirst Research Centre) are presented. The RF data in this section constitute the first published results of the RF performance of IHEMTs [2.2]. In Section 2.4, the dependence of the conduction band edge discontinuity on aluminium mole fraction is discussed. The dependence of this important parameter was independently deduced by the author and subsequently shown to be correct by recently published data. Limitations in HEMTs with aluminium gallium arsenide are discussed in Section 2.5, together with results on superlattice-based HEMTs. The chapter is concluded with a summary and a list of the cited references.

2.2 CHEMT and IHEMT

2.2.1 Fundamentals of operation

The structure and operation of HEMTs is similar to that of enhancement mode n-MOSFETs (Figure 2.1). In both HEMTs and nMOSFETs, the primary conduction path between source and drain is a two-dimensional electron gas (2DEG), whose sheet charge is modulated by the potential on the gate electrode. In both devices, the electron sheet is confined by band bending to a position, in the lower bandgap layer, adjacent to the interface between the two layers of differing bandgap. In MOSFETs, the smaller bandgap silicon layer is separated from the gate by the wide-bandgap, non-crystalline silicon dioxide insulating layer.

In HEMTs, the bulk silicon layer is replaced by nominally-undoped gallium arsenide (GaAs) and the larger-bandgap material is n-doped aluminium gallium arsenide (AlGaAs). The major structural difference between CHEMTs and IHEMTs is that in the former, the n-AlGaAs is grown on top of the GaAs layer, and in the latter this order is reversed (Figure 2.1). In both CHEMTs and IHEMTs, electrons from the ionized donors in the n-AlGaAs will move into lower energy states in the GaAs where they accumulate in a thin sheet (2DEG). In HEMTs, unlike MOSFETs, the interface between the two layers is atomically smooth because the AlGaAs layer is crystalline and lattice-matched to within 0.1 percent with the GaAs layer.

In this work, only GaAs/AlGaAs HEMTs are considered. Other material combinations such as InGaAs/AlGaAs HEMTs are emerging, with prospects of enhanced performance, but have received much less attention in the literature to date.

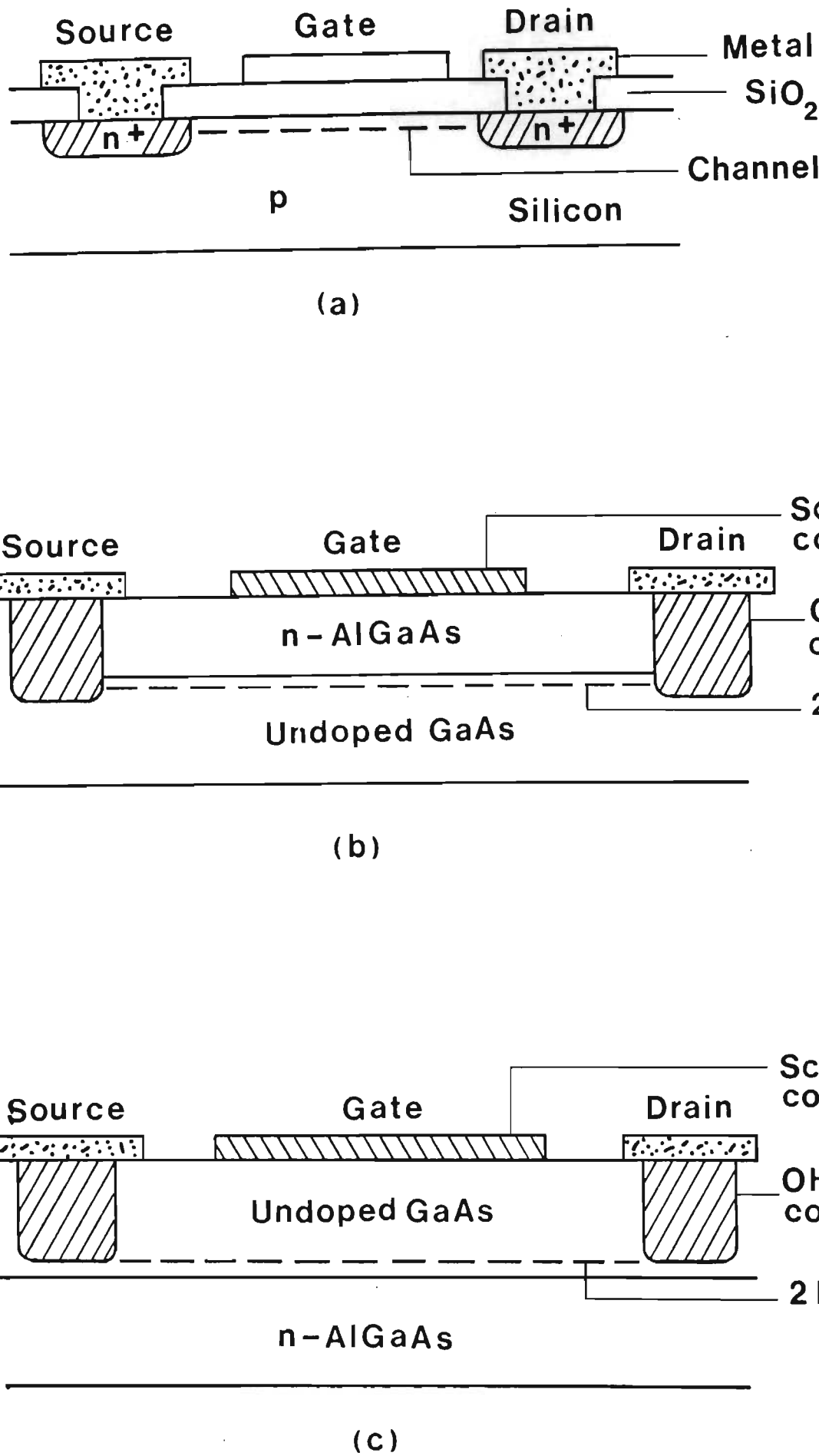


Figure 2.1. Simplified cross-section through (a) Silicon MOSFET (b) CHEMT and (c) IHEMT

An important consequence of the transfer of electrons from the AlGaAs to the GaAs layer is that the spatial separation of conduction electrons from their parent donors reduces the ionised impurity scattering rate. This results in enhanced electron mobility which in turn results in enhanced electrical performance. The ability to obtain a large sheet concentration of conduction electrons without incurring the penalty of a commensurately large impurity scattering rate is one of the major advantages of the HEMT structure over the standard GaAs MESFET.

A more detailed account of the principles of operation and electrical performance of HEMTs may be found in a review paper by the author [2.3], which is given in Appendix A. Since that paper was written (in early 1985) almost all the best-electrical performance figures given in that article have been bettered.

In digital circuits, the highest published operating frequency for a divide-by-two circuit, operating at room temperature, is 9.15 GHz with a power consumption of only 25 mW [2.4]. A 1.5 K gate array has been produced by Fujitsu, and an 8 by 8 bit multiplier implemented on the array has shown multiplication times of 3.1 ns at room temperature [2.5].

Analogue amplifiers with 7.5 dB gain at 61 GHz [2.6], and 3.6 dB at 94 GHz [2.7] have been demonstrated. One of the best noise performances achieved to date is 2.7 dB noise figure at 62 GHz [2.8].

Discrete CHEMTs with an extrinsic transconductance of 450 and 570 mS/mm, at 300 and 77K respectively, have been achieved by Camnitz et al. [2.9]. IHEMTs with an extrinsic transconductance of 1180 (1810) mS/mm at 300 (77K) have been demonstrated [2.10], but with an f_T of only 7.5 GHz. Given the

pace of HEMT development globally, all these performance figures will almost certainly soon be exceeded.

Since the review was written, considerable progress on the conduction band discontinuity question has been made. This subject will be further discussed in Section 2.4. Also, as a result of the development of the HEMT model, discussed in the next chapter, it is the author's view that velocity overshoot is less important in HEMTs than is widely believed. A more complete discussion of this issue is given in Chapter 3.

2.2.2 Advantages of IHEMT

The IHEMT has received far less attention than the CHEMT in the literature to date. This is mainly because of the higher mobilities obtained in normal than in inverted (GaAs on AlGaAs) heterojunction structures, during early work in this field [2.11]. Subsequent technological improvements (discussed in section 6 of Appendix A), and the general recognition that extremely high mobilities are of secondary importance in determining overall device performance, mean that the structural advantages inherent in the IHEMT make it a promising device for both digital and analogue applications.

In the rest of this section, some potential advantages of the IHEMT are discussed. Several others are contained in section 6 of the Appendix A. Some of these advantages have not appeared elsewhere in the literature to date.

Threshold voltage sensitivity

In digital circuits using enhancement mode (normally-off) switching devices, close control of the device threshold voltage, V_T, is essential to achieve reasonable yield. The threshold voltage is that applied gate-source voltage

which just reduces the drain-source current to zero (often termed pinch-off voltage in MESFETs). The threshold voltage is a function of several structural and material parameters in the device, one of the most important of which is the gate to 2DEG spacing, d. The threshold voltage equations for the IHEMT and CHEMT are derived in Chapter 3, and are not required directly for the purposes of this discussion. The sensitivity of the threshold voltage to the gate-2DEG spacing can be found by finding the derivative $\partial V_T/\partial d$. For the CHEMT, differentiating equation 3.56 yields:

$$\partial V_T/\partial d \simeq 2(V_T-\phi_m + \Delta E_c)/d \qquad (2.1)$$

where ϕ_m and ΔE_c are the Schottky barrier height and conduction band discontinuity (both in units of potential), respectively. For the IHEMT, differentiating equation 3.8 gives:

$$\partial V_T/\partial d \simeq (V_T-\phi_m)/d \qquad (2.2)$$

From these equations it is evident that for the same threshold voltage and gate-2DEG spacing (to get similar transconductance), the IHEMT exhibits around 50 percent less threshold sensitivity than the CHEMT. Therefore, for a given degree of process control, a higher yield can be expected from IHEMT digital circuits.

Transconductance compression

As the gate-source voltage is reduced towards the threshold voltage, CHEMTs typically exhibit a significant decrease in transconductance. This effect, termed transconductance compression, occurs because the effective gate to 2DEG spacing increases near pinchoff as a result of widening of the

triangular well containing the 2DEG. According to simple theory, the transconductance is inversely proportional to the gate to 2DEG spacing. Transconductance compression is undesirable in digital circuits, where, for maximum switching speed, the transconductance must be as large as possible over the full range of operating gate bias above V_T.

In IHEMTs, transconductance compression is less severe because of the enhanced electron confinement inherent in the device structure. The conduction band discontinuity is effective in confining the 2DEG so that as pinchoff is approached, the well width is decreasing, rather than increasing. The maximum gate to 2DEG spacing is limited to the distance from the gate to heterointerface.

Maximum transconductance

In order to maximize device transconductance, the gate-2DEG spacing, d, should be minimised. In HEMTs, the centroid of the 2DEG is around 8 nm from the heterojunction [2.12], which means that for a given gate to heterojunction spacing, the distance d is smaller in IHEMTs than in CHEMTs.

In CHEMTs, a lower limit on the AlGaAs layer thickness is set by the maximum doping density which can be achieved because the 2DEG-supplying layer is between the 2DEG and the gate. In IHEMT, this limitation does not exist because the donors are under the 2DEG.

In both IHEMTs and CHEMTs, the mobility of 2DEG electrons is increased by including an undoped AlGaAs layer (termed the spacer layer) between the doped AlGaAs layer and the undoped GaAs layer where the 2DEG is accumulated. In CHEMTs this leads to an increase in gate-2DEG spacing, a further deleterious effect which is absent in the IHEMT.

Das [2.13] has shown that a high aspect ratio (Lg/d), where L_g is the gate length, is necessary in devices required to yield useful gain at millimeter-wave frequencies. Since d is not limited in IHEMTs by the doping density, a high ratio can be achieved for shorter gate lengths in IHEMTs than in CHEMTs. This indicates that IHEMTs may offer a better prospect than CHEMTs for millimeter-wave simplification.

2.3 PRACTICAL IHEMT RESULTS

In this section practical results obtained on IHEMTs fabricated at GEC Hirst Research Centre are presented.

Several wafers were grown, by MBE, to a layer specification from the author. The layer specification is illustrated schematically in Figure 2.2. The spacer, between the doped AlGaAs and the undoped GaAs layer containing the 2DEG, consists of four thin undoped layers (2 nm GaAs, 3 nm AlGaAs, 2 nm GaAs and 2 nm AlGaAs). A short superlattice of this type has been shown by Fischer et al. [2.14] to yield enhanced 2DEG mobilities in inverted heterojunctions, by gettering impurities and acting as a diffusion stop to silicon donors (also discussed in Section 6.1, Appendix A). A thick cap layer is included to reduce the parasitic source and drain resistances [2.15]. The MBE growth rates were approximately 0.5 and 0.72 μm/hr for the GaAs and AlGaAs layers respectively.

Device fabrication after layer growth was similar to that used for processing conventional MESFETs. Device isolation was achieved with a proton implantation. Ohmic contact metallisation layers of 100 nm AuGe and 30 nm Ni were deposited, and after liftoff, alloyed in a thermal furnace for 20 seconds. Gate regions were defined by electron beam direct writing into PMMA resist. Gate recesses were etched using $NH_4OH:H_2O_2:H_2O$ (3:2:800). Finally gate metallisation of 100/400 nm of Ti/Au was deposited and lifted off to form the gates.

The layout (plan view) of a fabricated device is shown in Figure 2.3. The pad layout was arranged to allow DC and RF measurements to be made on the wafer using coplanar probes on a Cascade prober. The large pads on either

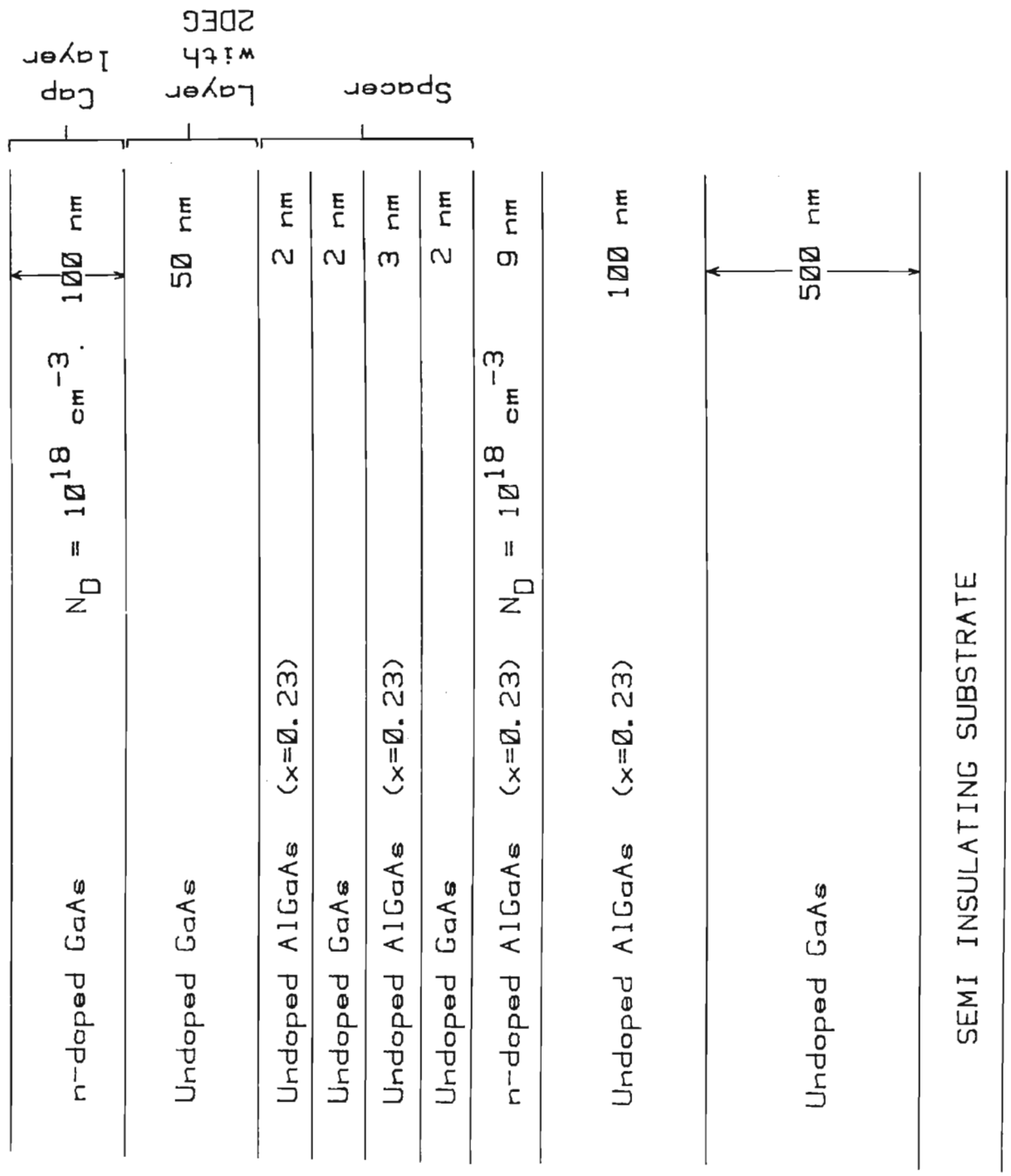


Figure 2.2. Layer structure of IHEMT wafers grown by MBE.

side of the active area are source contacts. The single gate and drain contacts are on opposite sides, down the centre, of the active area. The gate length was 0.5 μm and the total gate width was 120 μm. Devices with different gate widths but with the same gate length and similar layout geometry were also fabricated.

DC I_{DS}-V_{DS} characteristics obtained on a curve tracer for typical devices from wafers DB109b and DB110b are shown in Figures 2.4 and 2.5 respectively: The transconductance is around 195 mS/mm for both devices. The experimental data in Figure 2.4 will be used in chapter 3 for comparison with characteristics simulated using the model to be presented there. The curves of Figure 2.5 exhibit good saturated characteristics and abrupt pinch-off with little transconductance compression.

The maximum transconductance measured was around 300 mS/mm but the majority of devices, from several wafers, exhibited transconductances between 160 and 250 mS/mm. The gate-drain breakdown voltage was typically greater than 16V. Figure 2.6 shows the characteristics of a 120 μm gate width device (number 4335, wafer DB109b) operating up to a drain-source voltage of 18V, over a range of gate bias. The gate-drain breakdown voltage of greater than 18V for a 0.5 μm device compares favourably with the 4 to 8V typical of 1 μm gate length CHEMTs and 14 to 18V recently achieved in 1 μm double heterojunction HEMTs [2.16].

S-parameters were measured using a Cascade prober connected to an HP8510 network analyser. The measured magnitudes of S_{21}, for twelve devices from DB110b are shown in Figure 2.7. The rapid decrease in S_{21} with increasing frequency was attributed to the large parasitic gate-source capacitance present in the structure (Figure 2.3). Figure 2.5 shows the calculated

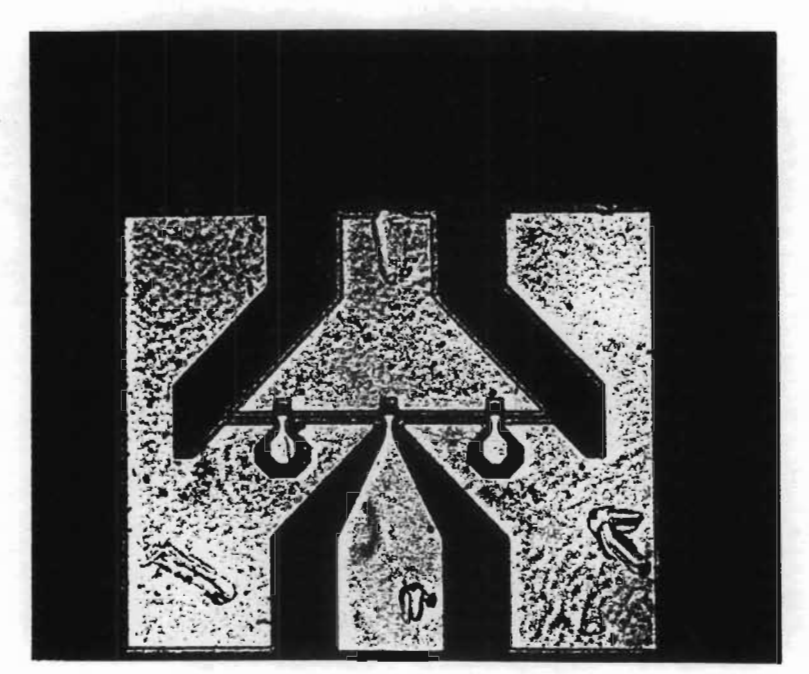


Figure 2.3 Plan view of fabricated IHEMT

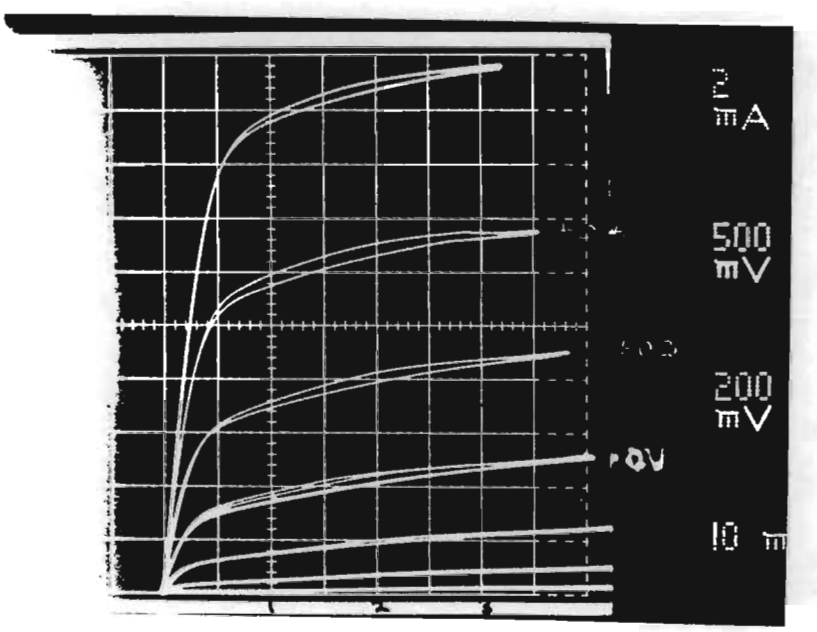


Figure 2.4 I-V characteristcs for device number 5453, from wafer DB109b. Gate width is 120 um. The gate voltage of the uppermost curve is +0.6 V

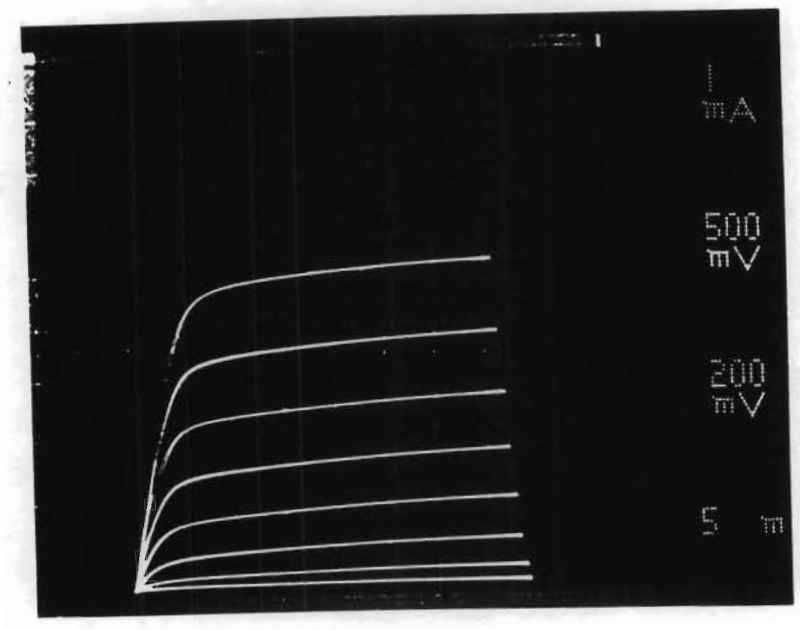


Figure 2.5 I-V characteristics for device number 2539, from wafer DB110b. Gate width is 30 micron.

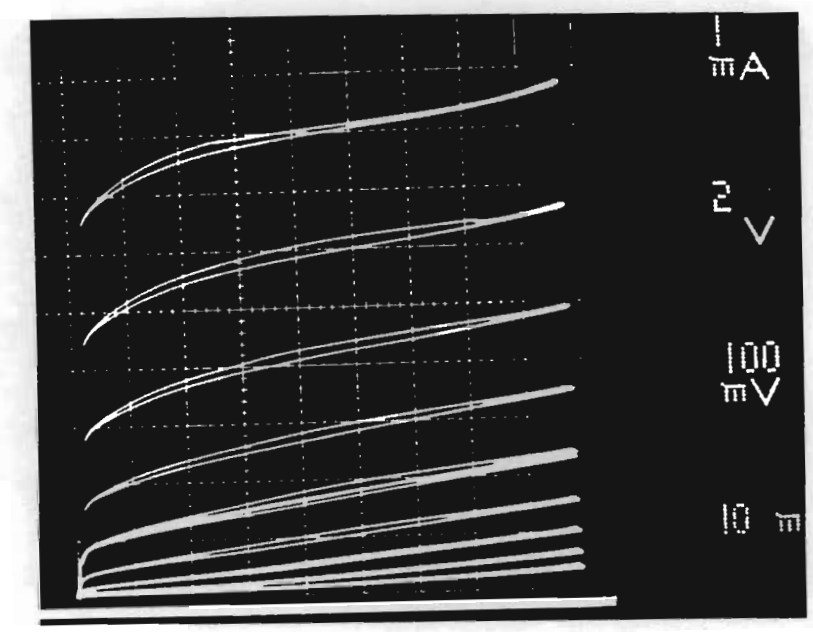


Figure 2.6 I-V characteristics for device number 4335, from DB109b, showing the high operating voltages possible with these devices. Gate width is 120 micron, and the uppermost curve is for a gate voltage of +0.4 V.

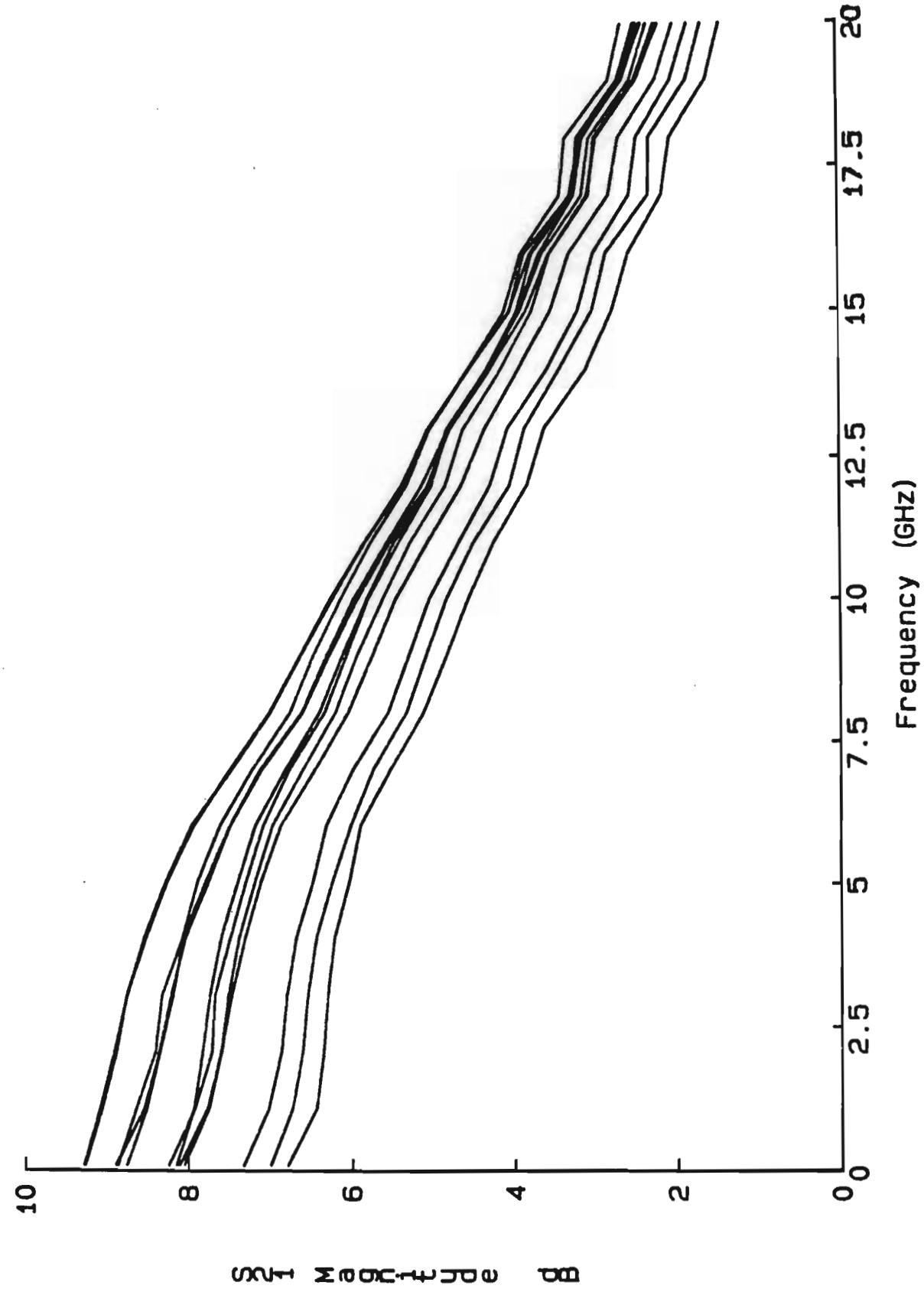


Figure 2.7 Measured magnitudes of S_{21} for twelve devices from DB110b.

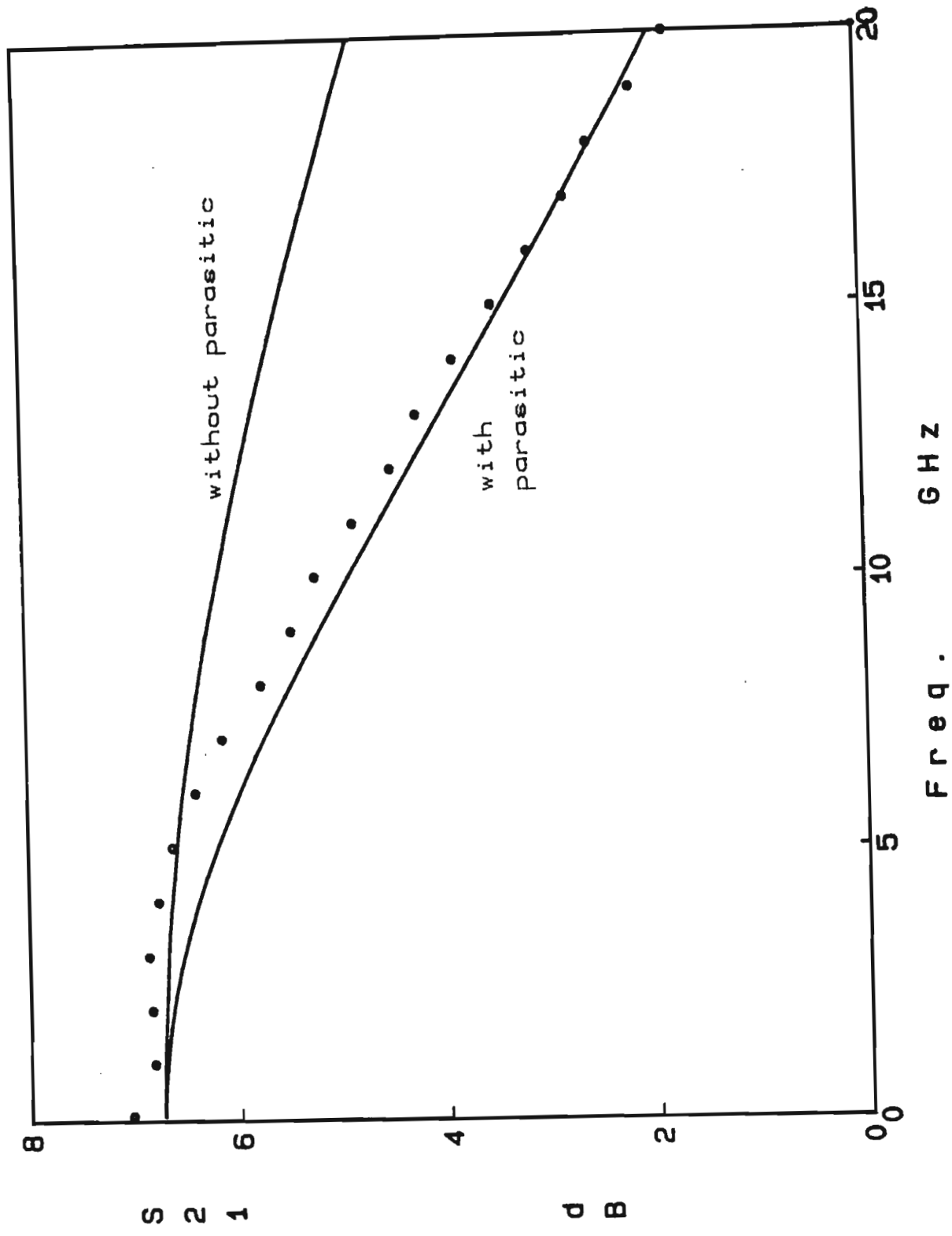


Figure 2.8. Comparison between simulated S_{21} for device with, and without, parasitic gate-source capacitance. Measured points are for device 5453, form DB109b.

magnitude of S_{21}, for device 5453, for the device as fabricated and for a device where the parasitic capacitance has been removed. The measured S_{21} is also shown. The useful performance of the intrinsic device is clearly demonstrated. Details of this calculation, the RF model and the other S-parameters are discussed in detail in Chapter 4. The calculated MAG for the full device was greater than 6 dB at 30 GHz (Figure 4.5).

These results illustrate the considerable potential of the IHEMT for both digital and analogue applications. It is probable that, with further technological development, the IHEMT will equal or exceed the best performance figures obtained with the CHEMT.

2.4 CONDUCTION-BAND DISCONTINUITY

The magnitudes of the valence-and conduction-band energy discontinuities (ΔE_v and ΔE_c) are two of the most important parameters determining the electrical properties of a heterojunction. In HEMTs, the conduction band discontinuity determines, in part, the amount of charge transfer from the n-AlGaAs to form the 2DEG (sheet concentration), and the threshold voltage. In addition, the conduction band discontinuity forms a barrier to real space transfer of 2DEG electrons from the 2DEG into the n-AlGaAs layer (to be discussed in detail in Chapter 3). In this section, the dependence of ΔE_c and ΔE_v on the aluminium mole fraction, x, in $GaAs/Al_xGa_{1-x}As$ heterojunctions is discussed.

The GaAs/AlGaAs heterojunction has a band line-up termed 'type I'. A type I heterojunction is one where the conduction band edge in the larger-bandgap material (in this case, AlGaAs) is higher in energy than the conduction band in the smaller-bandgap material (GaAs), and the valence band edge in the larger-bandgap material is lower in energy than the corresponding band in the smaller-bandgap material. Therefore, in a type I heterojunction, the sum of the conduction-and valence-band discontinuities is equal to the energy bandgap difference between the two materials:

$$\Delta E_g = \Delta E_c + \Delta E_v \qquad (2.3)$$

The energies of the conduction band minima in AlGaAs, relative to the top of the valence band at the Γ point (K=0,0,0), as a function of aluminium mole fraction [2.17], are shown in Figure 2.9. For mole fractions less than aproximately 0.45, AlGaAs has a direct bandgap. For larger mole fractions, the alloy is indirect, with the X-valley having the lowest energy. The L-valley is never the lowest energy valley and plays no part in this discussion.

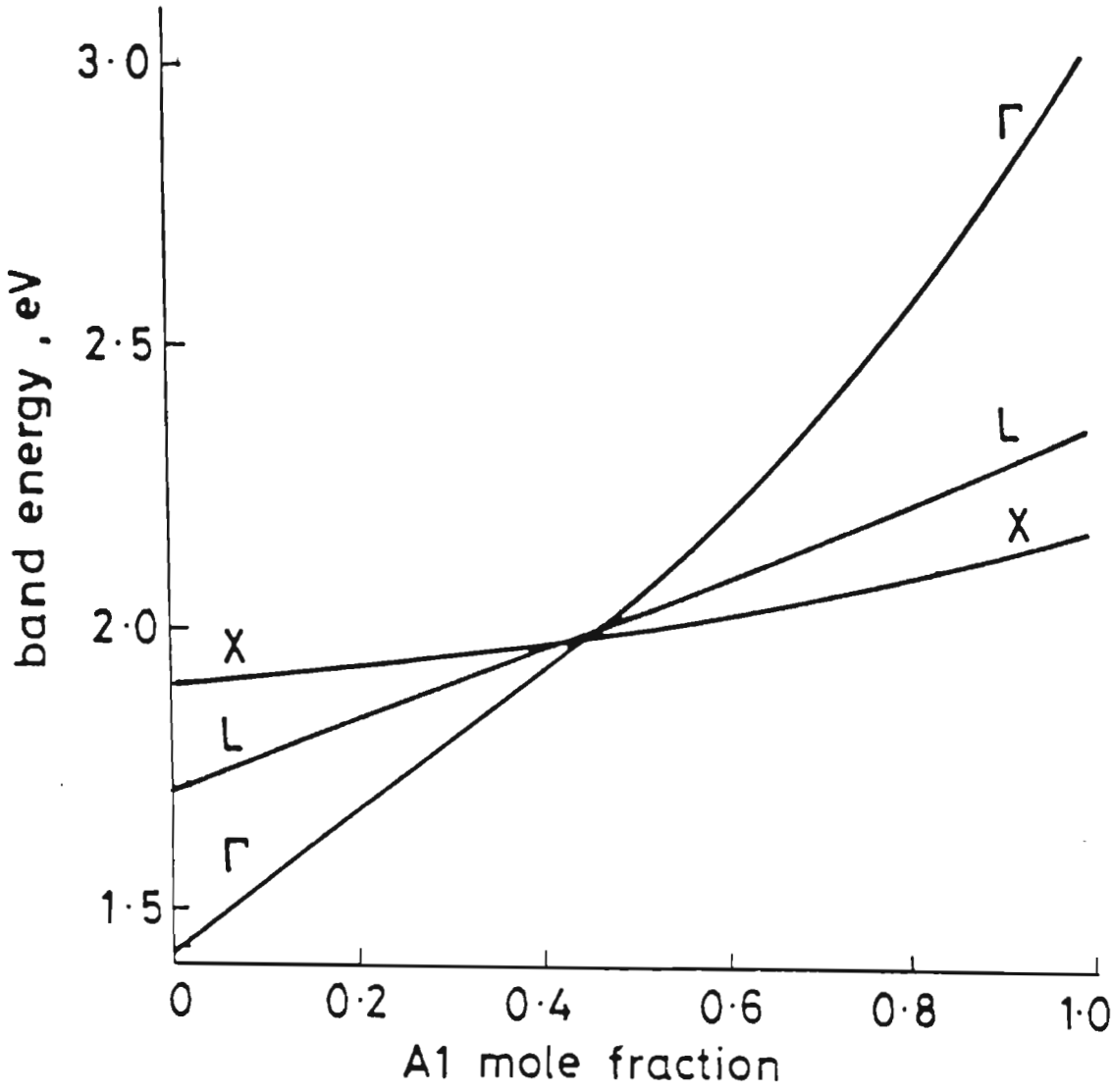


Figure 2.9 Energies of the AlGaAs conduction-band minima as a function of aluminium mole fraction, relative to the top of the valence band at the gamma point.

The ratio of the valence-band discontinuity to direct bandgap (Γ) difference, has been the subject of many recent studies (see, for example, refs. 2.18-2.20 and further references therein). With one exception [2.21], the ratio has been found to be independent of mole fraction, with the majority of values in the range 0.33 to 0.41. The valence band discontinuity as a function of mole fraction is shown Figure 2.10, assuming a ratio of 0.35. Several reported measured values are also shown in the Figure.

Since the sum of ΔE_v and ΔE_c must equal the total energy discontinuity, ΔE_g (equation 2.3), the ratio $\Delta E_c/\Delta E_g$ is equal to 0.65, assuming $\Delta E_v/\Delta E_g$ is 0.35. This is shown, for the direct gap region, in curve B, Figure 2.10. Several measured values are also shown for comparison.

For mole fractions greater than 0.45, the Γ-valley energy increases rapidly with increasing mole fraction (Figure 2.9) and hence the valence-band discontinuity also increases rapidly. The GaAs Γ-valley to AlGaAs Γ-valley discontinuity also shows a marked increase in this range. In quantum well structures, such as superlattices, quantum confinement occurs between bands with the same symmetry. Hence, in a GaAs/AlAs quantum well, the barrier height is approximately 0.65 times the direct energy bandgap difference between GaAs and AlAs, which (using Figure 2.9) is approximately:

$$\Delta E_c \simeq 0.65\ (3.0-1.4) = 1\ \text{eV} \qquad (2.4)$$

This barrier height is evidenced in many reported photoluminescence experiments, which measure the energy difference between confined electron and hole states in GaAs/AlAs quantum well structures.

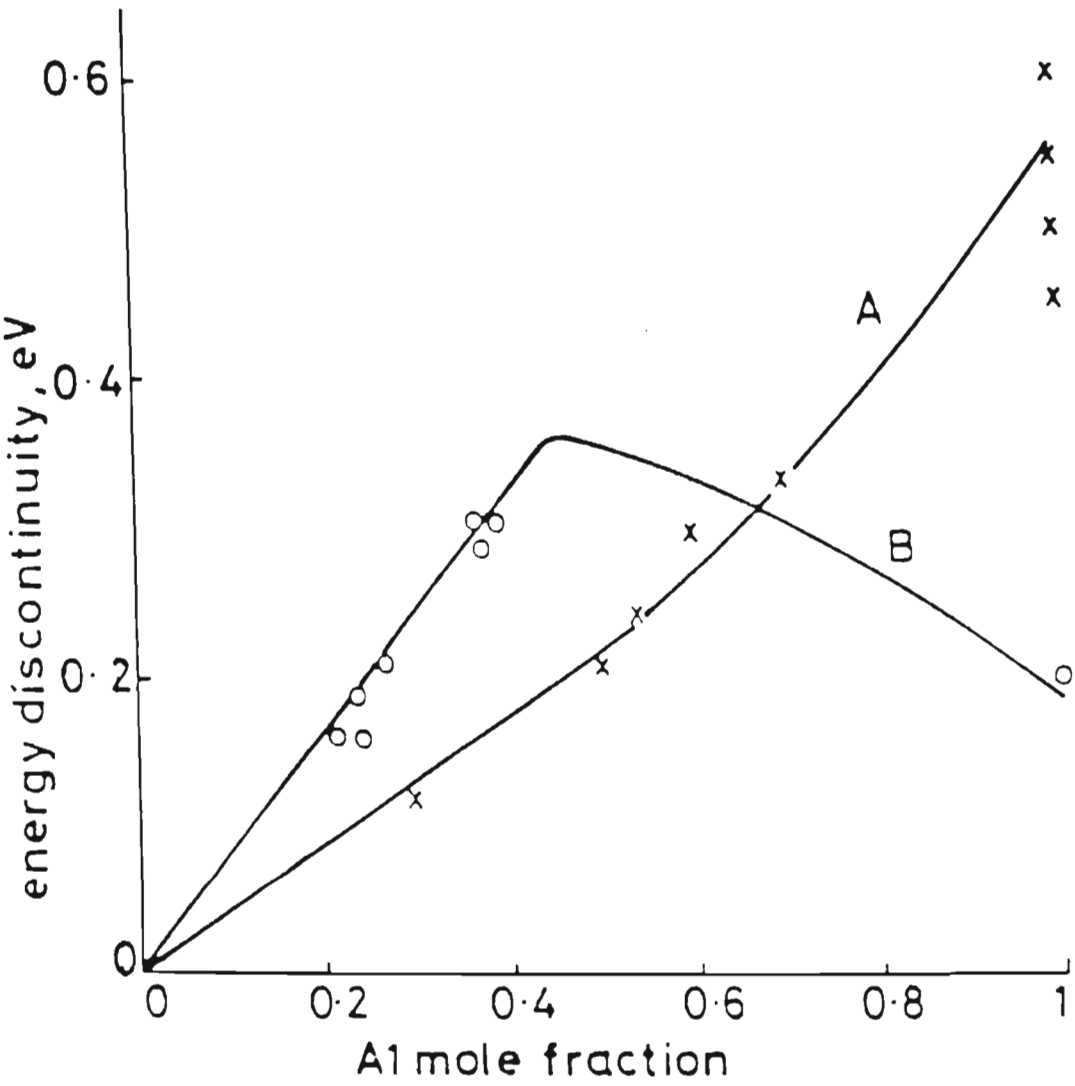


Figure 2.10 Dependence of valence- and conduction-band discontinuities on aluminium mole fraction. Some measured values from the open literature are also plotted.

In the indirect region ($x > 0.45$), the bandgap, determined by the X-valley, increases slowly in comparison to the rapid increase in the Γ-valley energy. Since equation 2.3 remains valid, the conduction band discontinuity to the X-valley in the AlGaAs must decrease with increasing mole fraction. This dependence is shown in curve B, Figure 2.10.

It is this discontinuity which is important in determining the electrical characteristics of HEMTs. In doped AlAs (or any AlGaAs alloy with $x > 0.45$) the majority of electrons will be found in the lowest energy states available, namely, in the X-valley. Hence in determining the amount of electron transfer from AlAs to GaAs, it is the smaller discontinuity which is important, rather than the Γ-Γ discontinuity.

This smaller discontinuity also forms the barrier to hot electron real space transfer out of the 2DEG, into the AlGaAs layer. Electrons in the 2DEG may gain sufficient energy from the longitudinal (source-drain) electric field to be scattered, by a phonon, towards the heterointerface and into the AlGaAs. This process, which is analagous to thermionic emission, does not require the conservation of the electrons crystal momentum because of the lack of translational symmetry at the heterojunction. Since the electrons are able to transfer to the X-valley, it is the smaller discontinuity, rather than the larger Γ-Γ one, which determines electron transport perpendicular to the heterojunction. Hence the maximum barrier height is obtained with an aluminium mole fraction of around 0.45.

The deduced dependence of conduction-band discontinuity on aluminium mole fraction (curve B in Feature 2.10), and published by the author ([2.22, 2.23], Appendices C and E), was contrary to the generally accepted dependence of increasing ΔE_c for increasing x for all mole fractions. However, subsequent published data has verified the deduced dependence.

Wang and Stern [2.24] estimated ΔE_c as around 0.2 eV, from the measured 2DEG density, for an n-AlAs/GaAs heterojunction. This point is shown on Figure 2.10. More recently, Maezawa et al. [2.25] reported values of 0.34, 0.32 and 0.26 for mole fractions of 0.48, 0.62 and 1.0 respectively, obtained from I-V measurements of current transport through GaAs/AlGaAs/GaAs structures.

With the dependences of Figure 2.10, staggered AlGaAs/AlAs (type II) heterojunction is possible. In a type II heterojunction, both the conduction and valence bands in one material (in this case AlAs) are lower than their respective bands in the other material (AlGaAs). For mole fractions in the AlGaAs layer of greater than approximately 0.3 (where the conduction band energies are almost the same in both layers) a type II heterojunction will result. This has been verified by the observation of electron transfer from the n-AlGaAs to the AlAs layer in an n-AlGaAs heterojunction [2.26], and from the photoluminescence spectra of an AlGaAs/AlAs multi-quantum well structure [2.27]. If ΔE_c increased with increasing mole fraction, for all mole fractions, a type II heterojunction could not be formed using AlGaAs.

While the qualitative dependence of ΔE_c on mole fraction is now widely accepted, the exact quantitative dependence is not well established. Measured values of 0.6 ± 0.2 eV [2.28] and 0.39 ± 0.07 eV [2.29] have been reported for ΔE_v at GaAs/AlAs heterojunctions and the corresponding values for ΔE_c from Wang (2.24] and Maezawa [2.25] differ by more than 0.05 eV. In addition, the accuracy of Casey and Pa nish's data [2.17] for the bandgap dependence on mole fraction, which has formed the basis of many published calculations (including the author's [2.30] and Appendix F), has also been called into question [2.21, 2.31]. More work is clearly required to establish the exact quantiative dependence of ΔE_c on aluminium mole fraction.

2.5 SUPERLATTICE-BASED HEMTs

From the discussion in the last section, it could be inferred that the optimum mole fraction for the AlGaAs layers in HEMTs is around 0.45, where the conduction band discontinuity is largest. However, the use of AlGaAs layers with mole fractions around this value introduces several undesirable effects in HEMT electrical performance. These effects include large donor activation energy, persistent photoconductivity and I-V collapse at cryogenic temperatures, and are primarily related to the large concentration of deep states present in the AlGaAs. (A more detailed discussion of these effects may be found in Section 7 of the review paper by the author [2.3], Appendix A).

The dominant donor activation energy increases from less than 10 meV for mole fractions less than 0.23, to a maximum of around 160 meV at a mole fraction of 0.48 [2.32]. Hence, the effect increasing the mole fraction beyond 0.23 to increase ΔE_c and hence the amount of charge transfer to the 2DEG, is negated by the rapid increase in donor ionisation energy. However, a large value of ΔE_c is still desirable to minimise the number of hot 2DEG electrons scattered into the AlGaAs layer.

One possible solution, to obtain low donor activation energy and a large barrier to hot electrons, is to use an aluminium mole fraction of around 0.2 in the doped AlGaAs layer and a mole fraction around 0.45 in the undoped spacer. Practical CHEMTs of this type have been fabricated at Cornell University [2.33]. One possible problem with these structures is that if the spacer is thin (around 2 nm), to obtain a high 2DEG sheet concentration, electrons may be able to tunnel through the spacer, into the lower mole fraction AlGaAs. This would remove any advantage of the high mole fraction spacer.

Another possible solution is to replace the bulk AlGaAs layers with a superlattice consisting of thin alternate layers of GaAs and AlAs (or AlGaAs with high Al mole fraction). If the GaAs layers are doped, the donor ionisation energies are low (around 10 meV) and there is almost complete donor ionisation even at 77K [2.34-2.36]. The electrons can move through the superlattice and accumulate in the bulk GaAs layer to form a 2DEG, in the same manner as in a HEMT with a bulk n-AlGaAs layer.

The effective bandgap of the superlattice (the energy difference between the lowest quantum levels for electrons and holes) is a function of the widths of the GaAs (well) and AlAs (barrier) layers. The donor activation energy, however, has been found to remain small, less than 10 meV, for bandgaps (at 300K) between 1.78 and 1.93 eV [2.35]. This offers the prospect of obtaining a large barrier to hot electrons, and low activation energy by varying the layer widths.

The choice of superlattice layer widths for a CHEMT involves a compromise which is not required for an IHEMT. For a given GaAs well width, the superlattice bandgap (and hence the barrier height for 2DEG electrons) increases as the AlAs layer width is increased. However, since only the GaAs wells are doped, the average doping density of the superlattice decreases as the AlAs width is increased. In the CHEMT, a high doping density is required to minimise the gate to 2DEG spacing and so maximise the transconductance. Hence in CHEMTs the choice of layer widths involves a compromise between transconductance and barrier height. The gate-2DEG spacing in the IHEMT does not depend on the doping density of the doped superlattice (or AlGaAs) layer, and hence the superlattice layers can be

chosen to maximise the barrier height to hot electron real space transfer. Therefore, in principle the IHEMT has the prospect of enhanced performance over the CHEMT in superlattice-based HEMTs, but this has yet to be realised in practical devices.

Superlattice-based IHEMTs have been fabricated at GEC and have exhibited transconductances of 180 mS/mm and gate-drain breakdown voltages of over 15V, in 0.5 μm gate length devices. These devices, which are discussed in more detail in Appendix C and D [2.23, 2.37], have not yielded as good results as those of the bulk AlGaAs-based IHEMTs discussed in Section 2.3. Clearly, however, there is scope for further work in the area of superlattice based devices.

2.6 SUMMARY

In this chapter some of the fundamental principles of IHEMT and CHEMT have been drawn and several potential advantages of the IHEMT discussed. Practical results on IHEMTs fabricated at GEC Hirst Research Centre have been presented, which indicate the considerable potential of the IHEMT. The RF measurements constitute the first published results on the RF performance on IHEMTs.

The dependence of the conduction-band discontinuity on aluminium mole fraction has been discussed and shown to attain a maximum value near the direct-indirect crossover. Finally, superlattice-based HEMTs were discussed as an alternative to AlGaAs-based devices.

In the next chapter, a new process-independent analytic model for IHEMTs and CHEMTs is presented.

CHAPTER 2 REFERENCES

2.1 Mimura T., Hiyamizu S., Fujii T. and Nanbu K., "A New Field-Effect Transistor with Selectively Doped GaAs/n-AlxGa1-xAs Heterojunctions", Jap. J. Appl. Phys., Vol. 19, No. 5, May 1980, pp. L225-L227

2.2 Ladbrooke P.H., Hill A.J., Frost M.S., Caddick C.J., Sansom C. and Kerr T.F., "Microwave Performance of 0.5 um Inverted HEMT", submitted to IEEE Elec. Dev. Letts.

2.3 Hill A.J. and Ladbrooke P.H., "High Electron Mobility Transistors (HEMTs) - A Review", GEC Jnl. Res., Vol. 4, No. 1, 1986, pp. 1-14

2.4 Shah N.J., Pei S-S, Tu C.W. and Tiberio R.C., "Gate-Length Dependence of the Speed of SSI Circuits Using Submicrometer Selectively Doped Heterostructure Transistor Technology", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 543-546

2.5 Watanabe Y., Kajii K., Nishiuchi K., Suzuki M., Hanyu I., Kosugi M., Odani K., Shibatomi A., Mimura T., Abe M. and Kobayashi M., "A High Electron Mobility Transistor 1.5K Gate Array", Proc. 1986 ISSCC, pp. 80-81

2.6 Berenz J., Nakano K., Hsu T-I., Goel J., "HEMT 60 Ghz Amplifier", Elec. Lett., Vol. 21, No. 22, Oct. 1985, pp. 1028

2.7 Smith P.M., Chao P.C., Duh K.H.G., Lester L.F. and Lee B.R., "94 GHz transistor amplification using an HEMT", Elec. Lett., Vol. 22, No. 15, July 1986, pp. 780-781

2.8 Duh K.H.G., Chao P.C., Smith P.M., Lester L.F., Lee B.R., "60 GHz Low-noise high-electron-mobility transistors", Elec. Letts., Vol. 22, No. 12, June 1986, pp. 647-649

2.9 Camnitz L.H., Tasker P.J., Lee H., van der Merwe D. and Eastman L.F., "Microwave Characterisation of Very High Transconductance MODFET", 1984 GaAs and Related Compounds, Inst. Phys. Conf. Ser. 74, pp.

2.10 Cirillo N.C., Shur M.S. and Abrokwah J.K., "Inverted GaAs/AlGaAs Modulation-Doped Field-Effect Transistors with Extremely High Transconductances", IEEE Elec. Dev. Lett., Vol. EDL-7, No. 2, Feb. 1986, pp. 71-74

2.11 Drummond T.J., Morkoc H. and Cho A.Y., "Dependence of electron mobility on spatial seperation of electrons and donors in AlxGa1-xAs/GaAs heterostructures", J. Appl. Phys., 52(3), March 1981, pp 1380-1386

2.12 Lee K., Shur M.S., Drummond T.J. and Morkoc H., "Current-Voltage and Capacitance-Voltage Characteristics of Modulation-Doped Field-Effect Transistors", IEEE Trans. Elec. Dev., Vol. ED-30, No. 3, March 1983, pp 207-212

2.13 Das M.B., "A High Aspect Ratio Design Approach to Millimeter-Wave HEMT Structures", IEEE Trans. Elec. Dev., Vol. ED-32, No. 1, Jan. 1985, pp. 11-17

2.14 Fischer R., Masselink W.T., Sun Y.L., Drummond T.J., Chang Y.C., Klein M.V., Morkoc H. and Anderson E., "Improvement of the inverted GaAs/AlGaAs heterointerface", J. Vac. Sci. Technol., B 2 (2), Apr. 1984, pp. 170-174

2.15 Dambkes H., Brockerhoff W., Heime K., Ploog K., Weimann G. and Schlapp W., "Optimisation of Modulation-Doped Heterostructures for TEGFET Operation at Room Temperature", Elec. Lett., Vol. 20, No. 15, July 1984, pp. 615-618

2.16 Hikosaka K., Hirachi Y. and Abe M., "Microwave Power Double-Heterojunction HEMT's", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 583-589

2.17 Casey H.C. and Panish M.B., "Heterostructure Lasers, Part A: fundamental principles", Academic Press, New York, 1978

2.18 Watanabe M.O., Yoshida J., Mashita M., Nakanisi T. and Hojo A., "Band discontinuity for GaAs/AlGaAs heterojunction determined by C-V profiling technique", J. Appl. Phys., Vol. 57(12), June 1985, pp. 5340-5344

2.19 Arnold D., Ketterson A., Henderson T., Klem J. and Morkoc H., "Electrical characterization of GaAs/AlGaAs semiconductor-insulator-semiconductor capacitors and application to the measurement of the GaAs/AlGaAs discontinuity", J. Appl. Phys., Vol. 57(8), April 1985, pp. 2880-2885

2.20 Batey J., Wright S.L. and DiMaria D.J., "Energy band-gap discontinuities in GaAs:(Al,Ga)As heterojunctions", J. Appl. Phys., Vol. 57(2), pp. 484-487, Jan. 1985

2.21 Batey J. and Wright S.L., "Energy band alignment in GaAs:(Al,Ga)As heterostructures: The dependence on alloy composition", J. Appl. Phys., Vol. 59(1), Jan. 1986, pp. 200-209

2.22 Hill A.J. and Ladbrooke P.H., "IHEMT, prospects and uncertainty", IEE Colloquium on GaAs ICs, London, 14 March 1985

2.23 Hill A.J., Ladbrooke P.H., Ransome S. and Westwood D., "Large Signal Modelling and Practical Performance of Inverted HEMT", Proc. IEEE/Cornell Conf., July 1985, pp. 128-135

2.24 Wang W.I. and Stern F., "Valence band offset in AlAs/GaAs heterojunctions and the empirical relation for band alignment", J. Vac. Sci. Technol. B 3(4), July 1985, pp. 1280-1284

2.25 Maezawa K., Mizutani T. and Yanagawa F., "Barrier Height in Indirect Bandgap AlGaAs/GaAs Hetero-Junction Determined with n-Semiconductor/Insulator/Semiconductor Diodes", Jap. J. Appl. Phys., Vol.25, No. 7, July 1986, pp. L557-L559

2.26 Drummond T.J. and Fritz I.J., "Modulation-doped (Al,Ga)As/AlAs superlattice: Electron transfer into AlAs", Appl. Phys. Lett., Vol. 47(3), Aug. 1985, pp. 284-286

2.27 Dawson P., Wilson B.A., Tu C.W. and Miller R.C., "Staggered band alignments in AlGaAs heterojunctions and the determination of valence-band offsets", Appl. Phys. Lett., Vol. 48(8), Feb. 1986, pp. 541-543

2.28 Kelly M.K., Niles D.W., Colavita E., Margaritondo G. and Henzler M., "Valence-band discontinuities at AlAs-based heterojunction interfaces", Appl. Phys. Lett., Vol. 46(8), April 1985, pp. 768-770

2.29 Katani A.D. and Bauer R.S., "Commutativity and transivity of GaAs-AlAs-Ge(100) band offsets", Phys. Rev. B, Vol. 33, No. 2, Jan. 1986, pp. 1106-1109

2.30 Hill A.J. and Ladbrooke P.H., "Dependence of conduction-band discontinuity on aluminium mole fraction in GaAs/AlGaAs heterojunctions", Elec. Lett., Vol. 22, No. 4, Feb. 1986, pp. 218-220

2.31 Kroemer H., "Band Offsets at Heterointerfaces: Theoretical Basis, and Review, of Recent Experimental Work", Supplement to MSS II, pp. 797-806, 1985

2.32 Chand N., Henderson T.J., Klem J., Masselink W.T., Fischer R., Chang Y-C and Morkoc H., "Comprehensive analysis of Si-doped $Al_xGa_{1-x}As$ (x=0 to 1): Theory and experiments", Phys. Rev. B., Vol. 30, No. 8, Oct. 1984, pp. 4481-4492

2.33 Huang J.C., Wicks G.W., Calawa A.R. and Eastman L.F., "Optimised HEMT structure with an $Al_{0.45}Ga_{0.55}As$ spacer and an $Al_{0.20}Ga_{0.80}As$ Doped Region", Elec. Lett., Vol. 21, No. 20, Sept. 1985, pp. 925-926

2.34 Baba T., Mizutani T., Ogawa M. and Ohata K., "High Performance (AlAs/n-GaAs Superlattice)/GaAs 2DEGFETs with Stabilized Threshold Voltage", Jpn. J. Appl. Phys., Vol. 23, No. 8, Aug. 1984, pp. L654-L656

2.35 Baba T., Mizutani T. and Ogawa M., "AlAs/n-GaAs superlattice and its application to high-quality two-dimensional electron gas systems", J. Appl. Phys., Vol. 59(2),, Jan. 1986, pp. 526-532

2.36 Arnold D., Henderson T.J., Klem J., Fischer R., Kopp W., Ketterson A., Masselink W.T. and Morkoc H., "High performance inverted and large current double interface modulation-doped field-effect transistors with the bulk (Al,Ga)As replaced by a superlattice at the inverted interface", Appl. Phys. Lett., Vol. 45(8), Oct. 1984, pp. 902-904

CHAPTER THREE

HEMT DC MODEL

3.1 INTRODUCTION

This chapter describes an IHEMT model developed by the author to predict DC electrical characteristics from the material and structural parameters of the device. In the succeeding chapter, IHEMT RF performance is also predicted by using this model to calculate the RF equivalent circuit elements of the device.

While several models for predicting the DC I_{DS}-V_{DS} characteristics of CHEMTs have been published, none for the IHEMT have appeared in the literature to date. The model presented here also includes several physical effects not included in other models, such as the extension of the active region beyond the drain edge of the gate and the effect of occupied surface states in the source-gate and drain-gate regions. The increase in drain-source current arising from real-space transfer of carriers out of the 2DEG into the underlying AlGaAs layer (IHEMTs) or into the substrate (CHEMTs) is also included.

The model uses approximations (to be discussed) to the detailed physical processes in the device. The result is a set of analytic expressions which must be solved simultaneously to obtain a solution for the currents, fields and potentials within the device. Implementation of the model in a computer program has resulted in a device simulation package which requires short execution times. Use of the program has provided new insight into HEMT performance by allowing the effect of structural and material parameter variations on DC characteristics to be simulated. The short execution times of the program will allow the model to be incorporated into circuit design software in the future.

A similar analytic approach to that described above has been used successfully by Ladbrooke for the GaAs MESFET (see Appendix G). These approaches avoid the large computation times characteristic of numerical 2D [3.1,3.2] and Monte Carlo HEMT [3.3, 3.4] simulators, neither of which necessarily provide more accurate or useful results than the analytic approach [3.5].

The model presented here is, however, considerably more complex than simple empirical curve-fitting models [3.6, 3.7, 3.8]. These latter models have severe limitations because they are not physically based and hence cannot predict the effect of structural and material parameter variations on device electrical characteristics.

This chapter is divided into six sections. In sections two and three, the details of the IHEMT model derivation are presented. In section four, results of simulations using the model and comparisons between simulated and measured DC characteristics are presented. Section five contains details of the extension of the analysis to model CHEMTs. Lastly, section six summarises the major aspects of the work and suggests directions for future development. A list of the references cited in this chapter appears after section six.

3.2 THREE-REGION MODEL DESCRIPTION

3.2.1 Introduction

A simplified cross-section through an IHEMT operating in the saturation regime is shown in Figure 3.1. The vertical scale is greatly exaggerated; in a typical device, the source-drain spacing would be around 3 micron and the total active device layer thickness only around 0.15 micron. The intrinsic device is divided into three regions (labelled I,II and III) which connect, via access regions, with the source and drain ohmic contacts.

The two-dimensional electron gas, 2DEG, is formed by the accumulation of electrons from the underlying n-AlGaAs layer. Although the thickness of the 2DEG is small (around 15 nm) and does not vary greatly with changes in sheet density, it is helpful in understanding device operation to represent the variation of sheet carrier density (n_s) along the channel in a manner similar to that used to illustrate the depletion edge in a conventional MESFET.

The gate electrode is shown in a gate recess etched in the heavily doped GaAs cap layer. Cap layers are often included in IHEMT and CHEMT designs to lower access and contact resistances, and in the IHEMT, to prevent depletion by the free surface potential of the 2DEG in the access regions. However, as will be seen in the following chapter, the cap layer may degrade the RF gain of IHEMT devices.

In the following subsection, the choice of velocity-electric field characteristic describing the 2DEG electron transport is discussed. In subsequent subsections, equations for the 2DEG current, as well as electric field and potential variations in the three intrinsic regions are derived.

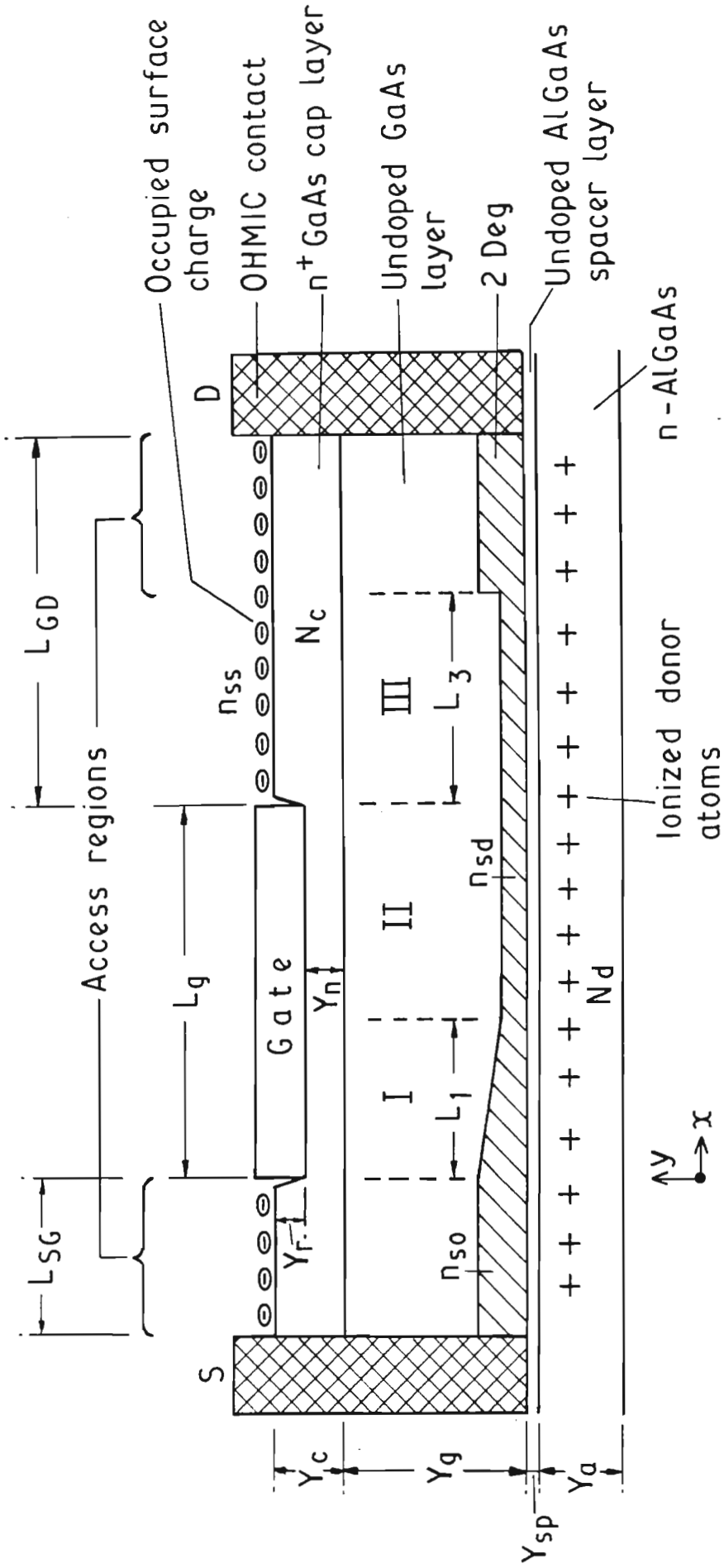


Figure 3.1 Simplified cross-section through the active region of an IHEMT operating in current saturation.

In this planar three-region model for the IHEMT, the approximation is made that the 2DEG and underlying n-AlGaAs layer can be treated as a charge sheet. This approximation is reasonable because the total thickness of these two layers is typically only around 20 nm.

Throughout this chapter, the device is assumed to be operating at 300 K.

3.2.2 Velocity-field characteristic

In spite of the intense interest in HEMTs, the dependence of 2DEG electron velocity on applied longitudinal electric field has only recently been experimentally determined on HEMT-like structures [3.9,3.10]. As a consequence, most HEMT models (including this one) have used the same velocity-field dependences as found in MOSFET and MESFET models.

The simplest form is the two-piece linear relation ((a) in figure 3.2), where the electron mobility is assumed constant for electric fields less than some critical field, F_c, and the electron velocity is assumed constant, v_s, for fields greater than the critical field. This characteristic was used in some of the first HEMT models [3.11] and in early versions of this model (Appendix C and [3.12]). The major limitation of this velocity-field characteristic is that it results in a 'knee' on the modelled I_{DS}-V_{DS} characteristics which is too abrupt and at too low a drain voltage in comparison with measured characteristics (Figures 4 and 5, Appendix C). In an attempt to avoid this problem, Lee et al. [3.13] used a three piece characteristic. While the abruptness was removed, the knee voltage was still too low.

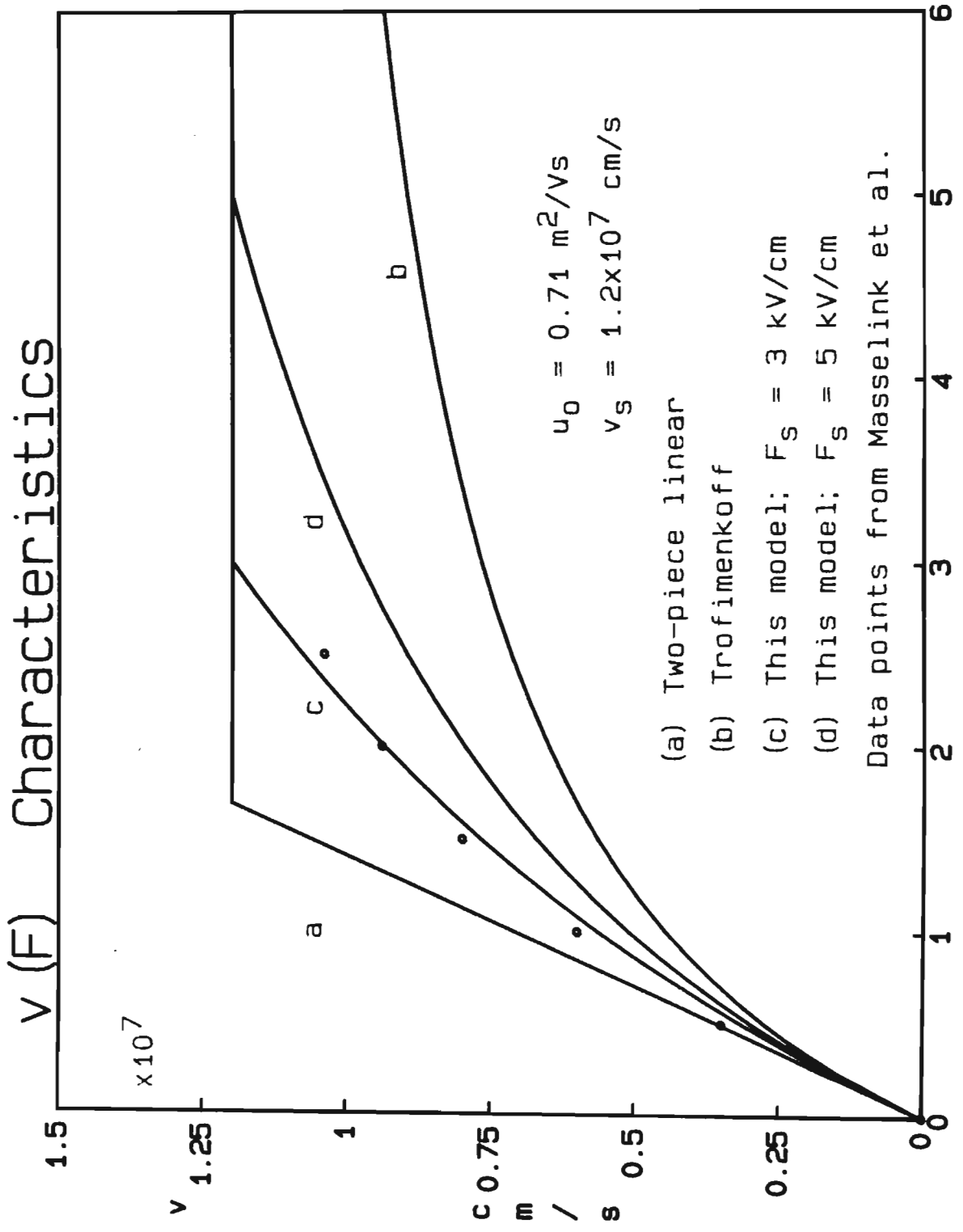


Figure 3.2. Graphs of several velocity vs. Electric field dependences discussed in the text.

The large low-field mobility of 2DEG electrons results in a rapid increase in kinetic energy of these carriers in applied fields. However, this increase is limited by the onset of spontaneous phonon emission from the hot carriers and consequently the electron mobility decreases as the electric field increases. To incorporate this field-dependent mobility effect in the model, a modified form of the expression developed by Trofimenkoff for silicon FET analysis [3.14] was used. The Trofimenkoff expression, which has been used directly by Hariu et al. [3.15] for GaAs MESFETs, gives the electron velocity as:

$$v = \mu_o F \, (1+PF)^{-1} \qquad (3.1a)$$

$$P = \mu_o . v_s^{-1} \qquad (3.1b)$$

where μ_o is the low field mobility, F is the applied electric field and v_s is the electron saturated drift velocity. This expression is shown as curve (b) in Figure 3.2.

One difficulty with this expression is that the velocity approaches the saturated velocity asymptotically, which means that it is impossible to define a point in the IHEMT channel where the carriers become velocity saturated. Such a point is required however, to determin the boundary between regions I and II in the intrinsic device (see next subsection). To overcome this problem, in this model the Trofimenkoff expression is modified to the form:

$$v = \mu_o F \, (1+PF)^{-1} \qquad F < F_s \qquad (3.2a)$$

$$P = \mu_o \, v_s^{-1} - F_s^{-1} \qquad (3.2b)$$

$$v = v_s \qquad F \geq F_s \qquad (3.2c)$$

where F_s is the critical field for velocity saturation. These expressions are plotted for two values of F_s in Figure 3.2 (curves (c) and (d)). If F_s is set equal to $\mu_o v_s^{-1}$, the relations reduce to the two-piece velocity field characteristic and if F_s is set to infinity, the Trofimenkoff expression is regained. These values for F_s represent the limits to the range of possible critical fields under this scheme.

For comparison, the data points from the recent measurements of Masselink et al.[3.9] are also plotted in Figure 3.2. These measurements were made on a CHEMT layer with a spacer layer, 3nm thick, under an n-AlGaAs layer with an aluminium mole fraction of 0.3 and a doping density of 10^{18} cm^{-3}. The measured low-field mobility was 7100 cm^2/Vs. It is not possible to predict how the measured velocity-field characteristic would change for any sample with different material parameters. No measured velocity-field characteristics for IHEMTs have been reported and nor have any data for fields beyond 2.5 KV/cm been published for either CHEMTs or IHEMTs.

When the complete IHEMT model is used for making comparisons with measured I_{DS}-V_{DS} characteristics, parameters for the velocity-field characteristic are determined as follows. The low field mobility is set equal to the value obtained from Hall measurements on samples from the same wafer as the device being simulated. This value is typically around 6000 cm^2/Vs in IHEMTs. In the absence of better information, the critical field is set equal to 8 kV/cm, which is around twice the generally accepted critical field in the two-piece approximation for bulk n-type GaAs. The saturated drift velocity is assumed in this work to be 10^7 cm/s; the reasons for this choice will be discussed in Section 3.4. The IHEMT velocity-field characteristic is not shown in Figure 3.2, which is plotted simply to show the general form of the equations 3.1 and 3.2.

3.2.3 Region I

Region I is defined in this model to be the intrinsic device region, starting at the source end of the gate, in which the longitudinal electric field, F_x, is less than the critical field for velocity saturation, F_s. This region is shown schematically in Figure 3.1 for a saturated device. For small applied drain-source voltages, this region may extend to the drain edge of the gate, (the linear mode of operation) in which case regions II and III disappear. Strictly, there would be some extension of the depleted region beyond the drain edge of the gate, but since F_s is small, the extension is negligible (see section 3.2.5, equation 3.29).

Since the electric field perpendicular to the 2DEG, F_y, is typically much greater than the longitudinal field, F_x, the fields in region I are assumed to be only y-directed. A one dimensional solution to Poisson's equation will be derived to determine the charge control relations in region I. The equation for the 2DEG drain-source current will then be derived to complete the analysis for this region.

Charge control

The combination of the doped AlGaAs layer and 2DEG are considered as a sheet of charge at the position of the 2DEG. For the sheet carrier densities typical in HEMTs, the centroid of the 2DEG is approximately 8 nm from the heterointerface [3.13] and hence the distance from the gate metal to the sheet charge is $Y_g + Y_n - 8$ nm, where Y_g is the undoped GaAs layer thickness and Y_n is the cap layer thickness under the gate (see Figure 3.1). The

maximum 2DEG sheet concentration, n_{so}, which occurs when the gate potential does not deplete the 2DEG, is assumed equal to the sheet density of the n-AlGaAs layer:

$$n_{so} = N_D Y_a \tag{3.3}$$

where N_D and Y_a are the doping (volume) density and thickness of the n-AlGaAs layer respectively. This is the optimum condition for an IHEMT; if $N_D Y_a$ is made greater than n_{so}, a parallel conduction path between source and drain is formed, which degrades the device output resistance in saturation. The dependence of n_{so} on layer thicknesses and doping has been studied more extensively than any other aspect of HEMT modelling (for example, [3.13, 3.16, 3.17, 3.18]) and will not be considered further here.

The combined sheet charge from the positive sheet charge of the ionized n-AlGaAs donors, and the negative sheet charge of the 2DEG is

$$n = n_{so} - n_s \tag{3.4}$$

where n_s is the 2DEG sheet density. This sheet charge is always greater than or equal to zero because the maximum value of n_s is n_{so}. The field arising from this net positive charge terminates on the gate electrode and is given by

$$F_y = \frac{q}{\varepsilon} (n_{so} - n_s) \tag{3.5}$$

where q is the electronic charge and ε is the permittivity of GaAs. Integrating from y = 0 to d = $Y_g + Y_n - 8$ nm (d is the distance from the charge sheet to the gate electrode), yields the gate potential:

$$V_G = - \frac{q}{\varepsilon} (n_{so} - n_s) d \qquad (3.6)$$

Adding in the effect of the cap layer charge and the Schottky barrier height, ϕ_m, provides the relationship between n_s and the applied gate-source voltage (for very low drain bias) as:

$$V_{GS} = - \frac{q}{\varepsilon} (n_{so} - n_s) d - \frac{q}{2\varepsilon} N_c Y_n^2 + \phi_m \qquad (3.7)$$

where N_c is the doping (volume) density in the cap layer.

The threshold voltage, defined as the applied gate-source potential at which the device is just switched off (often called pinch-off voltage) is then given by setting n_s to zero in (3.7):

$$V_T = - \frac{q}{\varepsilon} . n_{so} d - \frac{q}{2\varepsilon} N_c . Y_n^2 + \phi_m \qquad (3.8)$$

Substituting (3.8) in (3.7) and rearranging gives an expression for the 2DEG sheet density in terms of the applied gate-source voltage:

$$n_s = \varepsilon \frac{(V_{GS} - V_T)}{qd} \qquad (3.9)$$

This simple charge control model has been shown to be sufficiently accurate IHEMT [3.19], but is less accurate for CHEMT charge control near pinchoff. The CHEMT case will be discussed further in Section 3.5.

The gate-source voltage at which n_s reaches its maximum value n_{so} is given, from (3.7) and (3.8), by:

$$V_{GSM} = V_T + \frac{q}{\varepsilon} n_{so} d \qquad (3.10)$$

Applying gate-source voltages greater than V_{GSM} results in the n-AlGaAs layer becoming only partially depleted which in turn results in parallel conduction between source and drain. This parasitic MESFET effect is similar to that found in CHEMTs [3.20] and leads to a dramatic reduction in device transconductance. This effect is not considered further here because it is not usual to operate IHEMTs in this regime. Still further increase in V_{GS} may forward bias the Schottky junction and result in gate current flow from gate to source.

As can be seen from equations 3.7 and 3.8, the presence of the cap layer simply results in a decrease in the threshold voltage and an increase in the gate-2DEG spacing.

Drain-source current

The derivation of the drain-source current in region I follows a similar method to that found in many FET models. The drain-source current is given by:

$$I = qn_s(x)v(x)\ Z \tag{3.11}$$

where Z is the device gate width. $n_s(x)$ and $v(x)$ are the 2DEG sheet concentration and electron velocity at any point x along the 2DEG respectively. Assuming the source and drain resistances are zero, and defining V(x) as the potential at x in the 2DEG relative to the potential at the source end of region I, equation 3.9 can be recast as:

$$n_s(x) = \frac{\varepsilon\ [V_{GS} - V_T - V(x)]}{qd} \tag{3.12}$$

The electron velocity is obtained from equation 3.2a as:

$$v(x) = \mu_o\ F(x)\ \{1+PF(x)\}^{-1} \tag{3.13}$$

Substituting dV/dx for F(x) and combining equations 3.11 to 3.13 yields

$$I(dx + PdV) = \frac{\varepsilon\mu_o Z}{d}\ [V_{GS}-V_T-V(x)]\ dV \tag{3.14}$$

Integrating (3.14) and rearranging results in

$$I = \frac{\varepsilon\mu_o Z}{dL_1}\ [(V_{GS}-V_T)V_{DS}-0.5V_{DS}^2]\ \left[1+\frac{PV_{DS}}{L_1}\right]^{-1} \tag{3.15}$$

where V_{DS} is the applied drain-source voltage and L_1 is the length of region 1. Setting P to zero in (3.15) yields the standard expression for drain current assuming a constant mobility.

When the source resistance, Rs, and drain resistance, R_D, are not equal to zero, the effective gate-source voltage across the intrinsic device is reduced by the voltage across the source resistance, and the effective drain-source voltage across the intrinsic device is reduced by the voltage across the sum of the source and drain resistances.

Substituting

$$V_{GS} - IR_S \quad \text{for } V_{GS} \tag{3.16a}$$

and $$V_{DS} - I\,(R_S + R_D) \text{ for } V_{DS} \tag{3.16b}$$

in (3.15) yields, after lengthy algebra, a quadratic equation in I:

$$aI^2 + bI + C = 0$$

where $$a = 0.5K\,(R_S^2 - R_D^2) + P_1\,(R_S + R_D)$$

$$b = -\,[K(R_S+R_D)\,(V_{GS}-V_T) + (P_1 - KR_D)\,V_{DS} + 1]$$

$$c = K[V_{GS}-V_T)\,V_{DS} - 0.5\,V_{DS}^2]$$

$$K = \frac{\varepsilon\mu_o Z}{dL_1}$$

and $$P_1 = P/L_1 \tag{3.17}$$

The correct root is easily determined by selecting the root which is greater than zero and which yields a voltage drop across $R_S + R_D$ less than the applied V_{DS}. Unfortunately, it appears impossible to show algebraically that one root is always the correct one, independent of the choice of variables in the expressions for a, b and c.

The calculation of R_S and R_D from the material and structural parameters of the IHEMT is described in Section 3.2.6.

Onset of current saturation

In this model, the largest drain-source voltage for which the device can be considered to be operating in the linear mode, is that voltage which results in a peak longitudinal field equal to the critical field for velocity saturation, F_s. This peak field occurs at the drain edge of the gate. For greater drain-source voltages, the intrinsic device under the gate is divided into two regions, the boundary between them being that point in the channel where the longitudinal field equals F_s. Using equations 3.11 to 3.17, the peak field at the end of region 1 can be shown to be

$$F_x(L_1) = I.[KL_1(V_{GS}-V_T-V_{DS} - IR_D) + PI]^{-1} \qquad (3.18)$$

where I is found using equation 3.17.

The drain-source voltage at which $F_x(L_1)$ is equal to F_s occurs at a lower drain voltage than that calculated using the assumption that saturation occurs for dI/dV_{DS} equal to zero. This assumption has been used in some recent CHEMT models [3.21] but was originally shown to be incorrect by Grebene and Ghandi [3.22] in their FET analysis. The assumption is based on the idea that current saturation occurs because the channel is pinched off at the drain end of the gate.

Another approximation which is found in some models [3.13] is that saturation occurs for drain-source voltages greater than $F_s.L_g$, where L_g is the gate length. This approximation is based on the assumption that the longitudinal field is constant under the gate, whereas it actually increases towards the drain. This assumption also results in a drain-source voltage at the onset of saturation which is larger than predicted by (3.18).

The 'exact' equation (3.18) is used in this model because the decrease in the length of region I with increasing drain-source voltage in saturation (often termed 'channel length shortening') is a contributing factor to the output conductance of IHEMTs.

3.2.4 Region II

When the IHEMT is biased in saturation, the intrinsic device under the gate is considered to be divided into two regions, I and II (Figure 3.1). The boundary between these two regions at the charge sheet is the point where the longitudinal field is equal to F_s, the critical field for velocity saturation. In region II, therefore, the electrons are assumed to be moving at their saturated drift velocity, V_s. Furthermore, the sheet density of the carriers in this region is assumed constant, given by equation (3.12) as:

$$n_{sd} = \frac{\varepsilon\ (V_{GS} - V_T - V_1)}{qd} \tag{3.19}$$

where V_1 is the channel potential at the boundary between regions I and II. A similar method has been used in some MESFET models [3.22, 3.23].

The drain-source current is simply obtained from the general expression $I = qn_s vZ$ as

$$I = \frac{\varepsilon}{d}\ (V_{GS} - V_T - V_1) Z v_s \tag{3.20}$$

Clearly this current must equal the current in region I, which is calculated by substituting $V_1 - IR_s$ for V_{DS} and $V_{GS} - IR_s$ for V_{GS} in (3.15), which yields a modified form of (3.17):

$$aI^2 + bI + C = 0$$

$$a = 0.5KR_s^2 + P_1R_s$$

$$b = -(KR_s(V_{GS}-V_T) + P_1V_1 + 1)$$

$$c = K[V_{GS}-V_T).V_1 - 0.5\ V_1^2] \qquad (3.21)$$

It can be shown that the boundary condition $F_x = F_s$ at the interface between regions I and II is equivalent to the current continuity requirement that the current in region II (equation 3.20) is equal to that in region I (equation 3.21)

Electric field distribution

In region II, the electric field can no longer be considered only y-directed. The longitudinal field along the charge sheet in region II is required in order to estimate the hot electron temperature for the substrate current calculation (section 3.3) and to provide a boundary condition for the region III calculation (see next subsection).

To find the longitudinal, F_x, and transverse, F_y, field components, a solution for Poisson's equation:

$$\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} = \frac{\rho}{\varepsilon} \qquad (3.22a)$$

is found. The volume charge density, ρ, is obtained from the sheet charge as

$$\rho = \frac{q(n_{so}-n_{sd})}{Y_a} \qquad (3.22b)$$

The longitudinal electric field dependence on distance x from the source end of the gate is found (derivation in Appendix B) to be:

$$F_x(x) = -F_s \cosh\{(x-L_1)/Y\} \qquad (3.23a)$$

where $Y = \sqrt{Y_a.d}$ (3.23b)

The minus sign arises in (3.23a) because the field is directed in the -x direction. The potential along the sheet charge is easily found by integrating (3.23) and applying the condition $V(L_1) = V_1$, which yields:

$$V(x) = V_1 + F_s.Y \sinh \{(x-L_1)/Y\} \qquad (3.24)$$

Equations (3.23) and (3.24) are similar to those of Pucel et al. [3.23], derived for the GaAs MESFET, without the sheet charge assumption.

The peak longitudinal electric field acting on the charge sheet occurs under the drain edge of the gate and is given from (3.23) as:

$$F_p = - F_s \cosh \{(L_g-L_1)/Y\} \qquad (3.25)$$

where L_g is the gate length. The potential in the 2DEG at the drain edge of the gate, V_2, is found from (3.24):

$$V_2 = V_1 + E_s\ Y \sinh \{(L_g-L_1)/Y\} \qquad (3.26).$$

F_p and V_2 provide a set of boundary conditions for the solution of Poisson's equation in region III. The general shape of the field and potential profiles along the charge sheet in this region are shown in Figure 3.3.

3.2.5 Region III

Region III in this model is the region, starting at the drain edge of the gate, (see figure 3.1) where the longitudinal electric field decreases from its peak value at the drain edge of the gate, F_p, to approximately zero. This region is similar to the extension of the depletion region towards the drain in MESFETs. The depletion region extension has been included in some analytic MESFET models [3.24, 3.25] but has not been considered for the CHEMT or IHEMT, with the exception of this work [3.12].

It is important to include this region in HEMT (and FET) modelling because a significant fraction of the applied drain-source voltage may appear across it. Since the integral of the longitudinal field along the 2DEG must equal the drain-source voltage, omission of region III will lead to overestimation of the field in regions I and II. In addition, region II must be included to model the effect of occupied surface charge in the gate-drain region and to calculate the gate-drain capacitance.

As in region II, the positive charge sheet gives rise to both x-and y-directed fields. The 2DEG sheet charge in region III is assumed equal to that in region II and the electrons are assumed to still be moving at their saturated drift velocity. These assumptions satisfy the requirement of current continuity between two regions.

First, the equations for the field components for no cap layer are derived and then the extension to include the cap layer is discussed.

No cap layer

The field arising from the positive sheet charge is partitioned into x-and y-components by assuming that the latter component terminates on occupied surface charge of sheet density n_{ss}. Hence F_y is given by:

$$F_y = - \frac{q}{\varepsilon} \cdot n_{ss} \qquad (3.27)$$

The surface charge is usually negative charge and hence F_y is directed from the charge sheet towards the surface. F_y is assumed independent of x-coordinate in region III, that is, the surface charge is assumed to have a uniform distribution along the surface.

Solving Poisson's equation (3.22 and 3.23) in a similar manner to the solution for region II, the longitudinal field in region III is found to be (Appendix B):

$$F_x(x) = F_p + \frac{q}{\varepsilon} N_{eff}\, x \tag{3.28a}$$

where $$N_{eff} = N_D - \frac{n_{sd}}{Y_a} + \frac{n_{ss}}{Y_a} \tag{3.28b}$$

and where x is the distance from the drain edge of the gate. It can be seen from (3.28), that the effect of negative surface charge is that it reduces the rate of increase in longitudinal field by reducing the effective density of charge giving rise to x-directed field. At the drain end of region III, the longitudinal field is assumed equal to zero and hence from equation (3.28) the length of region III is given by:

$$L_3 = -F_p \left\{ \frac{q}{\varepsilon} N_{eff} \right\}^{-1} \tag{3.29}$$

Integration of (3.28) and using $V(0) = V_2$ gives the potential variation in region III as:

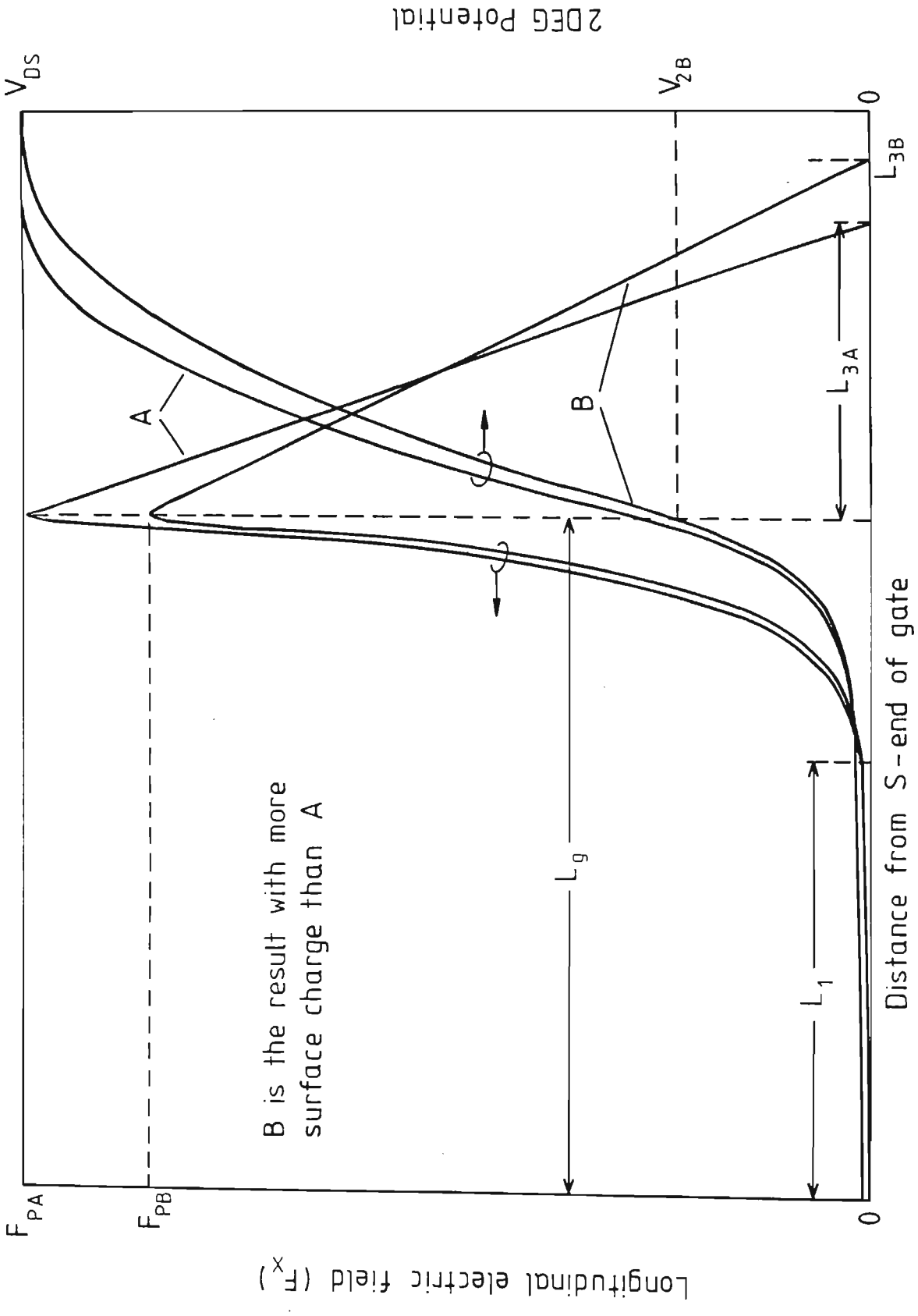
$$V(x) = V_2 - F_p x - \frac{q}{2\varepsilon} N_{eff}\, x^2 \tag{3.30}$$

The potential across region III is easily calculated by noting that F_x increases linearly across the region and so this potential is simply $-0.5 - 0.5\, E_p L_3$. The potential at the drain end of the region, V_3, is also given by $V_{DS} - IR_D$. Using these equations, the potential at the end of region III is given by

$$V_3 = V_2 + 0.5\, F_p^2 \left\{ \frac{q}{\varepsilon} N_{eff} \right\}^{-1} = V_{DS} - IR_D. \tag{3.31}$$

The shape of the electric field and potential profiles through the three intrinsic regions are illustrated in Figure 3.3, for the two values of surface charge. The dependences of the electrical parameters of surface charge are discussed further in section 3.4

Figure 3.3 Longitudinal electric field and potential distribution in the three intrinsic IHEMT regions, for a device operating in current saturation.



Cap layer effect

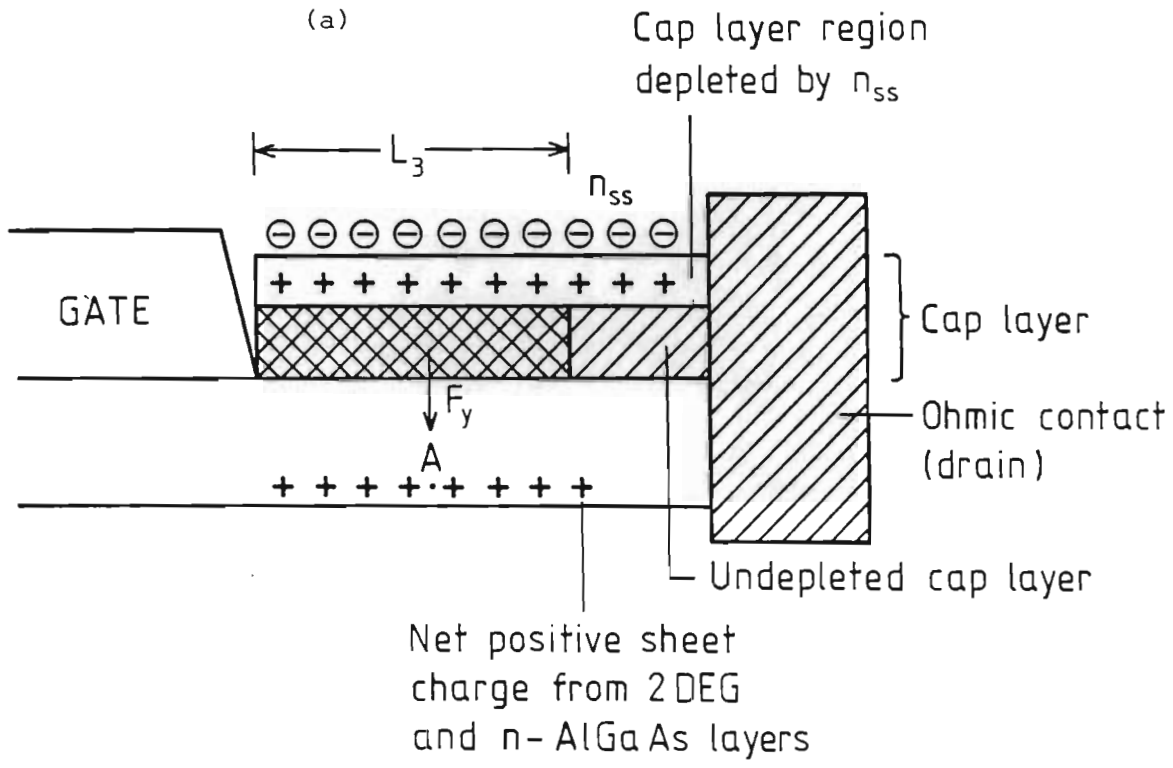
In order to incorporate the effect of the cap layer in the model, a qualitative argument is first presented to find the shape of the depletion region in the layer. Based on this argument, modifications to equations (3.28) to (3.31) are made.

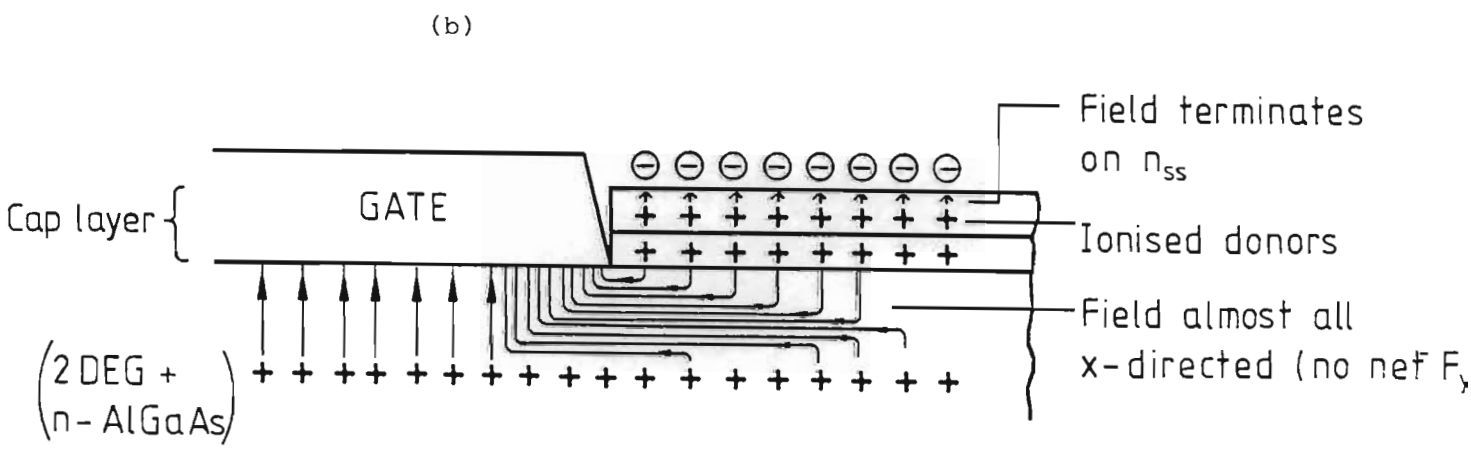
Figure 3.4a is a schematic diagram of the charge distributions in region III of a saturated device. The cap layer is shown partially depleted by the negative surface charge n_{ss}. (If the cap layer was fully depleted by the surface charge, then the situation would be equivalent to that for no cap layer, but with n_{ss} reduced by the sheet density of the cap, $N_c Y_c$). The net positive sheet charge $n_{so}-n_{sd}$ is also shown extending to a distance L_3 from the drain edge of the gate.

If the assumption is made that the cross-hatched area in the cap layer (Figure 3.4a) is undepleted, then because no current flows in the cap layer, this cross-hatched area must all have the same potential as the drain contact, V_D. At any point A along the charge sheet in region III, the potential must be less than V_D which in turn means that there must be a field, F_y, from the cross-hatched area towards A. However, this requires positive charge in that area, which is contrary to the initial assumption that the cross-hatched area is undepleted.

Therefore, the cap layer must be fully depleted throughout region III. The field arising from the positive charge in the cap layer in the cross-hatched area must terminate on the gate metal, in the same manner as field from the sheet charge of the 2DEG and n-AlGaAs layer. This is shown in Figure 3.4b. The planar nature of the IHEMT means that this field is almost entirely x-directed.

Figure 3.4. Schematic representat ions of the charge distributions in region III, as discussed in the text. The vertical scale is greatly exaggerated.





The effective surface charge in the presence of a cap layer is therefore

$$n'_{ss} = n_{ss} + N_c Y_c \tag{3.32}$$

and n'_{ss} can be substituted for n_{ss} in equation (3.28b). The addition of a cap layer introduces additional positive sheet charge in region III, which results in a decrease in L_3 (equation 3.29). For very thick cap layers, the length of region III is negligible and the IHEMT becomes similar to a self-aligned gate device.

3.2.6 Source and drain resistances

The source and drain resistances are both considered to consist of two parts, a contact resistance from the ohmic contact and an access resistance from the finite conductance of the access regions.

The contact resistance, R_c, cannot be calculated from the material parameters and therefore, in this model, it is assumed that R_c is known from measurements made on transmission line structures on the IHEMT wafer.

The sheet conductance of the 2DEG and cap layer (if one is present) is used to calculate the resistance of the access region. The 2DEG electrons are assumed to maintain their low-field mobility, μ_o, and the 2DEG sheet density is assumed to be n_{so}. The cap layer sheet density is given by equation 3.32 to account for the partial depletion of the cap in both source and drain access regions. The mobility of the cap layer electrons, μ_c, is assumed to be 2500 cm^2/Vs, a typical value in highly doped n-GaAs at 300K.

The source resistance, R_s, is given then by:

$$R_s = R_c + L_{SG}\,[qZ(n_{so}\mu_o + (N_c Y_c - n_{ss})\mu_c)]^{-1} \tag{3.33}$$

where L_{SG} is the source-gate spacing.

The drain resistance, R_D, is given by:

$$R_D = R_C + (L_{GD} - L_3).[qZ(n_{so}\mu_o + (N_cY_c - n_{ss})\,\mu_c)]^{-1} \qquad (3.34)$$

where L_{GD} is the gate-drain spacing. The length of this access region is reduced by the length of region III. Since the length of region III is bias dependent, it is clear that R_D is also bias dependent.

3.2.7 Summary

Equations for the 2DEG current and the field distribution along the 2DEG have been found in the three regions of the intrinsic device. For a given bias point in saturation several equations must be solved simultaneously to find the current and fields. When the device is operating in the linear region, it is only necessary to solve a quadratic (equation 3.17) to find the 2DEG current.

In the next section, the model is extended to include the effect of scattering of 2DEG electrons over the heterojunction into the AlGaAs layer. The inclusion of region III has resulted in a model which is a closed-charge system. By Gauss's Law, the total flux from region II should equal the charge enclosed in region III:

$$\oint_A \underline{D}.\underline{dA} = \int_v \rho\, dv \qquad (3.35)$$

The flux density at the interface is $-\varepsilon F_p$ and the effective area is $Y_a.Z$. The volume enclosed in region III is the sheet density multiplied by the area $L_3.Z$.

Hence from (3.35)

$$\varepsilon F_p Y_a Z = q\ (n_{so}-n_{sd}+n_{ss})L_3 Z \qquad (3.36)$$

Dividing both sides by Z and making L_3 the subject of the formula gives

$$L_3 = -\varepsilon F_p Y_a\ [q(n_{so}-n_{sd}-n_{ss})]^{-1}$$

Using equation 3.28b yields

$$L_3 = -Fp\left\{\frac{q}{\varepsilon} N_{eff}\right\}^{-1}$$

which is equation 3.29, and hence the model represents a closed-charge system.

3.3 SUBSTRATE CURRENT IN IHEMT

3.3.1 Qualitative Description

Electrons in the 2DEG are confined, at low applied drain-source voltages, by the conduction-band energy discontinuity which results from the difference in energy bandgap between the GaAs and $Al_xGa_{1-x}A_s$. At larger drain-source voltages however, 2DEG electrons may gain sufficient energy from the longitudinal field to be scattered over the energy barrier, into the AlGaAs layer. This effect is called hot-electron real-space transfer [3.26, 3.27]; the 'real-space' term is used to distinguish the effect from momentum (k) -space transfer which describes intervalley transfer in bulk material.

The electrons scattered into the AlGaAs result in a component of drain current in parallel with the 2DEG current, leading to an increase in device drain current, output conductance and transconductance. Since the effect of this current is similar to that of substrate current in MESFETS, the term 'substrate current' is retained for the AlGaAs component in IHEMTs.

The substrate current described here is distinct from that current in the AlGaAs caused by applying a gate-source voltage greater than V_{GSM} (equation 3.10), which results in the formation of an undepleted region in the AlGaAs. With the exception of this work, no attempts at the incorporation of substrate current in analytic models for either IHEMTs or CHEMTs have appeared in the literature to date. In spite of the approximations used in modelling the substrate current, good agreement between predicted and measured device performance is obtained (section 3.4), without the need to use non-physical values as input parameters to the model.

In IHEMTs, the barrier to hot electron transfer is determined by the conduction-band energy discontinuity at the interface. In MESFETs, the effective barrier is determined by the dopant profile and impurity concentration in the substrate [3.28]. The resulting substrate current has been modelled in a 2D simulator by Barton et al. [3.29]. In CHEMTs, the effective barrier is determined by the impurity concentration in the nominally undoped GaAs layer and will be discussed in section 3.5. Experimentally, reduced drain conductance in both MESFETs and CHEMTs has been obtained by the use of an undoped AlGaAs buffer layer under the active channel [3.30, 3.31], which increases the barrier height in these structures.

In the next subsection, the calculation of the substrate current is presented. In broad outline, the ratio of the density of electrons in the 2DEG to the density of those scattered into the AlGaAs is calculated. From this, the ratio of 2DEG to substrate current is obtained. Since the 2DEG current can be calculated (section 3.2), the actual substrate current can easily be found from this ratio. The model presented here contains several enhancements over the previously published model by the author [3.12].

3.3.2 Calculation of ratio of substrate to 2DEG current

2DEG Electron temperature

The distribution of electrons in energy is assumed to be given by the Fermi-Dirac distribution function,

$$f(E) = [1+\exp((E-E_F)/kTe)]^{-1} \tag{3.37}$$

where E_F is the Fermi energy, k is Boltzmann's constant and Te is the electron temperature. The electron temperature is obtained using the simple energy conservation equation [3.32]:

$$\frac{3}{2} k(Te-To)\, \tau_e^{-1} = qv_s \langle F \rangle \tag{3.38}$$

where T_o is the lattice temperature (assumed 300k), τ_e is the energy relaxation time and $\langle F \rangle$ is the average field. The left hand side of the equation is the rate of change of electron energy due to lattice collisions, and the right hand side is the energy per unit time that the electron gains from the applied electric field. The energy relaxation time is given by [3.33]:

$$\tau_e = \frac{\mu_o}{v^2} \left(\frac{kT_o}{q}\right) \tag{3.39}$$

The electrons become hot mainly in region II where the longitudinal field is high (figure 3.3) and the carriers travel at their saturated drift velocity, v_s. The average field in region II is given by:

$$\langle F \rangle = (V_2-V_1)\, (L_g-L_1)^{-1} \tag{3.40}$$

From the distribution function (3.37), the number of electrons with energy greater than the barrier energy can be calculated once the density of states function and the Fermi energy are established. It is interesting to note that, in spite of the approximations implicit in equations 3.38 and 3.39, the electron temperatures calculated when simulating devices are similar, for comparable structures, to those calculated using more elaborate energy transport equations in a numerical 2D CHEMT simulator [3.34].

Barrier energy

The barrier energy to real-space transfer, B, is equal to the conduction-band edge energy discontinuity, ΔE_c, at the heterojunction. The dependence of this discontinuity on aluminium mole fraction, x, has been shown by the author ([3.35] and Appendix F) to be given by:

$$\Delta E_c \text{ (ev)} = 0.81x \qquad x \leq 0.45 \qquad (3.41a)$$

$$= 0.395 + 0.05x - 0.258x^2 \qquad x > 0.45 \qquad (3.41b)$$

These dependences mean that the maximum barrier height obtainable in the GaAs/AlGaAs system is 0.365 eV, and occurs at a mole fraction of 0.45.

Density of states function

The density of available states function is assumed to be that of a perfectly parabolic conduction band bulk material:

$$g(E) = \frac{4\pi}{h^3} \sqrt{(2m_c)^3} \sqrt{E-Ec} \qquad (3.42)$$

where m_c is the directionally-averaged density of states effective mass for conduction band electrons, h is Planck's constant and E_c is the energy of the conduction band edge in the GaAs. This equation neglects the non-parabolicity of the GaAs conduction band [3.36], but this simplification is reasonable for non-degenerate material (the effective volume densities typical of 2DEG sheet densities in region II are equivalent to those of non-degenerate material).

At first sight, it would appear invalid to use a three dimensional (3D) density of states function when it is firmly established that the electrons occupy discrete energy levels in a quantized system, formed by the approximately-triangular potential well at the heterojunction [3.16, 3.17, 3.20]. However, when electrons become hot due to heating in the electric field, a large fraction of the electrons may have energies greater than 0.36 eV, the maximum barrier height. Electrons with these energies are no longer confined in the potential well and behave essentially as bulk 3D GaAs electrons.

The widely-used method of Lee et al. [3.16], where the 2DEG sheet density is calculated assuming only the lowest two subbands in the well are occupied, has been shown by many workers (for example [3.17]) to be a reasonable approximation, even at 300K, when no electron heating occurs. However, calculations by the author have shown that as the temperature is increased, either by lattice heating or by electric field, occupation of higher subbands becomes important due to 'broadening' of the Fermi-Dirac function. 'Broadening' of the Fermi function refers to the fact that as the temperature increases, the kTe-term increases, and hence the energy interval

over which the value of the function decreases by a given amount, increases. To incorporate the effect of higher subbands, the 2D density of states envelope function, $g_{2D}(E)$, proposed by Pierret [3.37] for an infinite triangular well, was tried in preliminary calculations. It was found that at high electron temperatures, the sheet density of electrons at a given energy (calculated from the product $g_{2D}(E).f(E)$) increased as the energy was increased. This problem occurred because of the $E^{3/2}$ dependence in the $g_{2D}(E)$ function.

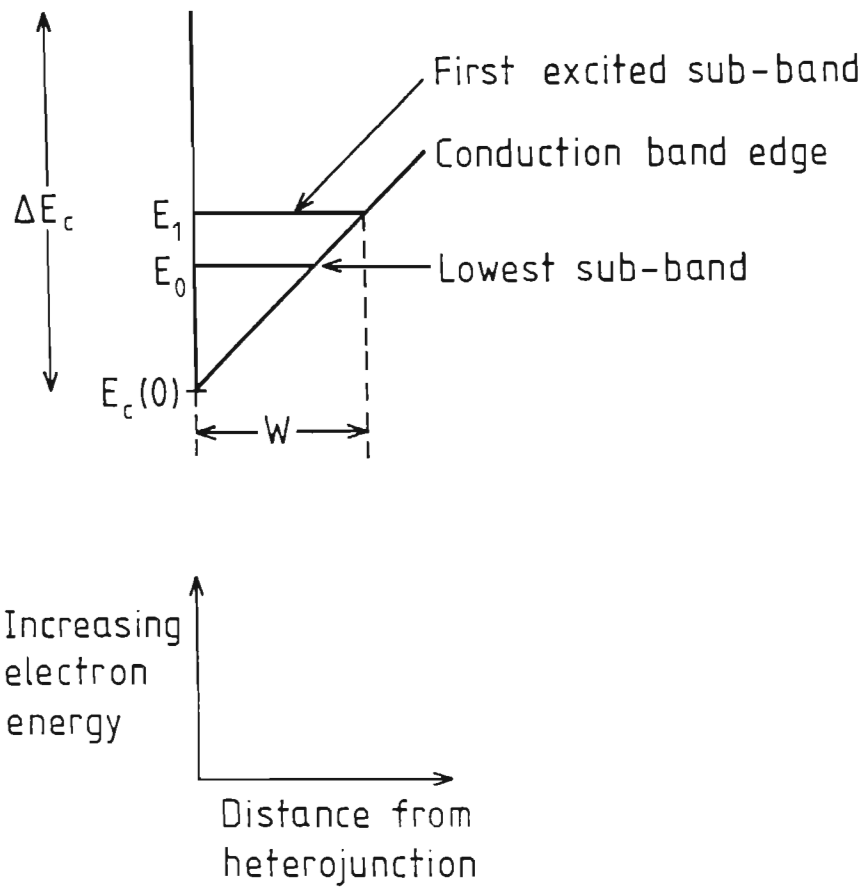
As a result of these considerations, it was decided to use the 3D (bulk) density of states function, equation 3.42. Interestingly, recently reported calculations by Yoshida [3.38] have shown that the use of the 3D density of states and Fermi-Dirac function for calculating the 2DEG sheet concentration at 300K agrees well with 'exact' quantum mechanical calculations. This result is in contrast to that of Lee et al. [3.39] which found poor agreement between the bulk and triangular well calculations.

Equivalent volume density of 2DEG

In order to find the ratio of the number of carriers in the AlGaAs (substrate) to the number in the 2DEG, the position of the Fermi energy must be found using the Fermi function and 3D density of states (equations 3.37 and 3.42). To find the Fermi energy, the equivalent volume density of 2DEG electrons must be obtained from the 2DEG sheet density, N_{sd}. The volume density is given by

$$N = n_{sd}/W \qquad (3.43)$$

where W is the well width. The well width is assumed to be determined by the width at the first excited subband, E_1, (see Figure 3.5) because at 300K more than 80 percent of the electrons are in the first two subbands [3.17]. The slope of the conduction band in the well is assumed to be constant and determined by the field at the interface [3.40]:



$$E_j = \left[\left(\frac{\hbar^2}{2m_c}\right)\left(\frac{3}{2}\pi\frac{q^2}{\varepsilon}\right)^2\left(j+\frac{3}{4}\right)^2 n_S^2\right]^{1/3}$$

$$j = 0, 1, 2 \ldots\ldots\ldots$$

Figure 3.5. Simplified conduction-band diagram in the region of the 2DEG, used to calculate the effective well width. E_j is the energy of the j-th subband. (Equation is from reference 3.37).

$$F = \frac{q}{\varepsilon} n_{sd} \tag{3.44}$$

The energy of the first subband, E_1, is obtained by substituting (3.44) into Pierret's equation 7:

$$E_1 = \left[\left(\frac{\hbar^2}{2m_c} \right) \left(\frac{3}{2} \pi \frac{q^2}{\varepsilon} \right)^2 \left(7/4 \right)^2 n_{sd} \right]^{1/3} \tag{3.45}$$

The well width, W, is then given by $E_1 . F^{-1}$

Scattering

Electrons with energy greater than the barrier energy are assumed to be scattered into the AlGaAs by collisions with phonons. Electrons scattered towards the barrier which have energy less than the barrier energy are assumed to be reflected by the barrier, that is, zero tunnelling is assumed.

Isotropic scattering is assumed, which means that 50 percent of 2DEG electrons are scattered towards the substrate and 50 percent are scattered away. Hence the number of carriers in the substrate is assumed to be 50 percent of those carriers which have energy greater than the barrier height. The remainder of the electrons are therefore the 2DEG electrons.

An implicit assumption here is that there is a sufficient density of available states in the AlGaAs layer for the density of states not to be a limiting factor in the transfer of electrons from the GaAs to the AlGaAs. This assumption is justified because the density of available states is greater in AlGaAs than GaAs, in part because of the increased contribution of upper valleys due to lower intervalley energy differences in the conduction band (Figure 3.6).

The density of carriers in the substrate is then given by:

$$N_{subs} = \tfrac{1}{2} \int_B^\infty g(E) \, f(E) \, dE \tag{3.46}$$

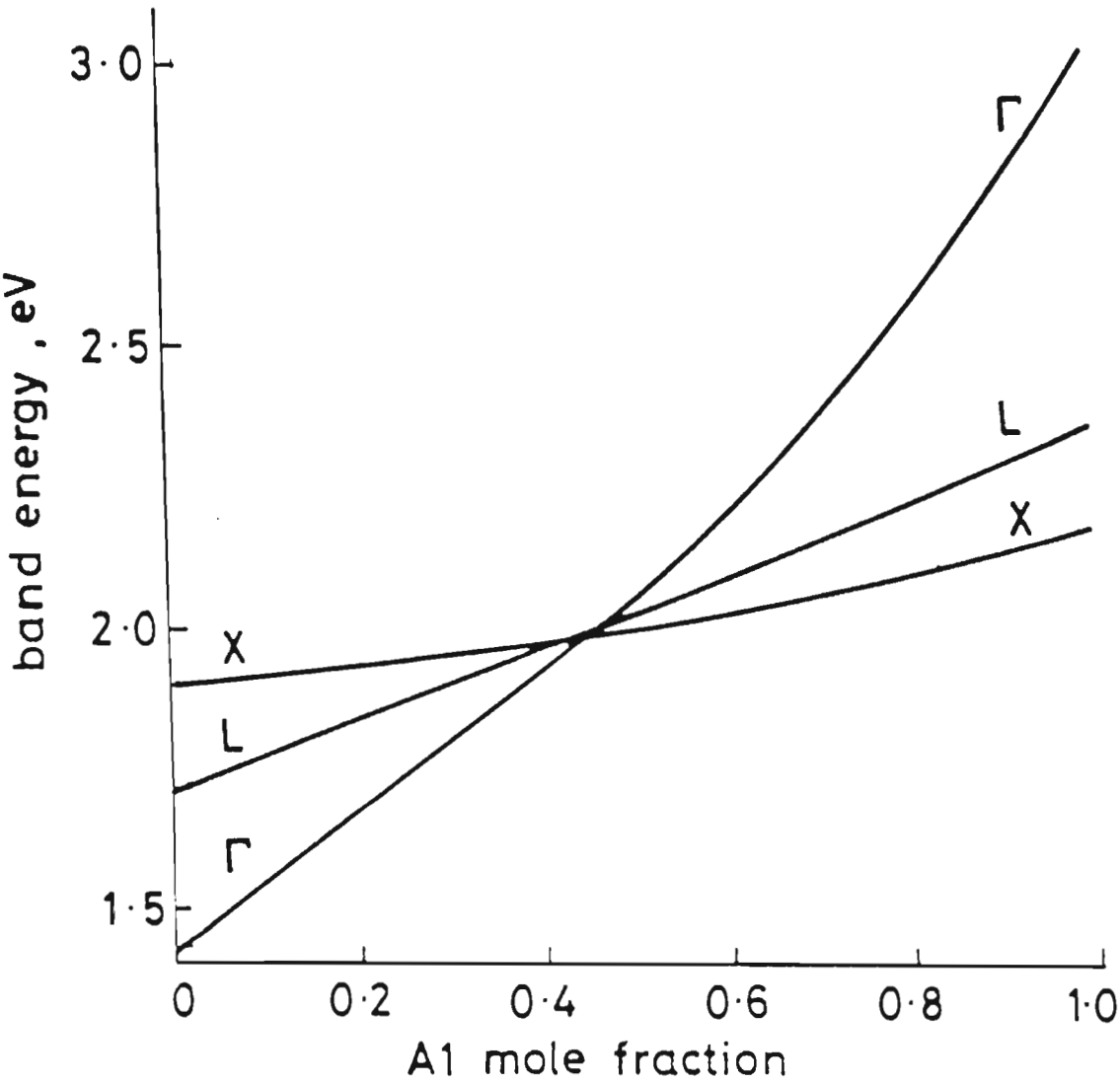


Figure 3.6. Energies of the conduction-band valley minima in AlGaAs, as a function of aluminium mole fraction. Energies are relative to the top of the valence band at the gamma point (K=0,0,0).

where B is the barrier energy. The total density of electrons, N_{tot}, is given by

$$N_{tot} = \int_{E_c}^{\infty} g(E)\, f(E)\, dE \qquad (3.47)$$

where E_c is the energy at the conduction band edge. The ratio N_{subs}/N is then given by $N_{subs}/(N_{tot}-N_{subs})$

Evaluating the integrals

The Fermi energy is not known directly for the evaluation of these integrals. Consequently, to find the ratio N_{subs}/N, the integrals are evaluated for different values of Fermi energy until the energy is found at which the calculated 2DEG volume density is equal to the 'known' value from equation 3.43. This iterative process is discussed further in Appendix B.

The integrals in (3.46) and (3.47) cannot be evaluated directly. Their solution, using approximations to the Fermi-Dirac integral 3.41] and numerical integration are detailed in Appendix B.

Current ratio

The ratio, R, of the substrate current, I_{subs}, to 2DEG current, I_{2D}, is assumed to be given by the ratio of the carrier concentrations multiplied by the ratio of the carrier saturation velocities:

$$R = (N_{subs}.\ va)\ (N.v_s)^{-1} \qquad (3.48)$$

where v_a is the saturated drift velocity of carriers in the AlGaAs (substrate) layer. The large longitudinal fields present in the AlGaAs layer will accelerate the carriers to their saturated velocity, which makes the use of the velocity ratio a reasonable assumption. A more rigorous

treatment would be to characterise the electron transport in AlGaAs by some form of velocity-field characteristic. However, the lack of reliable data on transport parameters in AlGaAs means that this approach offers no greater potential accuracy than (3.48) at present.

Widely differing data on the saturated drift velocity of electrons in AlGaAs have appeared in the literature, and the dependence of this velocity on aluminium mole fraction is also not well established. Sugeta et al [3.42] reported values of 4.3×10^6 and 1.5×10^6 cm/s for mole fractions of 0.34 and 0.43 respectively while Hill and Robson [3.43] reported a value of 4.8×10^6 cm/s for a mole fraction of 0.45. Hirano et al. [3.44] inferred values of 7×10^6 and 3×10^6 cm/s for mole fractions of 0.24 and 0.3 respectively, from a simple device model.

AlGaAs has a direct bandgap for mole fractions less than about 0.45 and has an indirect bandgap for mole fractions between 0.45 and 1. As the mole fraction is increased towards crossover, the intervalley energy differences decrease [3.45] (Figure 3.5) and hence intervalley scattering rates will increase, which will in turn tend to decrease the saturated drift velocity. Beyond crossover the L and X-valleys, which have low mobilities and high effective masses, are lower in energy than the high-mobility Γ valley. Hence, electron transport will be dominated by the high-mass valleys and the velocity is expected to remain low.

In the absence of better knowledge, the AlGaAs velocity was assumed to be given by:

$$v_a = 0.3\times10^7 \text{ cm/s for } x > 0.4 \qquad (3.49a)$$

$$= (1.1-2x)\times10^7 \text{ cm/s for } x \leq 0.4 \qquad (3.49b)$$

3.3.3 Inclusion of substrate current in 3-region model

The total IHEMT drain current, I_{DS}, is increased from the value of current, (I), calculated in section 3.2 for the 2DEG, by the addition of the substrate current component. The total drain current is given by

$$I_{DS} = I\,(1+R) \qquad (3.50)$$

where R is the ratio of substrate to 2DEG current (equation 3.48). The value of R is determined, in part, by the variation in longitudinal field (and hence potential) in region II (equation 3.40). However these fields were calculated assuming only current I flows through region I and the source and drain resistance, while physically it is clear that the substrate current must also flow through these areas. The 'extra' current flowing through these regions has a negative feedback effect; by increasing the potential difference across the resistances, the field $\langle F \rangle$ (used to calculate to the substrate current initially) will decrease (equation 3.40). This in turn would reduce the ratio, R, as calculated from the new average field $\langle F \rangle$.

Therefore the current ratio, R, and hence drain current, must be calculated by iteration. The following procedure is used:

(a) assuming a value for R, a solution for V_1, V_2 V_3 and I_{DS} is found so that the current in region I is equal to the sum of the region II and substrate current.

(b) from this solution, a new value R' for the ratio is obtained.

(c) if the difference between R and R' is greater than a certain percentage, the procedure is repeated from (a) again.

Hence the final ratio, R, is consistent with the potentials in region II (equation 3.40). However, one element of self consistency in the model has been lost; the current in region one, (1+R)I, is no longer equal to the region II 2DEG current alone (I), but is now equal to the sum of the 2DEG and substrate current. In actual devices, it is expected that the length of region I will decrease to allow the extra current required in region I (equation 3.17). Hence the loss of a true self-consistent solution by the addition of substrate current to the model is not considered serious.

3.4 IHEMT SIMULATION RESULTS

3.4.1 Introduction

In this section, the results of simulations using the IHEMT model are presented. Comparisons between simulated and measured IHEMTs are given to demonstrate the usefulness of this model. In addition, some of the predicted dependences of electrical characteristics on material and structural parameters are shown. The choice of 2DEG electron saturation velocity for HEMT simulations is discussed, and it is concluded that the inclusion of substrate current in this model obviates the need to overestimate the velocity to achieve good agreement between simulations and experiment.

A menu-driven interactive simulation program incorporating the model was written in FORTRAN on a VAX 11/875. The short execution time of the program (approximately 10 CPU seconds to calculate the drain current at 40 bias points, used to produce figure 3.7) means that this software could readily be written to run on a small personal computer or incorporated in a circuit simulation program.

3.4.2 I_{DS}-V_{DS} Characteristics

A comparison between predicted and measured I_{DS}-V_{DS} characteristics for an IHEMT, fabricated at GEC, is shown in Figure 3.7. The measured currents were scaled to 1 mm gate width to facilitate comparison. The material and structural parameters used in the simulation are shown in Table 3.1.

The agreement is good, in spite of the fact that the only parameter adjusted to obtain this fit was the gate recess depth. The exact gate recess depth was not known because the etch rate of the wet chemical etch used for the recess etch had not been accurately determined.

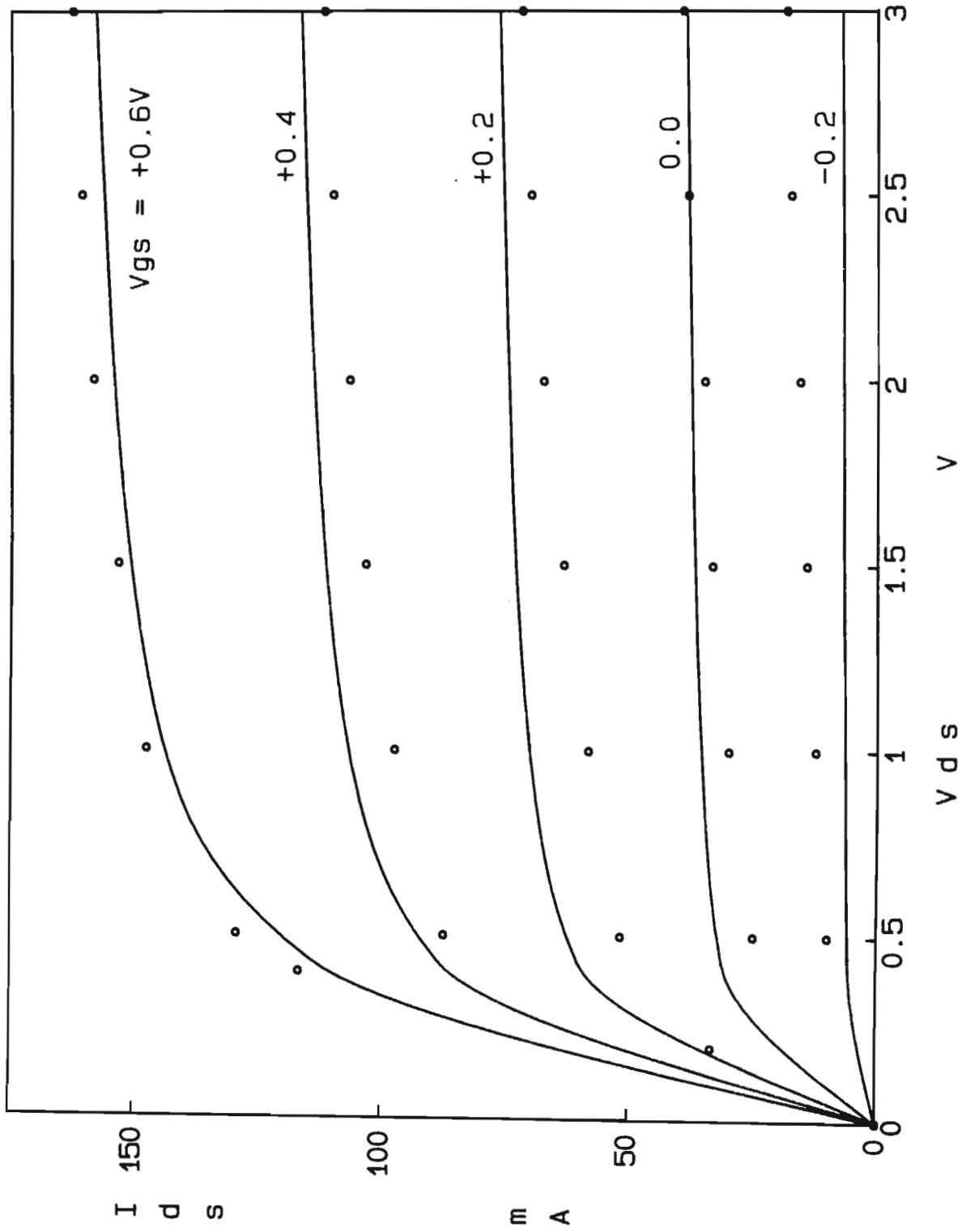


Figure 3.7. Comparison between simulated and measured DC characteristics for a 0.5 micron gate length IHEMT. Data points are from device 5453, shown in Figure 2.4.

The first five parameters in Table 3.1 are the nominal values for the IHEMT wafer, as specified to the MBE grower. The nominal cap layer thickness was 100 nm, but it was estimated that approximately 10 nm was removed during device fabrication. The nominal gate length of 0.5 micron was used, together with estimates for the gate-drain and gate-source spacings. The 2DEG sheet concentration and low-field mobility were obtained from measurements on Hall samples from the same wafer as the measured device. The contact resistance was measured on a transmission line structure while the Schottky-barrier height was taken from Okamoto et al. [3.46]. The occupied surface state density was taken as 3.5×10^{12} cm^{-2} [3.47] and the recess depth used was 67 nm. The saturated drift velocity was assumed to be 10^7 cm/s; this choice will be discussed further in the next subsection.

An even better fit could be obtained by allowing some of the parameters to be changed to values other than their measured or nominal values, but there is no justification for this procedure. The discrepancy between measured and calculated values in Figure 3.7 is therefore a measure not only of the model accuracy, but also the precision in MBE growth, processing tolerances and measurement accuracy.

Figure 3.8 shows the effect (simulated) of overetching the gate recess by 10 nm. The data points are the same as in Figure 3.7. Qualitatively, the large shift in predicted current is due to an increase (more positive) in threshold voltage. In Figure 3.9, the effect (again, simulated) of underetching the recess by 10 nm, is shown. In this case, the threshold votlage has decreased, leading to greater drain current for a given gate bias. The simulated transconductance has also decreased, which is consistent with the expectation that transconductance will decrease as the gate to 2DEG separation is increased.

1.	Undoped GaAs thickness	50	nm
2.	n-AlGaAs doping density	10^{18}	cm^{-3}
3.	Cap layer doping density	10^{18}	cm^{-3}
4.	Al mole fraction	0.28	
5.	Spacer thickness	6	nm
6.	Cap layer thickness	90	nm
7.	Gate length	0.5	μm
8.	Gate-drain spacing	0.6	μm
9.	Gate-source spacing	0.5	μm
10.	2DEG sheet concentration	0.9×10^{12}	cm^{-2}
11.	Hall mobility	4300	cm^2/Vs
12.	Contact resistance	0.5	ohm.mm
13.	Schottky barrier height	0.9	eV
14.	Surface state density	3.5×10^{12}	cm^{-2}
15.	Recess depth	67	nm
16.	Saturated drift velocity	10^7	cm/s

Table 3.1 Structural and material parameters used in the simulation shown in Figure 3.7

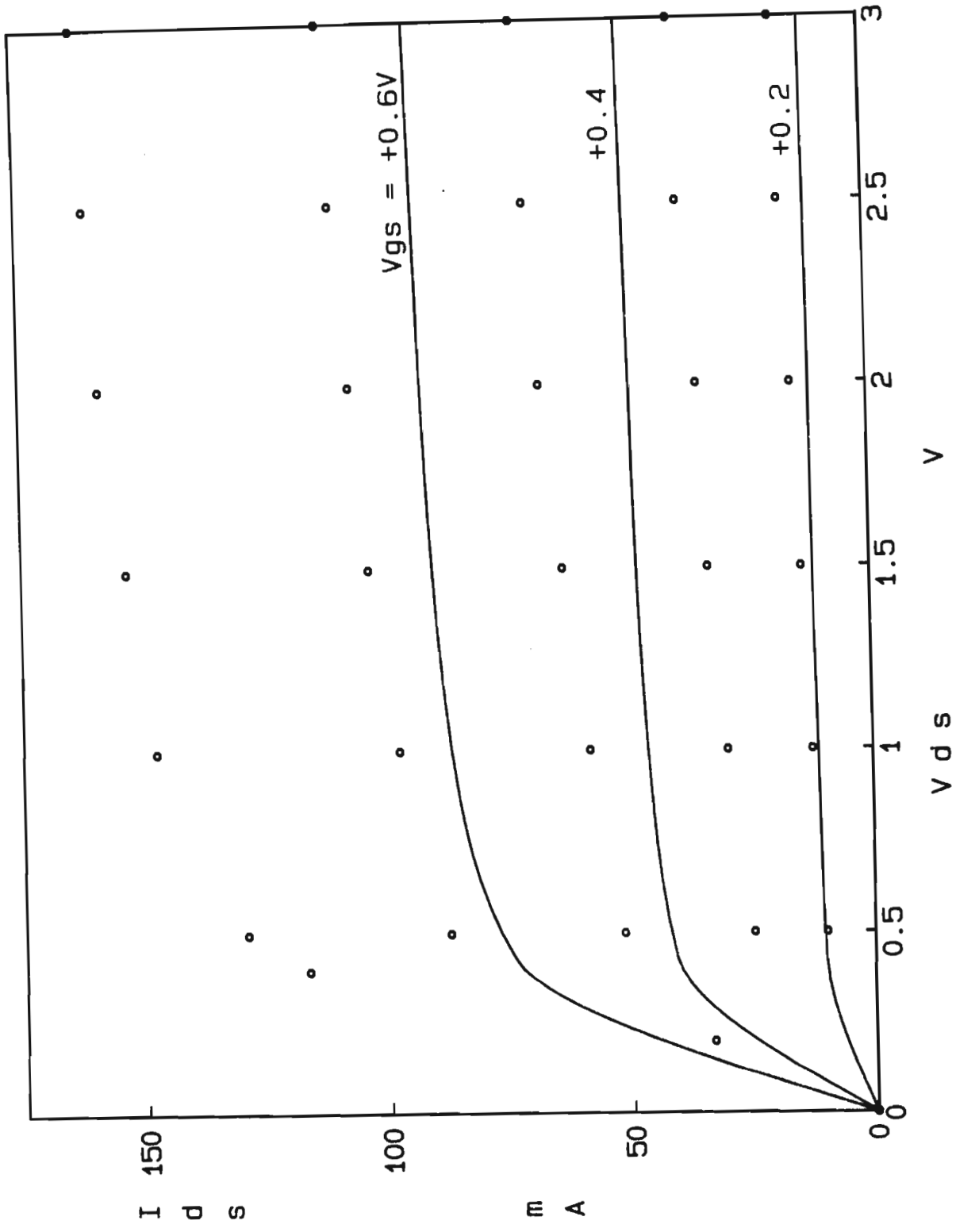


Figure 3.8. Simulated I-V characteristics for device with the gate recess overetched by 10 nm. Data pionts are the same as in Figure 3.7.

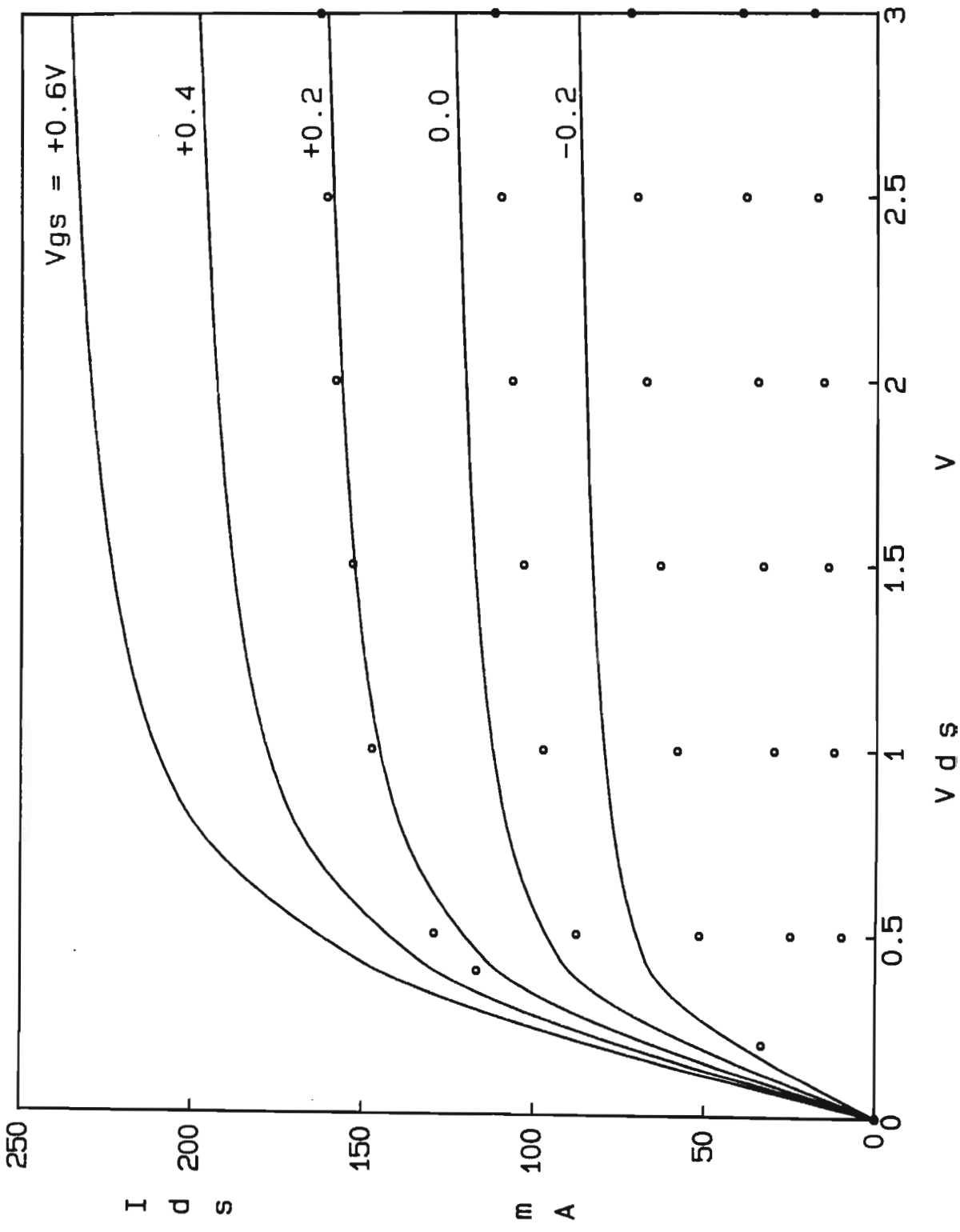


Figure 3.9. Simulated DC characteristics for IHEMT with gate recess underetched by 10 nm. Data points are from Figure 3.7.

These figures illustrate the critical nature of the gate recess etching step in device processing. To fabricate devices reproducibly, the etch process must be well controlled to avoid variations across the wafer and from wafer to wafer. One approach to this problem is to exploit the excellent uniformity achievable in MBE layers [3.48] by either growing an AlGaAs etch-stop layer in the GaAs or by growing the layers only to the thickness required under the gate. The first method requires the development of a selective etch system [3.49] and introduces an unwanted extra heterojunction which will increase the parasitic resistances of the device [3.50]. The second method means that the cap layer thickness in the gate-source and gate-drain regions cannot be chosen independently of the considerations (primarily transconductance and threshold voltage) governing the choice of cap layer thickness under the gate. Increasing the cap layer thickness outside the gate region will decrease the access resistances (equations 3.33 and 3.34) but also decrease the breakdown voltage. This latter trend occurs because as the cap layer thickness increases, the peak electric fields in the channel increase because L_3 decreases, and hence avalanche breakdown will occur for a lower drain-source voltage.

The agreement between the measured and simulated slope of the I_{DS}-V_{DS} characteristic in saturation (DC output resistance) in figure 3.7 is good. The finite simulated output conductance is caused by two effects; reduction in the length of region I (often termed 'channel length modulation') and an increase in substrate current as the drain bias is increased. Output conductance is probably the most difficult DC parameter to predict in any model and many models either predict infinite output conductance [3.13, 3.40] or use a fitting parameter to get agreement with experiment [3.6, 3.21, 3.54]. One limitation to the accuracy which can be obtained in this model is the lack of data on the dependence of surface charge on bias and structure (see Section 3.4.4).

Further simulated dependences of device electrical characteristics on material and structural characteristics on material and structural parameters will be presented in Section 3.4.4.

3.4.3 Saturated drift velocity in 2DEG

The saturated drift velocity of bulk GaAs is approximately 10^7 cm/s and occurs for electric fields (in the direction of transport) greater than about 15 kV/cm [3.51]. The longitudinal fields present in HEMTs (or FETs) are of the order of 300 kV/cm and heat the 2DEG electrons to high energies which means that the electrons are no longer confined in the 2D well. These electrons then behave as bulk (3D) carriers and consequently will have the same saturated drift velocity as in bulk GaAs [3.52], determined primarily by spontaneous phonon emission from the electrons.

However, reported values for saturated drift velocity in CHEMTs have ranged between 1.3 and 2.5×10^7 cm/s [3.13, 3.44, 3.53-55]. These values have been obtained by choosing the value for v_s which best fits the measured transconductance and current assuming a simple model of the form:

$$g_{mo} = \partial I_{DS} / \partial V_{GSi} = \varepsilon v_s Z / d \qquad (3.51a)$$

$$V_{GSi} = V_{GS} - I_{DS} R_s \qquad (3.51c)$$

where g_m and g_{mo} are the measured (extrinsic) and intrinsic transconductances respectively, Z is the gate width, d is the gate-2DEG spacing, R_s is the source resistance and V_{GSi} is the effective gate-source bias across the intrinsic device.

The reason given for the large difference between these deduced values and the expected value of 10^7 cm/s is that the saturated drift velocity represents the average transit velocity under the gate, rather than the saturated drift velocity itself. This implies that the carriers must experience considerable velocity overshoot in the channel. However, the direct measurements of velocity-field characteristics by Masselink et al. [3.9] (Figure 3.2) showed that the high Hall mobility, characteristic of HEMTs, did not yield enhanced drift velocities. Monte Carlo simulations show varying amounts of velocity overshoot, depending on the assumptions and device structure used; two recent results [3.56, 3.57] found negligible overshoot in devices with a gate length of 1 micron. Qualitatively, this result is expected because the length of the charge control region (including the length of region III) is much greater than the mean free path between scattering events.

In this work, good agreement between measured and simulated DC characteristics is obtained, without the need to assume velocities greater than 10^7 cm/s, because of the inclusion of the substrate current component in the model. Figure 3.10 shows the measured and simulated IHEMT characteristics for a gate-source bias of +0.4V, taken directly from Figure 3.7. The lower curve is the 2DEG current component and the upper curve is the total drain current. At a drain bias of 2.5V, the predicted total drain current and predicted 2DEG current component are approximately 5 percent larger and 25 percent smaller than the measured value, respectively.

The lower curve is obtained by simulating the device with an infinitely high barrier at the heterojunction, which sets the substrate current to zero. This condition is implicit in analytic models which neglect substrate current. To obtain a better fit, while maintaining this condition, it is

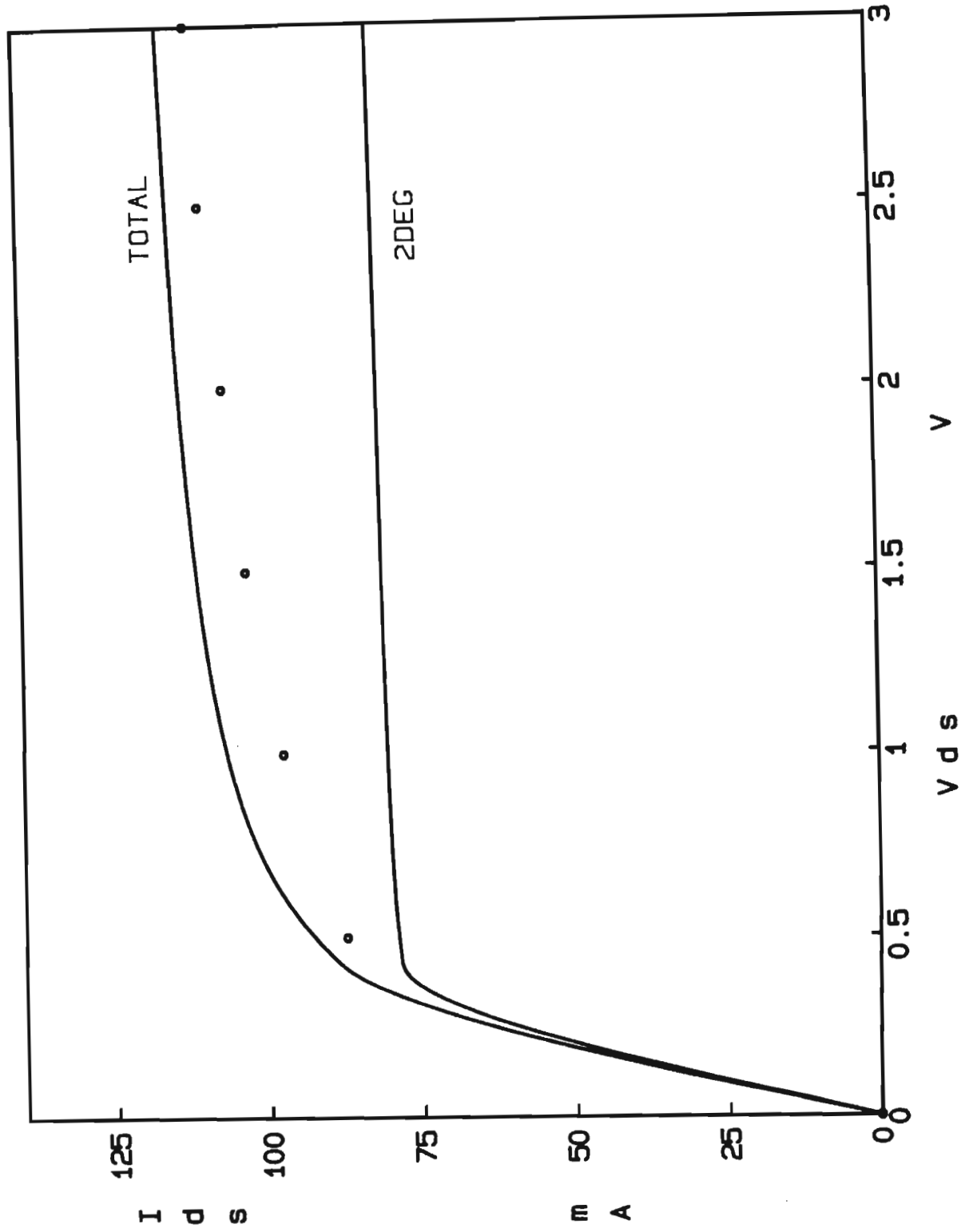


Figure 3.10. Graph showing difference between the 2DEG current and total drain current in IHEMT simulation. V_{GS} is 0.4V and the data points are for the same device (5453) as the previous three Figures.

necessary to either decrease the threshold voltage or increase the saturated drift velocity. This point is clearly illustrated by the simple approximate expression (which form the basis of equation 3.51a):

$$I_{DS} = qn_s.v_s.Z \tag{3.52a}$$

$$n_s = \varepsilon(V_{GSi}-V_T)/qd \tag{3.52b}$$

The first option is unsatisfactory because the threshold voltage can be directly measured. Therefore, to obtain a better fit, while maintaining the known value of threshold voltage, it is necessary to increase the saturated drift velocity. However, simply increasing the velocity results in a knee voltage which is too low and a knee which is too abrupt, (the knee voltage is the drain voltage where the transition from linear to saturated operation occurs; in the lower curve of figure 3.10, this value is about 0.4V). To increase the knee voltage, the parasitic resistances must be increased, which requires a further increase in saturated drift velocity to maintain the required current value.

The procedure of increasing source resitance and saturated drift velocity has been used directly in two models [3.54, 3.58] to obtain good agreement with the experimental CHEMT data of Drummond et al [3.11]. In section 3.5 a fit to the same data, using a modified form of the IHEMT model, is presented. The inclusion of the substrate current has obviated the need to increase either velocity or resistances to provide good agreement with the experimental data.

It has been shown that to obtain good agreement between experiment and simulation, while maintaining physically-based parameter values for the simulation, it is necessary to include the substrate current component.

The origin of saturation velocities greater than 10^7 cm/s in other models can be illustrated as follows: setting $R_S = 0$, for simplicity, and using equation 3.50 and 3.52, gives:

$$I_{DS} = qn_s v_s Z(1+R) \tag{3.53}$$

Hence the transconductance, using equation 3.52b is given by:

$$g_{mo} = \varepsilon v_s Z(1+R)/d \tag{3.54}$$

For a given transconductance, the omission of substrate current leads directly to the need for an overestimation of saturated drift velocity.

3.4.4 Some predicted DC dependences

In this subsection, the simulated dependences of some IHEMT DC electrical characteristics on the device material and structural parameters are presented. Only dependences which are not predicted by any other models (CHEMT or IHEMT) are discussed. It is clear that the model correctly predicts the more obvious dependences, such as the increase in drain current caused by increasing mobility or decreasing contact resistance.

Dependence on aluminium mole fraction

The simulated total drain-source current for two values of aluminium mole fraction, x, in the n-$Al_xGa_{1-x}As$ layer is shown in Figure 3.11. The gate-source voltage is +0.5V and the remainder of the device parameters are the same as used in Figure 3.7 (Table 3.1).

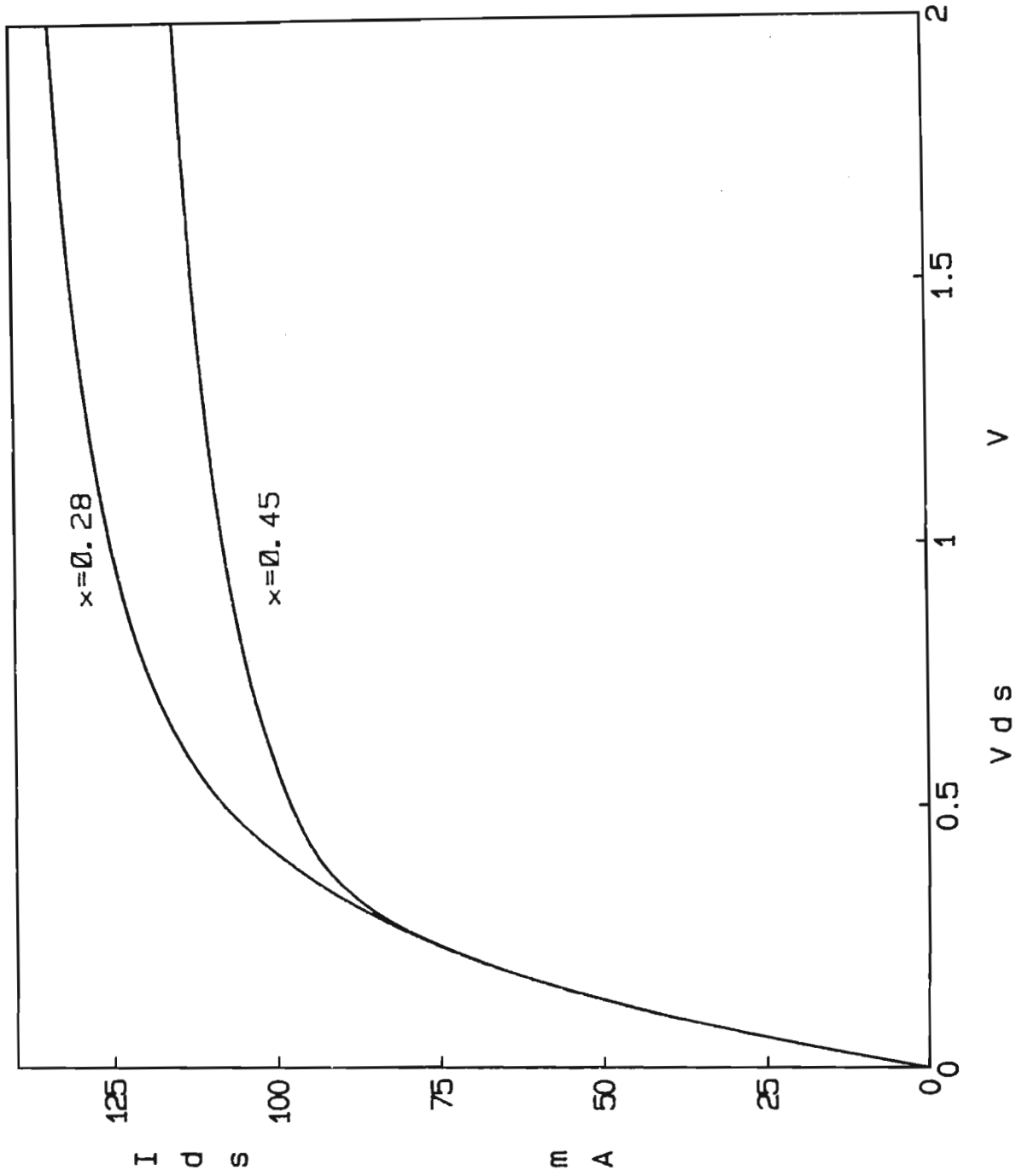


Figure 3.11. Simulated drain current dependence on aluminium mole fraction for a fixed value of gate bias.

The decrease in drain current, caused by the increase in aluminium mole fraction from 0.28 (upper curve) to 0.45 (lower curve), occurs because of the increase in barrier height to hot electron transfer from the 2DEG to the AlGaAs layer (equation 3.41) and the resultant decrease in substrate current. For a given 2DEG sheet carrier density, the IHEMT threshold voltage (unlike the CHEMT) is independent of aluminium mole fraction (equation 3.8) and hence the only source of a change in drain current is the change in the substrate current component. Consequently, the IHEMT charge control model of Lee et al. [3.19] would find no change in device current for a change in aluminium mole fraction only. The minimum ratio of substrate to 2DEG current will occur for a mole fraction of 0.45 (equation 3.41).

Transconductance compression

In GaAs MESFETs, as the gate-source voltage is decreased towards pinch-off, the device transconductance decreases. This effect is called transconductance compression. The exact dependence of transconductance on gate bias is a function, in part, of the doping profile in the channel. In general terms the transconductance decreases because the depletion layer width increases, so increasing the gate to channel spacing.

In CHEMTs, the effective gate to 2DEG spacing also increases as the gate bias decreases towards pinchoff because the triangular well widens as the 2DEG sheet concentration decreases. Consequently, CHEMTs also exhibit transconductance compression.

The effective gate to 2DEG spacing in IHEMTs remains almost constant because the well width narrows as the gate bias approaches pinch-off, but the maximum gate-2DEG spacing is defined by the position of the heterojunction. As a result, IHEMTs exhibit less transconductance compression than either FETS or CHEMTs (see,for example, the measured characteristics of Chapter 2).

In this work, the gate to 2DEG spacing is assumed to be constant (equation 3.6) and hence the model predicts less compression than is measured in practical devices (Figure 3.7). Some compression is predicted however because the ratio of substrate to 2DEG current, R, is a function of gate bias, not a constant as implicitly assumed in the deviation of the simple transconductance expression, equation 3.54.

Effect of occupied surface states

Two main effects result from an increase in the density of occupied surface states (more negative charge) in the source-gate and gate-drain access regions. The possible origins of such a change will be discussed after the effects have been considered.

The first effect is an increase in the source and drain access resistances due to the increased depletion of the cap layer in these regions (equations 3.3 and 3.34). This change, which has also been recently included in a 2D MESFET model [3.59], clearly leads to a decrease in drain current.

The second main effect from an increase in negative surface charge is an increase in the length of region III, as a result of the decreased net positive sheet charge in this region (equation 3.28). This in turn leads to a decrease in the average field in region II and hence a decrease in the substrate to 2DEG current ratio R. Consequently the total drain current also decreases from this effect. (A secondary result of the increase in the length of region III is a decrease in the drain resistance, but this trend is less significant than the other effects described here).

Figure 3.12, (a repeat of Figure 3.3) shows the variation of longitudinal electric field and potential along the 2DEG, for two values of surface state density, in a 0.5 micron gate length IHEMT with a drain bias of 10V. The device is biased near pinchoff to reduce the voltage drops across the source and drain resistances, in order to emphasize changes in the intrinsic device regions. The decrease in peak longitudinal field and increase in length of region III from the increase in negative suface charge are clearly evident.

Figure 3.13 shows the normal electric field distribution along the gate electrode. The expression for this field in region II is obtained, using equations 3.5 and those in Appendix B1 as:

$$F_y(x) = \frac{q}{\varepsilon}(n_{so}-n_{sd}) + F_s\,(Ya/Y)\,\sinh((x-L_1)/Y) \qquad (3.55)$$

The large increase in field at the drain edge of the gate has a similar form to that predicted in 2D MESFET simulations [3.47]. Again, it can be seen that the negative surface charge reduces the peak field at the drain edge of the gate.

Figure 3.14 shows the simulated increase in drain current for the IHEMT from Figure 3.7 (Table 3.1), at a gate-source voltage of +0.4V, as the occupied surface state density is reduced to zero. In this particular structure, the current increases by only around 3 percent. The simulated change in drain current is greater in devices where the change in surface charge is a larger proportion of the cap layer sheet density (equation 3.32).

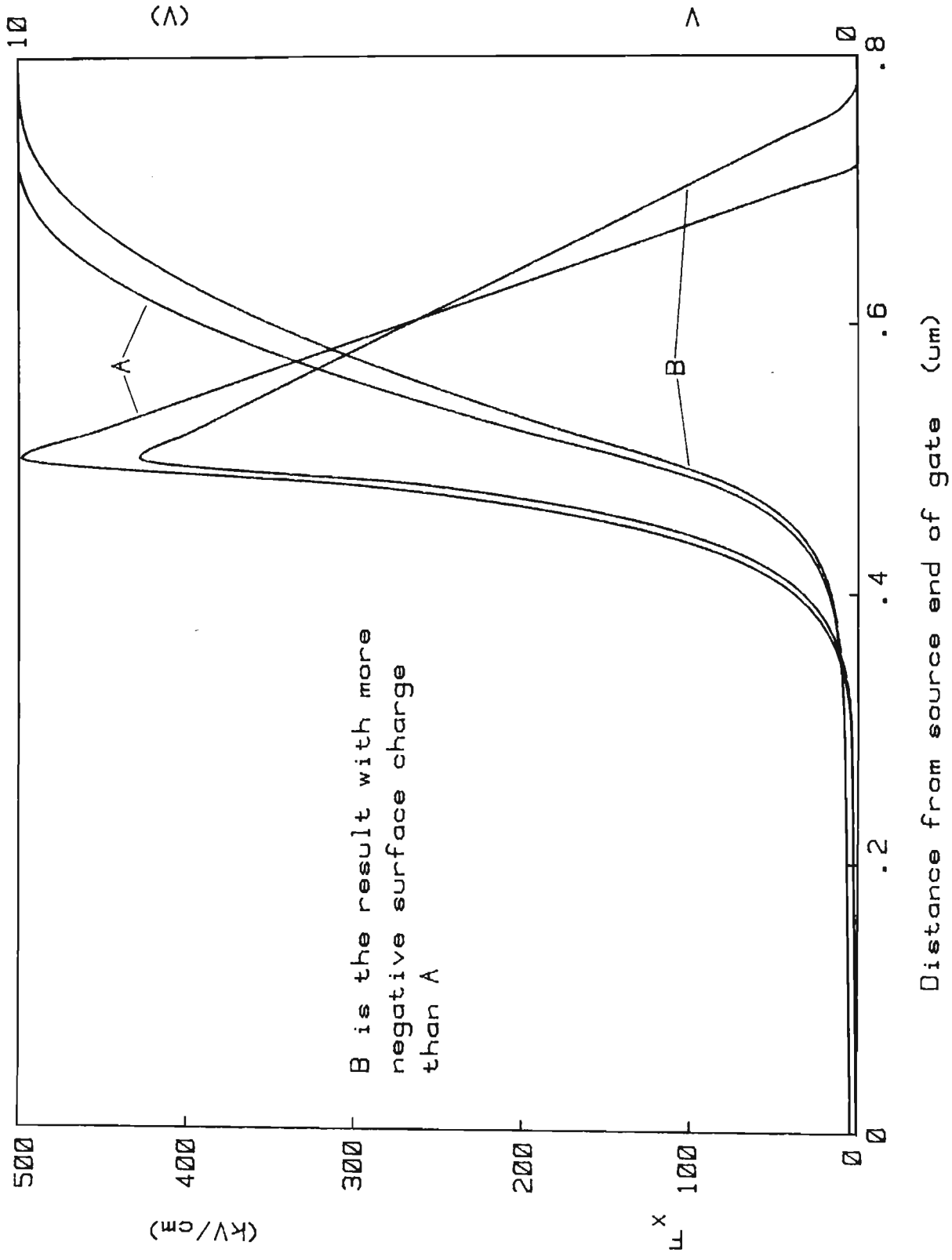


Figure 3.12. Longitudinal electric field and potential distribution in the three intrinsic regions of a 0.5 micron gate length IHEMT with V_{DS} = 10 V.

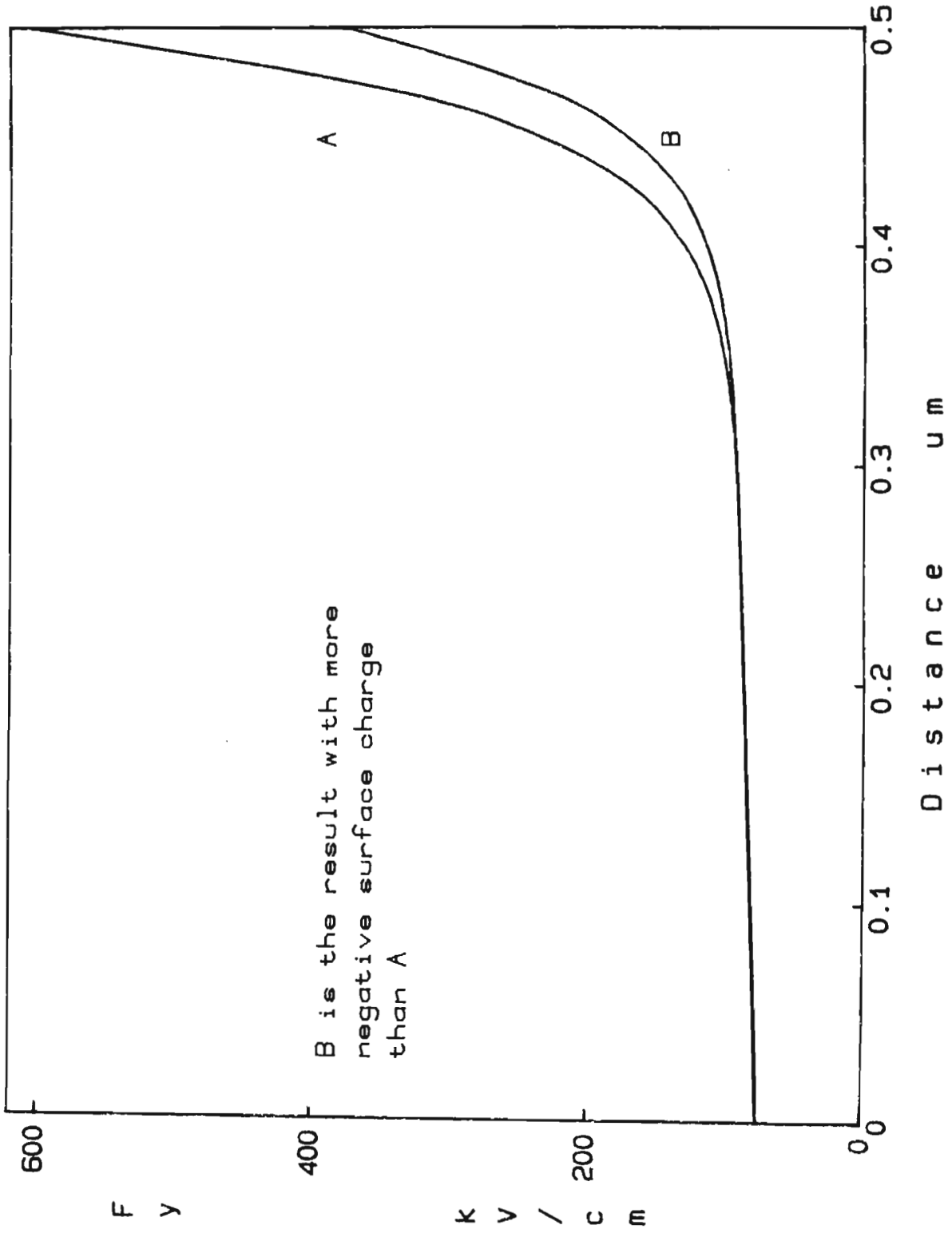


Figure 3.13. Normal electric field distribution along the gate electrode in IHEMT, for two values of surface charge.

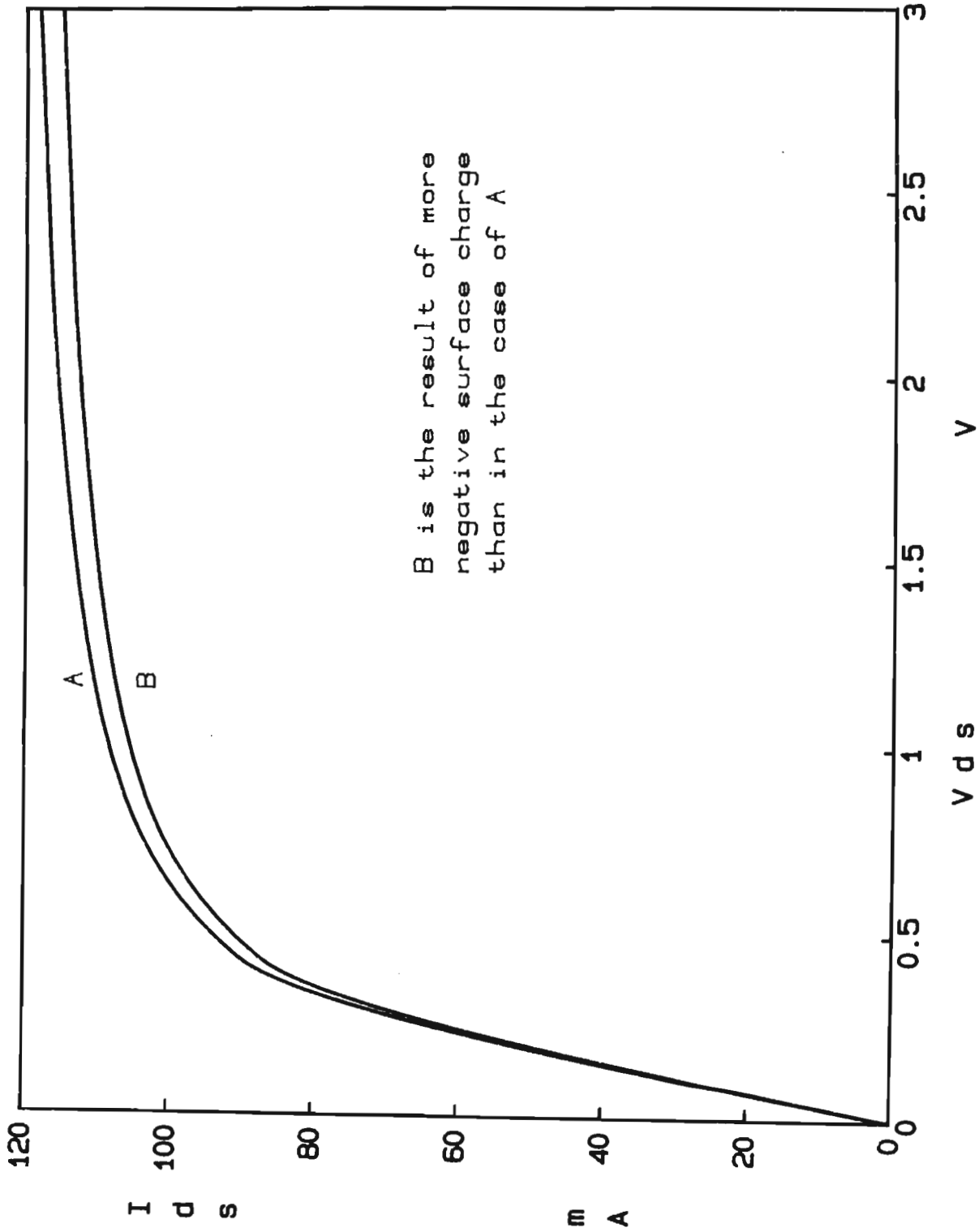


Figure 3.14. Simulated drain-source current for two values of surface charge, at constant gate bias.

Surface charge arises from the trapping, in acceptor-like surface states, of free electrons from the doped GaAs layer. The surface states are allowed energy states near the centre of the energy bandgap, caused by the change in crystal bandstructure at the surface due to the crystal discontinuity. The Fermi-level in the GaAs is pinned at approximately mid-gap and the resultant band-bending forms a depletion region adjacent to the surface [3.60]. For a cap layer doping density of 10^{18} cm^{-3}, the surface potential of approximately 0.7 eV (one half of the GaAs bandgap at 300K) will deplete around 32 nm of the cap, which is equivalent to a sheet surface charge density of 3.2×10^{12} cm^{-2}. This surface charge density is not an upper limit however, because electrons may tunnel from the high field region at the drain end of the gate electrode into unoccupied surface states, and so yield excess surface charge [3.47]. The density of available states on the GaAs surface may be as high as 10^{14} cm^{-2} [3.60].

Hence, the surface state density itself may be a function of bias, and not a constant as assumed in this (and other [3.59]) model. In addition, if excess charge exists near the drain edge of the gate [3.47], then the assumption of a uniform charge distribution is also incorrect (equation 3.27). A further complication is that there may be a difference between the static and dynamic occupancy of surface states because of their relatively slow capture and emission rates.

The time dependence of surface state occupation has been observed by making pulsed I_{DS}-V_{DS} characteristic measurements on IHEMTs made at GEC. Figure 3.15 shows two sets of I-V characteristics, measured by applying gate and drain bias as pulses. The increase in drain current as the pulse width is decreased from 10 to 0.2 μs is due to the greater excess surface charge present when the longer pulse is applied, because the surface states are

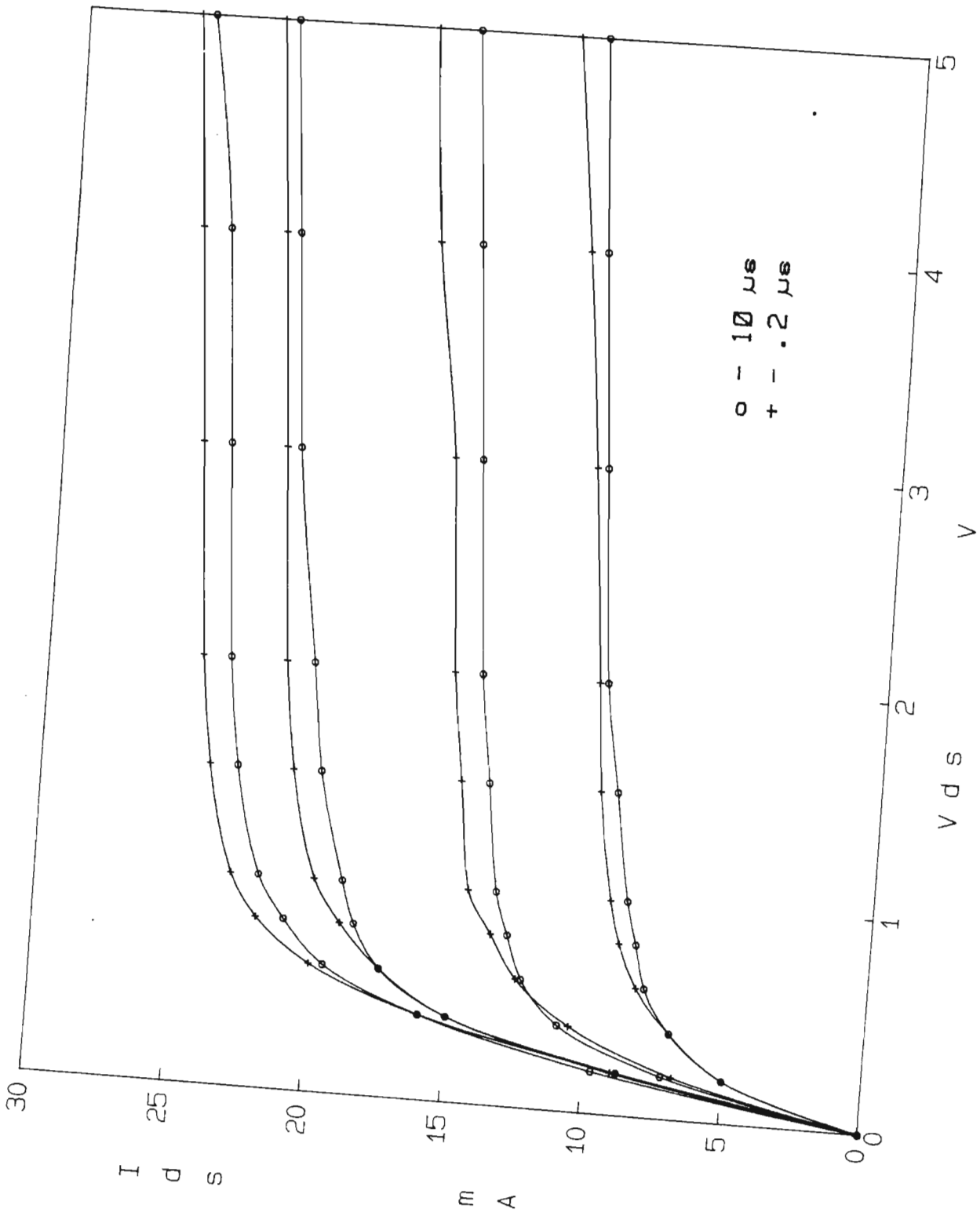


Figure 3.15. Measured I-V characteristics for 0.5 micron gate length IHEMT, using biasing pulse widths of 10 and 0.2 μs.

able to capture tunnelling electrons within this time. (The undulation in the measured curves do not represent a real effect but are the result of measurement errors from estimating the current in the presence of some ringing on the pulse seen on a curve tracer).

In MESFETs similar changes between measured currents at different pulse widths are observed. However in these devices the cause can be either changes in surface charge or barrier lowering at the active channel -substrate interface caused by charging of deep states [3.29]. In IHEMTS, the barrier to the 'substrate' is determined by the conduction-band discontinuity and hence barrier lowering will play no part.

By assuming a constant surface charge density and uniform distribution, the time and bias dependence of surface charge has not been included in this model. The model does however provide a basis for explanation of the effects of surface charge changes on device performance, which are in qualitative agreement with experiment. The effects cannot be addressed in models which ignore the depletion extension beyond the drain edge of the gate.

The surface state density in practical MESFETs is a function of frequency, applied bias, fabrication processes, passivation layer properties (if one is used) and orientation of the channel with respect to the crystallographic axes [3.61-3.63]. No comprehensive quantitative analysis of these effects has yet appeared. The situation is the same for HEMTs. Passivation of devices may lead to change (either increase or decrease) in the occupied surface state density and hence alter the breakdown voltage of the device. [3.47]. A qualitative understanding of this may be gained from Figures 3.12

and 3.13 which show a significant decrease in the peak electric fields in the device from an increase in negative surface charge. This leads to the expectation of a lower value for the ionization integral [3.25] and hence an increased breakdown voltage.

3.5 CHEMT ANALYSIS USING PLANAR IHEMT MODEL

3.5.1 Introduction

In this section, the IHEMT model developed in sections 3.2 and 3.3 is extended to analyse the CHEMT. The major structural difference between the two devices is that in the CHEMT, the doped AlGaAs layer is between the GaAs layer containing the 2DEG and the gate electrode whereas in the IHEMT the AlGaAs layer is under the 2DEG layer. The lateral structure of the two devices is the same. Like the IHEMT, the CHEMT is a planar device with the lateral dimension of the active region typically an order of magnitude greater than the vertical dimension.

The principles of operation of the two devices are also the same; the dependence of 2DEG sheet charge density on gate-source voltage allows the gate to effectively modulate the current between drain and source. Electron transport in the two devices is identical and the measured I_{DS}-V_{DS} characteristics are very similar.

With these similarities between the two devices, it is reasonable to expect that, with suitable modification, the IHEMT model analysis should apply to CHEMT as well.

3.5.2 CHEMT Intrinsic regions

In the IHEMT analysis, the n-AlGaAs layer and 2DEG were treated as a charge sheet displaced from the gate electrode. In extending the IHEMT analysis to the CHEMT, the assumption is made that the operation of a CHEMT with uniform

doping density in the AlGaAs layer (Figure 3.16a) would not differ greatly (except for a change in threshold voltage) from the operation of a CHEMT where the dopant was placed in a thin sheet near the heterojunction (figure 3.16b). With this assumption, the thin dopant sheet and 2DEG can be treated together as a charge sheet in exactly the same way as in the IHEMT analysis. The dopant density in the doping plane would be increased from the uniform dopant density so that the dopant sheet charge in the two cases is the same. Hence the region II and III analysis for the IHEMT can be used directly for the CHEMT.

CHEMTs with a doping plane in the AlGaAs layer, as described above, have been fabricated at Hewlett-Packard Laboratories[3.68] and called 'pulse-doped MODFETs'. Apart from a difference in the threshold voltage, the devices have the same characteristics as conventional HEMTs, but with some reduction in gate leakage current and light sensitivity. Another variation on the pulse-doped MODFET has also been reported by Hida et al [3.65].

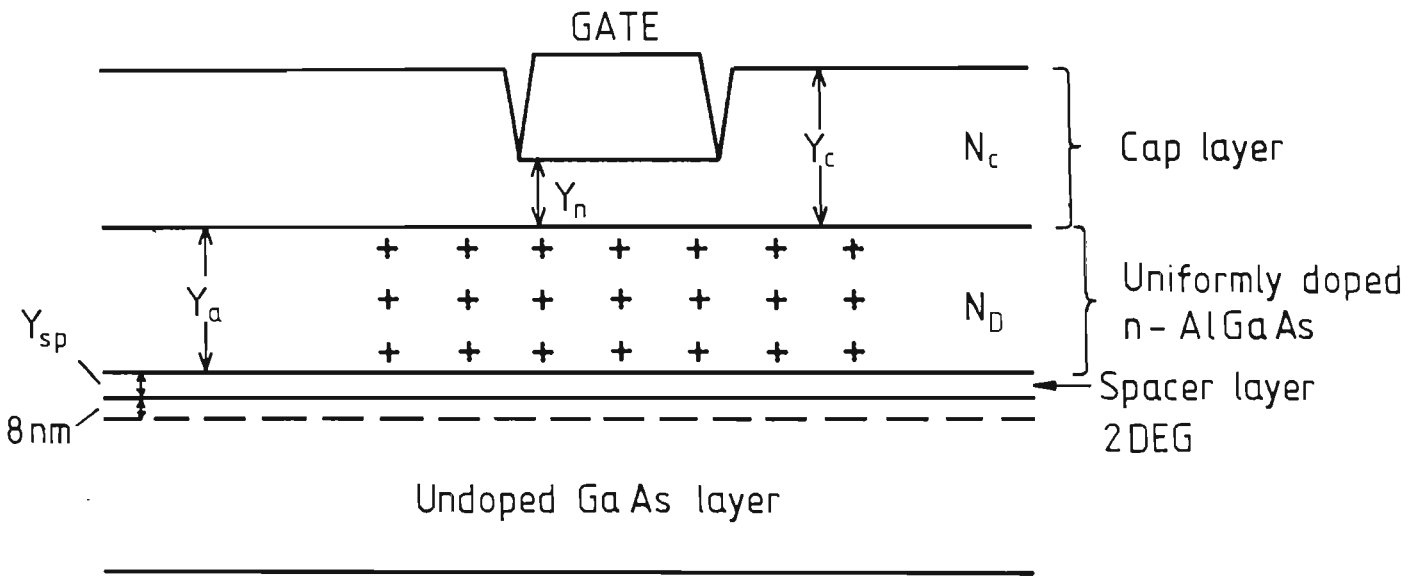
Threshold voltage

The threshold voltage for a CHEMT with uniform doping and no cap layer is found from the one-dimensional solution of Poisson's equation as [3.11, 3.13]:

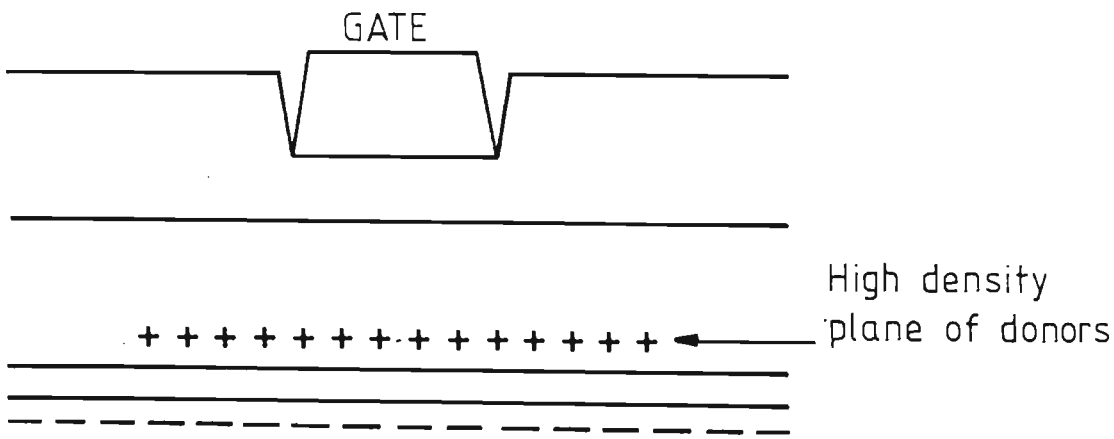
$$V_T = \phi_m - \Delta E_c - qN_D Y_a^2/2\varepsilon \qquad (3.56)$$

where ϕ_m and ΔE_c are the Schottky-barrier height and conduction band discontinuity in eV units, N_D is the AlGaAs doping density, Y_a is the doped AlGaAs thickness and ε is the AlGaAs permittivity.

Figure 3.16. CHEMT structures discussed in the text: (a) uniformly-doped AlGaAs and (b) plane-doped AlGaAs



(a)



(b)

When the gate electrode is on the cap layer (Figure 3.16a), the threshold voltage expression is modified to the form:

$$V_T = \phi_m - \Delta E_c - \frac{q}{2\varepsilon}\left(N_D Y_a^2 + 2N_D Y_a Y_n + N_c Y_n^2\right) \qquad (3.57)$$

where N_c is the cap layer doping density and Y_n is the cap layer thickness under the gate.

Charge control

The same linear charge control relation as in the IHEMT (equation 3.9).

$$n_s = \varepsilon(V_{GS}-V_T)/qd \qquad V_{GS} \leq V_{GSM} \qquad (3.58a)$$

where $d = Y_a + Y_{sp} + Y_n + 8$ nm (3.58b)

is used for the CHEMT. The gate electrode to 2DEG spacing, d, is given by the sum of the doped AlGaAs, undoped spacer (Ysp) and cap layer thickness together with the 8 nm correction for the displacement of the 2DEG centroid from the heterointerface [3.13]. CHEMTs exhibit some deviation from this linear relation near pinchoff because the triangular well widens significantly at this point (section 3.4.4). In spite of this, all published analytic HEMT models to date, which encompass more than just charge control theory, use this approximation [3.11, 3.13, 3.21, 3.40, 3.54, 3.58, 3.66].

As in the IHEMT, the gate voltage at which n_s reaches it maximum value, n_{so}, is given by (equation 3.10):

$$V_{GSM} = V_T + \frac{q}{\varepsilon} n_{so} d \qquad (3.59)$$

where d is given by equation 3.58b. For gate-source voltages geater than this value, the n-AlGaAs layer will become partially undepleted and result

in a parallel conduction path [3.20] between source and drain (see section 3.23).

2DEG current

Since the same velocity-electric field dependence is assumed for the IHEMT and CHEMT, the expression for the CHEMT 2DEG current in regions I and II have the same form as the IHEMT equations, with the substitution of the CHEMT relations for V_T and d (equations 3.56 -3.58).

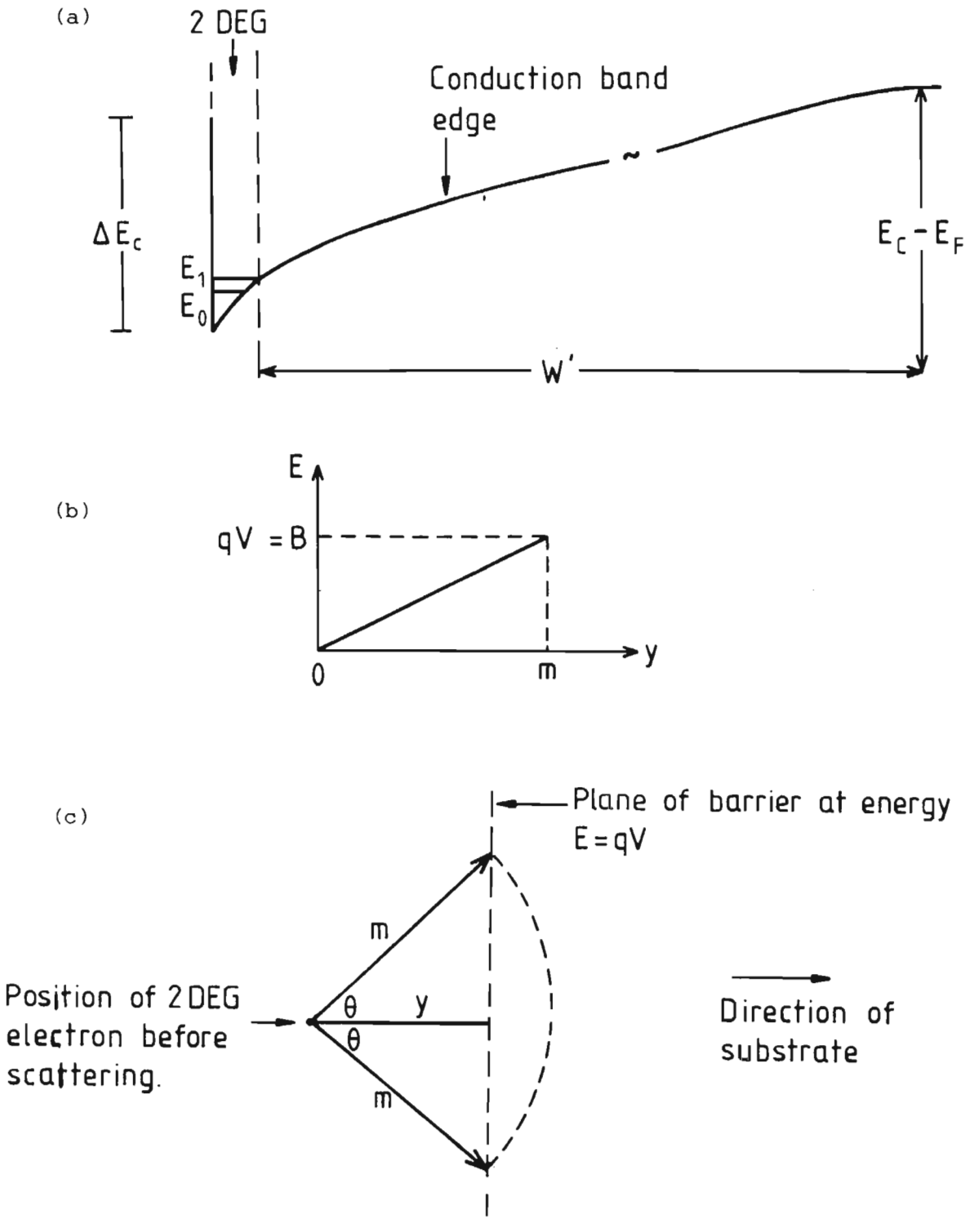
The longitudinal field and potential along the 2DEG in the three regions are also given by the same expressions as in the IHEMT analysis.

The CHEMT source and drain resistances are calculated using the expressions derived (equations 3.33 and 3.34) in the IHEMT analysis. These expressions will underestimate the CHEMT access resistances because the heterojunction at the cap layer -AlGaAs interface presents a barrier to electrons and hence increases the resistance [3.50]. However, the analysis of Feuer [3.50] is too complex to incorporate in this model as it would lead to an unacceptably large increase in computation time of the software program.

3.5.3 CHEMT substrate current

The origin of substrate current in CHEMTs is the same as in IHEMTs; electrons in the 2DEG become hot due to the high longitudinal fields in the device and may gain sufficient energy to be scattered out of the 2DEG and form a parallel current component in the substrate. The derivation of analytic expressions to estimate the substrate current component is, however, more complex in the CHEMT (and MESFET) than in the IHEMT because of the lack of a clearly defined barrier. In the IHEMT, the conduction band discontinuity at the heterojunction provides a clearly defined step in the energy band diagram whereas in the CHEMT (Figure 3.17a) the conduction band bends gradually towards the substrate.

Figure 3.17. (a) CHEMT band structure in GaAs layer (b) linearized barrier energy dependence on distance from 2DEG (c) scattering angle for 2DEG electron



Estimation of barrier height

The bandbending in the GaAs layer, outside of the 2DEG, is caused by ionised acceptors. Nominally undoped GaAs when grown by MBE is lightly p-type, with carbon as the main acceptor impurity [3.67]. The n-AlGaAs layer, which is the source of the 2DEG electrons, also supplies the electrons for the acceptors in the GaAs. In the neutral region, the energy difference between the Fermi-level and conduction band can be found from Boltzmann statistics (because the material is lightly doped) and noting that carbon is a shallow acceptor, as

$$E_c - E_F = E_G + kT \ln (N_a/N_v) \qquad (3.60)$$

where E_c and E_F are the conduction band energies, E_G is the GaAs energy bandgap, N_A is the acceptor concentration and N_v is the valence band of available states. Using the depletion approximation, the distance from the 2DEG to the neutral GaAs is [3.68]:

$$W = \sqrt{2\varepsilon (E_c - E_F)/q\ N_A} \qquad (3.61)$$

It can then be shown [3.69] that the increase in potential over any distance y from the 2DEG, towards the substrate, is given by:

$$V = qN_A y\ (w - y/2)/\varepsilon \qquad (3.62)$$

The average distance that an electron travels between scattering events is the mean free path which is approximately given by

$$m = \tau_e v_s \qquad (3.63)$$

where τ_e is the energy relaxation time (equation 3.39) and v_s is the saturated drift velocity of the 2DEG electrons. The increase in potential over this distance is obtained from (3.63) and (3.64) as:

$$V_b = q\, N_A\, m\, (w-m/2)/\varepsilon \qquad (3.64)$$

The assumption is made that any electron with energy greater than $B = qV_b$, and scattered towards the substrate, will give rise to substrate current because within one mean free path after scattering, the electron does not get reflected by the barrier.

Electrons with energy lower than the barrier, and scattered towards the substrate, will not necessarily impinge on the graded barrier, depending on the angle in which the electron is scattered. The potential dependence on distance from the 2DEG (equation 3.62) is linearized over the distance m (Figure 3.17b) to yield:

$$V(y) = \frac{y}{m} V_b \qquad (3.65)$$

This can be rearranged to give the effective distance from the 2DEG to the barrier at any potential, V, as:

$$y = mV/V_b \qquad (3.66)$$

Assuming isotropic scattering, as in the IHEMT, the probability that an electron will be scattered in angle 2θ is θ/π. At any energy $E = qV$, less than B, the angle θ is given by (figure 3.17c) $\cos^{-1}(y/m)$ and hence the probability that an electron will hit the barrier is given by:

$$P(E) = \frac{1}{\pi} \cos^{-1}\left(\frac{E}{B}\right) \qquad (3.67)$$

since it is assumed that the electron will travel a distance m before being scattered again.

The density of carriers in the substrate is then given by the total number of electrons scattered towards the substrate but which do not get reflected by the barrier. The fraction which hit the barrier is given by:

$$n_b = \int_0^B P(E)\, g(E)\, dE \tag{3.68}$$

and hence the density of carriers in the substrate is given by:

$$Ns_{ubs} = \tfrac{1}{2} \int_B^\infty f(E)g(E)\, dE + \tfrac{1}{2} \int_0^B f(E)g(E)\, dE - \int_0^B P(E)\, g(E)f(E)\, dE \tag{3.69}$$

The functions f(E) and g(E) are the Fermi-Dirac distribution function and 3D density of states functions, as used in the IHEMT analysis.

Equation 3.69 can be used in place of equation 3.46 in the IHEMT analysis to estimate the ratio N_{subs}/N for the CHEMT. The ratio (R) of substrate to 2DEG current is equal to the ratio N_{subs}/N because the substrate current flows in GaAs and consequently the saturated drift velocity for both current components is the same.

Limitations

The substrate current analysis presented above incorporates some assumptions also made for the IHEMT analysis but which are considered less valid for the CHEMT.

The most important of these is the assumption that scattering is isotropic. However, it is known [3.70, 3.71] that scattering by polar optical phonons, the dominant scattering mechanism for hot electrons in GaAs, is anisotropic. In the IHEMT, this presents no difficulty because the exact scattering angle is unimportant; the 2DEG centroid to barrier spacing is a small fraction of the mean free path which means that any electron scattered towards the heterojunction will cross the heterojunction (if the energy is large enough)

within one mean free path. In the CHEMT however, the scattering angle enters directly into the equations to calculate the number of electrons reflected by the barrier.

Also, effects which play no role in the IHEMT, such as barrier lowering and electron trapping in deep states in the substrate [3.29] are neglected in this CHEMT analysis.

In spite of these limitations, the substrate current analysis presented here provides a useful basis for incorporating substrate current in a CHEMT model. As in the IHEMT model, the inclusion of substrate current allows good agreement with experiment to be obtained without the use of saturated drift velocities greater than 10^7 cm/s.

3.5.4 CHEMT simulation results

To assess the usefulness of the CHEMT model, fits to published experimental data for a normally-on [3.13] and a normally-off (V_T > 0V) [3.11] CHEMT were made. These particular devices were chosen because several other models have been published with fits to the same data, and hence these devices provide a basis for comparing different models.

Normally-on CHEMT

Figure 3.18 shows the fit obtained for the normally-on device. The measured data [3.13] were scaled for a gate width of 1 mm. Fits to the same data were made in the original paper using a simple three piece model [3.13], and have also been made by Cil and Tansal [3.72] using a model based on the Trofimenkoff expression (section 3.2.2 and reference 3.14). The fits to the experimental data obtained with these models is shown, unscaled, in Figure

3.19 [3.72]. A comparison of the parameters used to compute these fits in the three models, together with the experimentally determined parameters, is tabulated in Table 3.2.

The first five entries in the table are the nominal values for the fabricated devices. The recess depth was unknown and hence was chosen in this work to give the best fit. The mobility and source resistance were measured by Lee et al. [3.13] and the measured values were used directly to obtain the fit in Figure 3.18.

Parameter	Experimental data [3.13]	Lee model [3.13]	Cil model [3.72]	This work	Units
L_g	1	1	1	1	μm
Z_g	145	145	145	145	μm
N_D	1	1	1	1	10^{18} cm^{-3}
Y_{SP}	6	6	4	6	nm
x	0.3	0.3	0.3	0.3	
Y_r	?	?	25	18	nm
μ_o	6800	6800	6800	6800	cm^2/Vs
R_S	7	7	10	7	ohm
R_D	?	?	10	8	ohm
v_s		2	2	1	10^7 cm/s

? means parameter not measured or not stated.

Table 3.2 Parameter comparison for three models used to fit data for normally-on CHEMT.

Figure 3.18. Fit to measured data (from ref. 3.13) for normally-on CHEMT, using the model described in this work.

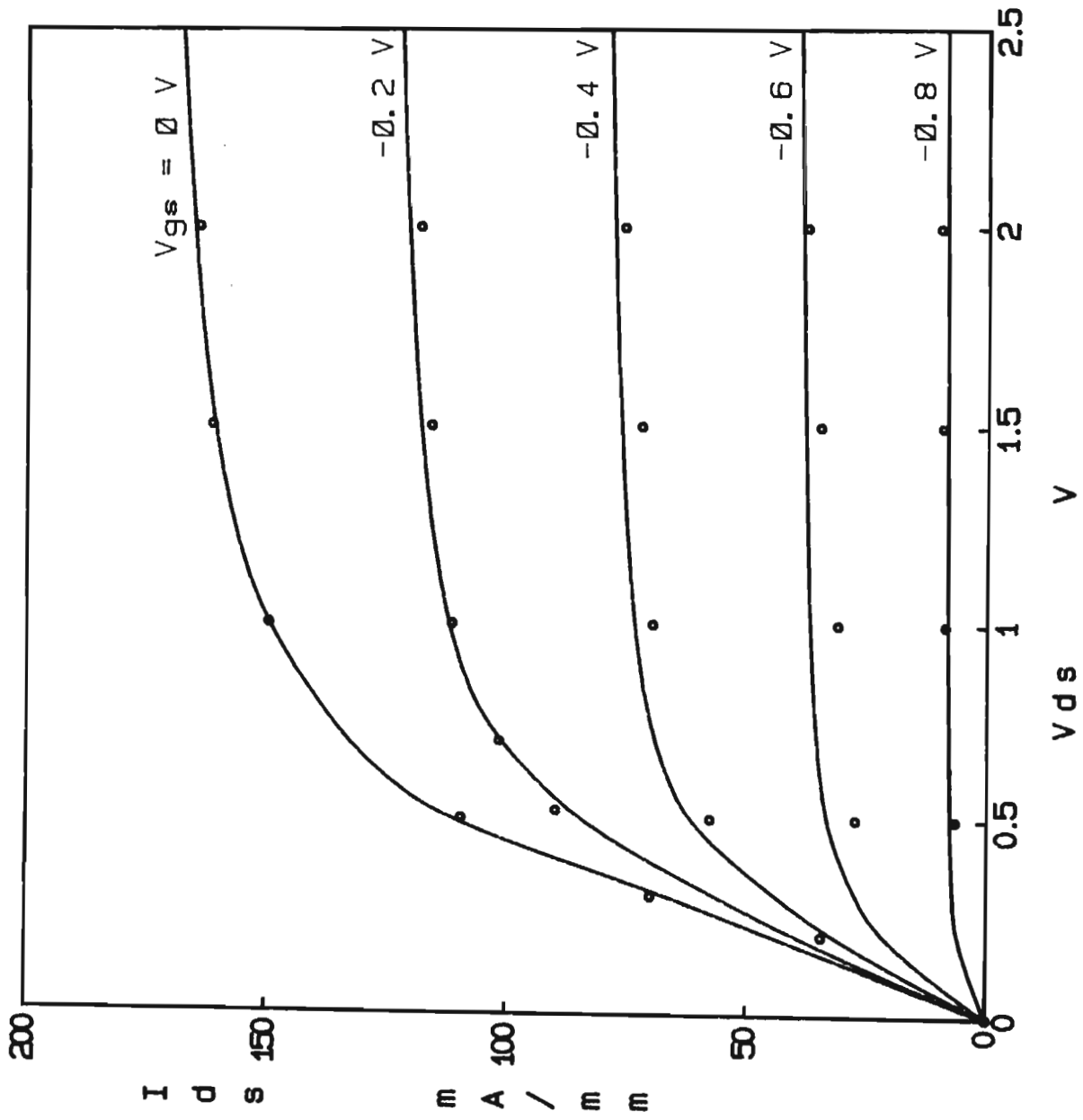
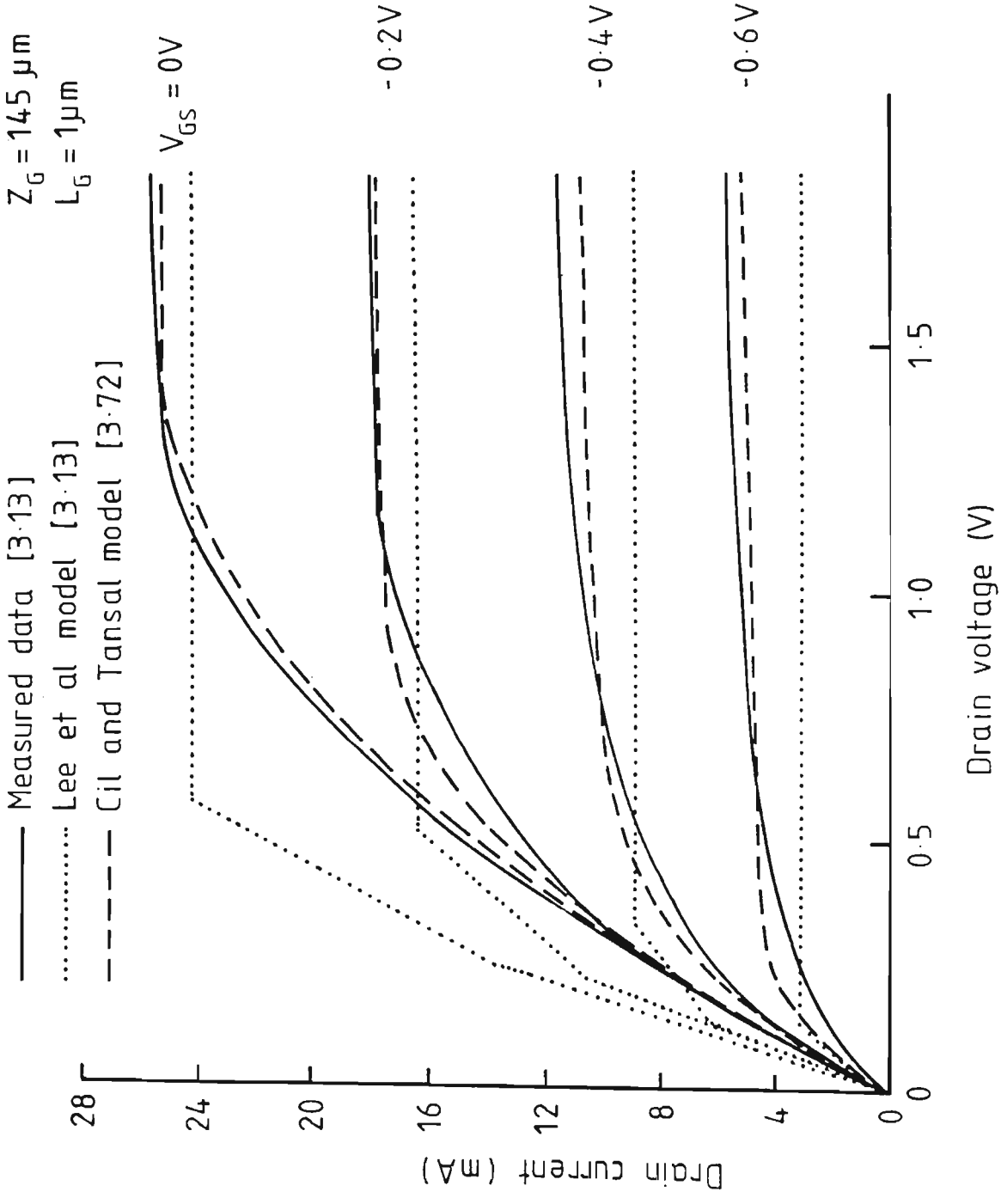


Figure 3.19. Fits to the data (unscaled) of Lee et al., as obtained by Lee et al. and by Cil and Tansal. From re. 3.72.



The fit obtained by Cil and Tansal [3.72] is good for the V_{GS} = 0 to V_{GS} = -0.6V curves. To obtain this good fit, the source resistance had to be increased by nearly 50 percent and a saturated drift velocity of $2x10^7$ cm/s was required. In addition, the threshold voltage was set at -0.8V, which resulted in a prediction of zero current for V_{GS} = -0.8V.

The inclusion of the substrate current estimation in the CHEMT model of this work results in a model which requires none of the parameter adjustment discussed above.

Normally-off CHEMT

Figure 3.20 shows the comparison between the simulated I_{DS}-V_{DS} characteristics from this work and the measured data of Drummond et al. [3.11] for a normally-off CHEMT. The measured data was scaled for a gate width of 1 mm. A similar comparison, unscaled,for the models of Drummond et al.[3.11] and Weiler and Ayasli [3.54] is shown in figure 3.21. An almost 'exact' fit to the measured data was obtained by Park and Kwack [3.58]. Their model parameters, together with those for the other models, are tabulated in Table 3.3

In common with the other models, it was found that a good fit could not be obtained using the nominal device parameters. It was necessary to increase the source resistance to one third above its measured value, and to assume a low value for the doping density obtain a fit. These adjustments are smaller than those of Weiler and Ayasli, who also used a low doping density, but more than doubled both mobility, source resistance and saturation velocity. In addition, finite drain conductance in saturation was obtained in their model by assuming a resistance of 3600 ohm in parallel with the modelled CHEMT. Park and Kwack also doubled the source resistance but did not adjust the doping density. The value of saturation velocity used in the their model was not stated.

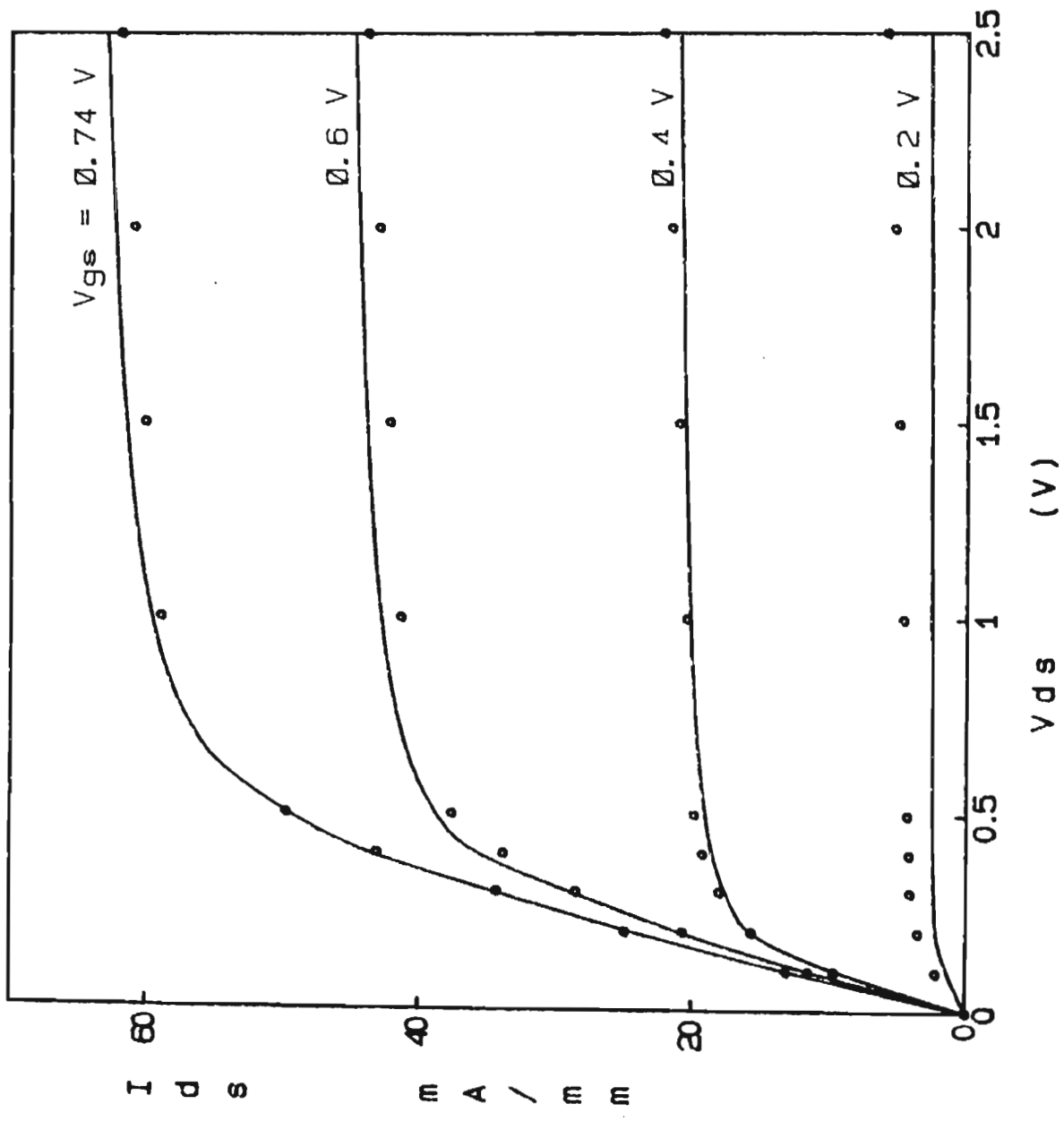
Parameter	Experimental data [3.11]	Weiler model [3.54]	Park model [3.58]	This work	Units
L_g	1	1	1	1	μm
Z_g	145	145	145	145	μm
N_D	1	0.65	1	0.41	10^{18} cm^{-3}
Y_{sp}	10	10	6	10	nm
x	0.3	0.3	0.3	0.3	
ΔE_c	0.24	0.32	0.32	0.24	eV
Y_r	?	25	35	10	nm
μ_o	4300	8000	4300	4300	cm^2/Vs
ϕ_m	?	1.1	1.1	1.1	eV
R_s	12	25	25	16	ohm
v_s		2.5	?	1	10^7 cm/s

Note: Drummond et al. used the experimental data in their model, and set v_s to 8×10^6 cm/s.

? means not stated. or unknown.

Table 3.3. Parameter comparison for three models used to fit data for normally-off experimental data.

Figure 3.20. Fit to measured data from reference 3.11 for normally-off CHEMT, using the model described in this work.



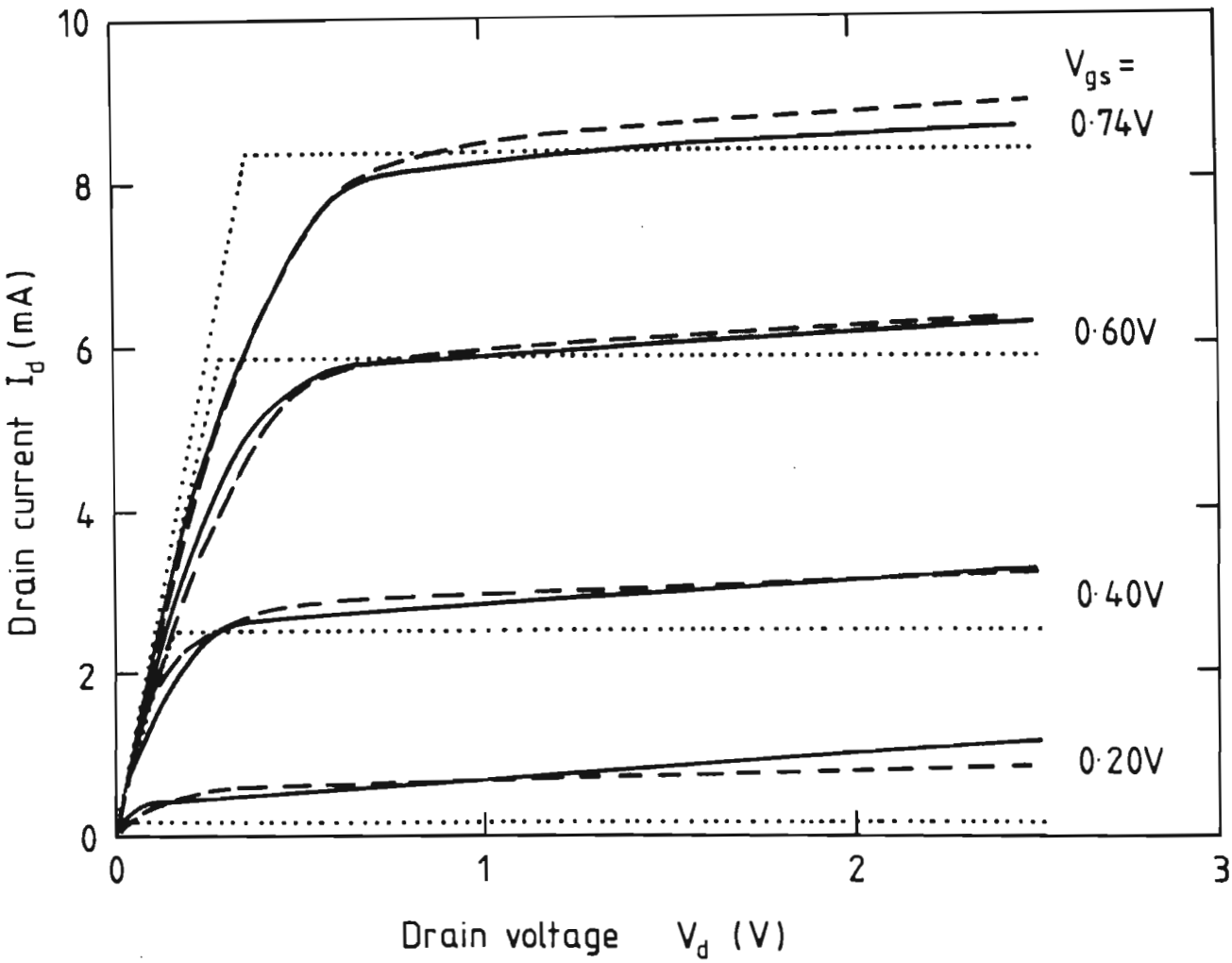


Figure 3.21. Fits to the data of reference 3.11 (dashed line), as obtained by Drummond et al. (dotted line) and by Weiler and Ayasli (solid line). From reference 3.54.

It is not clear why the agreement is poor using the nominal values. The measured transconductance of this device is unexpectedly low (around 100 ms/mm) in comparison with other reported device characteristics obtained on similar devices [3.73], and hence it is considered likely that the nominal layer parameters were not realised in the practical devices.

A further possible source of discrepancy is the assumed value of the Schottky barrier height. All the models discussed above assumed that the dependence of Schottky-barrier height on aluminium mole fraction is given by [3.74]:

$$\phi_m = 0.92 + 0.62x \quad (eV) \tag{3.70}$$

This expression was obtained for a gold gate on AlGaAs. It is well known however, that the barrier height is a function of both doping density and surface preparation, as well as mole fraction, and hence (3.70) may introduce significant errors here. Others [3.75] have deduced values between 0.6 and 1.2 eV for the barrier height from measurements on a range of HEMT structures.

3.6 SUMMARY

In this chapter, a new model for the IHEMT has been presented. Apart from the charge control analysis of Lee et al. [3.19], no other model for the IHEMT has been published to date. The development of the model has been discussed and the model's usefulness in predicting IHEMT DC performance, demonstrated.

The extension of the analysis to incorporate CHEMT modelling has also been presented and has been shown to provide reasonable agreement with published experimental data.

The analysis includes several effects which are not considered in other HEMT models; in particular, the extension of the active region beyond the drain edge of the gate, the effect of surface charge and the real space transfer of hot electrons into the layer under the 2DEG (substrate current).

The model calculates the DC characteristics from the fundamental physical properties of the device and, in spite of the overtly approximate nature of some assumed dependences, the agreement between measured and simulated HEMT performance is good. Importantly, for a physically-based model, the physical quantities used in the model are not required to assume non-physical values to obtain good agreement. In particular, it is not found necessary to use the saturated drift velocity as a fitting parameter.

This work provides a basis for explaining the effects of surface charge in HEMTs. More work by materials scientists is required to determine the exact dependences of surface charge on bias, frequency and processing techniques, before it can be demonstrated that this model gives an accurate quantitative prediction of the relationship between DC electrical parameters and occupied surface state density.

The IHEMT analysis could readily be applied to the analysis of GaAs MISFETs [3.76] , which are receiving increasing attention in the literature. The surface state density is better understood in MIS structures than in HEMTs because of the similarity of the former structure to silicon MOS devices.

In this model, good agreement between simulation and measured results were obtained using a saturated drift velocity equal to the electron saturation velocity measured in bulk GaAs, 10^7 cm/s [3.51]. This agreement does not constitute proof of no velocity overshoot in HEMTs, but does indicate that there is need for a drastic lowering of the assumed velocities in these devices from the values of around $2x10^7$ cm/s commonly reported in the literature. No attempt has been made in this work to simulate devices operating at 77K, where velocities as large as $3x10^7$ cm/s have been fitted to measured data. This is clearly an area for further investigation in the future.

The assessment of the value of any model, and the validity of the approximations in it, can only be finally determined by an extensive series of experiments in which accurate measurements are made on devices which can be repeatably produced. The experimentally determined dependences of electrical characteristics on single changes in device structural and material parameters will determine the suitability of each approximate dependence used in the model. These procedures require many man-years of effort. It is the author's view that the model presented here constitutes a sound basis for undertaking such a study in the future.

CHAPTER 3 REFERENCES

3.1 Widiger D.J., Kizilyalli I.C., Hess K. and Coleman J.J., "Two-Dimensional Transient Simulation of an Idealized High Electron Mobility Transistor", IEEE Trans. Elec. Dev., Vol. ED-32, No. 6, June 1985, pp. 1092-1102

3.2 Yoshida J. and Kurata M., "Analysis of High Mobility Transistors Based on a Two-Dimensional Numerical Model", IEEE Elec. Dev. Lett., Vol. EDL-5, No. 12, Dec. 1984, pp. 508-510

3.3 Ravaioli U. and Ferry D.K., "MODFET Ensemble Monte Carlo Model Including the Quasi-Two-Dimensional Electron Gas", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 677-681

3.4 Tomizawa M., Yoshii A. and Yokoyama K., "Modeling for an AlGaAs/GaAs Heterostructure Device Using Monte Carlo Simulation", IEEE Elec. Dev. Lett., Vol. EDL-6, No. 7, July 1985, pp. 332-334

3.5 Bennett H.S., "Modeling GaAs/AlGaAs Devices: A Critical Review", IEEE Circuits and Devices Magazine, Jan. 1985, pp. 35-42

3.6 Kacprzak T. and Materka A., "Compact dc Model of GaAs FET's for Large-Signal Computer Calculation", IEEE Jnl. Solid-State Circuits, Vol. SC-18, No. 2, April 1983, pp. 211-213

3.7 Taki T., "Approximation of Junction Field-Effect Transistor Characteristics by a Hyperbolic Function", IEEE Jnl. Solid-State Circuits, Vol. SC-13, No. 5, Oct. 1978, pp. 724-726

3.8 Abuelmaatti M.T., "Modelling DC Characteristics of HEMTs", Elec. Lett., Vol. 21, No. 2, Jan. 1985, pp. 69-70

3.9 Masselink W.T., Kopp W., Henderson T. and Morkoc H., "Measurement of the Electron Velocity-Field Characteristic in Modulation-Doped Structures Using the Geometrical Magnetoresistance Method", IEEE Elec. Dev. Lett., Vol. EDL-6, No. 10, Oct. 1985, pp. 539-541

3.10 Masselink W.T., Henderson T.S., Klem J., Kopp W. and Morkoc H., "The Dependence of 77K Electron Velocity-Field Characteristic on Low-Field Mobility in AlGaAs-GaAs Modulation-Doped Structures", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 639-644

3.11 Drummond T.J., Morkoc H., Lee K. and Shur M., "Model for Modulation Doped Field Effect Transistor", IEEE Elec. Dev. Lett., Vol. EDL-3, No. 11, Nov. 1982, pp. 338-341

3.12 Hill A.J., Ladbrooke P.H., Ransome S. and Westwood D., "Large Signal Modelling and Practical Performance of Inverted HEMT", Proc. IEEE/Cornell Conf., July 1985, pp. 128-135

3.13 Lee K., Shur M.S., Drummond T.J. and Morkoc H., "Current-Voltage and Capacitance-Voltage Characteristics of Modulation-Doped Field-Effect Transistors", IEEE Trans. Elec. Dev., Vol. ED-30, No. 3, March 1983, pp 207-212

3.14 Trofimenkoff F.N., "Field-Dependent Mobility Analysis of the Field Effect Transistor", Proc. IEEE, 53, 1965, pp. 1765-1766

3.15 Hariu T., Takahashi K. and Shibata Y., "New Modeling of GaAs MESFET's", IEEE Trans. Elec. Dev., Vol. ED-30, No. 12, Dec. 1983, pp. 1743-1749

3.16 Lee K., Shur M.S., Drummond T.J. and Morkoc H., "Electron Density of the Two-dimensional electron gas in modulation doped layers", J. Appl. Phys., Vol. 54, 1983, pp 2093-2096

3.17 Vinter B., "Subbands and charge control in a two-dimensional gas field-effect transistor", Appl. Phys. Lett., Vol. 44(3), Feb. 1984, pp. 307-309

3.18 Hihara H. and Hamaguchi C., "Self-consistent calculations of the density of 2D electron gas in AlxGa1-xAs/GaAs heterointerface as a function of the doping level in n-AlxGa1-xAs layer", Solid State Comm., Vol. 54, No. 6, 1985, pp. 485-488

3.19 Lee K., Shur M., Drummond T.J. and Morkoc H., "Charge control model of inverted GaAs-AlGaAs modulation doped FET's (IMODFET'S), J. Vac. Sci. Technol., B 2 (2), Apr. 1984, pp. 113-116

3.20 Lee K., Shur M.S., Drummond T.J. and Morkoc H., "Parasitic MESFET in (Al,Ga) As/GaAs Modulation Doped FET's and MODFET Characterization", IEEE Trans. Elec. Dev., Vol. ED-31, No. 1, Jan 1984, pp. 29-35

3.21 Hida H., Itoh T. and Ohata K., "An Accurate dc Model of 2-DEG FET for Implementation on a Circuit Simulator", IEEE Elec. Dev. Lett., Vol. EDL-7, No. 6, June 1986, pp. 393-395

3.22 Grebene A.B. and Ghandhi S.K., "General theory for pinched operation of the junction-gate FET", Sol.- St. Elec., Vol. 12, 1969, pp. 573-589

3.23 Pucel R.A., Haus H.A. and Statz H., "Signal and Noise Properties of Gallium Arsenide Microwave Field-Effect Transistors", Advances in Electronics and Electron Physics, Vol. 38, New York: Academic Press, 1975, pp. 195-265

3.24 Sone J. and Takayama Y., "A Small-Signal Analytical Theory for GaAs Field-Effect Transistors at Large Drain Voltages", IEEE Trans. Elec. Dev., Vol. ED-25, No. 3, March 1978, pp. 329-337

3.25 Lehovec K. and Miller R.S., "Field Distribution in Junction Field-Effect Transistors at Large Drain Voltages", IEEE Trans. Elec. Dev., Vol. ED-22, No. 5, May 1975, pp. 273-281

3.26 Hess K., Morkoc H., Shichijo H. and Streetman B.G., "Negative differential resistance through real-space transfer", Appl. Phys. Lett., Vol. 35(6), Sept. 1979, pp. 469-471

3.27 Keever M., Shichijo H., Hess K., Banerjee S., Witkowski L., Morkoc H. and Streetman B.G., "Measurements of hot-electron conduction and real-space transfer in GaAs-AlxGa1-xAs heterojunction layers", Appl. Phys. Lett., Vol. 38(1), Jan. 1981, pp. 36-38

3.28 Matsumoto K., Hashizume N., Atoda N. and Awano Y., "Effects of p-type barrier layer on characteristics of sub-micron gate self-aligned GaAs FET", Proc. Inst. Symp. GaAs and Related Compounds 1984, Inst. Phys. Conf. Ser. No. 74, pp. 515-520

3.29 Barton T.M., Richardson J.S. and Snowden C.M., "Modelling of Recessed Gate MESFET Structures", 3rd. GaAs Simulation Workshop, Duisberg, W. Germany, Oct. 1986

3.30 Judaprawira S., Wang W.I., Chao P.C., Wood C.E.C., Woodard D.W. and Eastman L.F., "Modulation-doped MBE GaAs/n-AlxGa1-xAs MESFETs", IEEE Elec. Dev. Lett., Vol. EDL-2, No. 1, Jan. 1981, pp. 14-15

3.31 Camnitz L.H., Maki P.A., Tasker P.J. and Eastman L.F., "Sub-Micrometer Quantum Well HEMT with $Al_{0.3}Ga_{0.7}As$ Buffer Layer", 1984 GaAs and Related Compounds, Inst. Phys. Conf. Ser. 74, pp. 333-338

3.32 Stratton R., "Diffusion of Hot and Cold Electrons in Semiconductor Barriers", Physical Review, Vol. 126, No. 6, June 1962, pp. 2002-2014

3.33 Huang R.S. and Ladbrooke P.H., "The physics of excess electron velocity in submicron-channel FET's", J. Appl. Phys., Vol. 48, No. 11, Nov. 1977, pp. 4791-4798

3.34 Loret D., Baets R., Snowden C.M. and Hughes W.A., "Two-Dimensional Numerical Models for the High Electron Mobility Transistor", in "Simulation of Semiconductor Devices and Processes Vol. 2", eds. K. Board and Owen D.R.J., Pineridge Press, 1986

3.35 Hill A.J. and Ladbrooke P.H., "Dependence of conduction-band discontinuity on aluminium mole fraction in GaAs/AlGaAs heterojunctions", Elec. Lett., Vol. 22, No. 4, Feb. 1986, pp. 218-220

3.36 Blakemore J.S., "Intrinsic density ni(T) in GaAs: Deduced from band gap and effective mass parameters and derived independently from Cr acceptor capture and emission coefficients", J. Appl. Phys., Vol. 53(1), Jan. 1982, pp. 520-531

3.37 Pierret R.F., "Extension of the Approximate Two-Dimensional Electron Gas Formulation", IEEE Trans. Elec. Dev., Vol. ED-32, No. 7, pp. 1279-1287

3.38 Yoshida J., "Classical Versus Quantum Mechanical Calculation of the Electron Distribution at the n-AlGaAs/GaAs Heterointerface", IEEE Trans. Elec. Dev., Vol. ED-33, No. 1, Jan. 1986, pp. 154-156

3.39 Lee K., Shur M.S., Drummond T.J., Su S.L., Lyons W.G., Fischer R. and Morkoc H., "Design and fabrication of high transconductance modulation-doped (Al,Ga) As/GaAs FETs", J. Vac. Sci. Technol., Vol. JVST B1, 1983, pp 186-189

3.40 Delagebeaudeuf D. and Linh N.T., "Metal-(n) AlGaAs-GaAs Two-Dimensional Electron Gas FET", IEEE Trans. Elec. Dev., Vol. ED-29, No. 6, June 1982, pp 955-960

3.41 Aymerich-Humet X., Serra-Mestres F. and Millan J., "An Analytical Approximation for the Fermi-Dirac Integral F3/2(ℸ)", Sol.-St. Elec., Vol. 24, No. 10, 1981, pp. 981-982

3.42 Sugeta T., Majerfeld A., Saxena A.K., Robson P.N. and Hill G., "High-Field Transport Properties of $Ga_{1-x}Al_xAs$", Proc. 6th. Biennial IEEE/Cornell Conf., 1977, pp. 45-53

3.43 Hill G. and Robson P.N., "Hot Electron Transport in the $Ga_{1-x}Al_xAs$ system", Proc. 3rd. Intl. Conf. on Hot Carriers in Semiconductors, Journal de Physique, Oct. 1981, pp. (C7)335-(C7)341

3.44 Hirano M., Takanashi Y. and Sugeta T., "Current-Voltage Characteristics of an AlGaAs/GaAs Heterostructure FET for High Gate Voltages", IEEE Elec. Dev. Lett., Vol. EDL-5, No. 11, Nov. 1984, pp. 496-499

3.45 Bhattacharya P.K., Das U. and Ludowise M.J., "Transport properties of n-type metalorganic chemical-vapor-deposited AlxGa1-xAs (0ẋ0.6)", Phys. Rev. B, Vol. 29, No. 12, June 1984, pp. 6623-6631

3.46 Okamoto K., Wood C.E.C. and Eastman L.F., "Schottky-barrier heights of molecular beam epitaxial metal-AlGaAs structures", Appl. Phys. Lett., Vol. 38, 1981, pp. 636-638

3.47 Barton T.M. and Ladbrooke P.H., "The Role of the Device Surface in the High Voltage Behaviour of the GaAs MESFET", Sol.-St. Elec., Vol. 29, No. 8, 1986, pp. 807-813

3.48 Miller D.L., "MBE material considerations for HEMT circuits", Proc. IEEE GaAs IC Symposium, Boston 1984, pp. 37-40

3.49 Hikosaka K., Mimura T. and Joshin K., "Selective Dry Etching of AlGaAs-GaAs Heterojunction", Jpn. J. Appl. Phys., Vol. 20, No. 11, Nov. 1981, pp. L847-L850

3.50 Feuer M.D., "Two-Layer Model for Source Resistance in Selectively Doped Heterojunction Transistors", IEEE Trans. Elec. Dev., Vol. ED-32, No. 1, Jan. 1985, pp. 7-11

3.51 Ruch J.G. and Kino G.S., "Measurement of the Velocity-Field Characteristics of Gallium Arsenide", Appl. Phys. Lett., Vol. 10, 1967, pp. 40

3.51a Braslau N. and Hauge P.S., "Microwave Measurement of the Velocity-Field Characteristic of GaAs", IEEE Trans. Elec. Dev., Vol. ED-17, No. 8, Aug. 1970, pp. 616-622

3.51b Fawcett W., Boardman A.D. and Swain S., "Monte Carlo Determination of Electron Transport Properties in Gallium Arsenide", J. Phys. Chem. Solids, Vol. 31, 1970, pp. 1963-1990

3.52 Tomizawa M., Yokoyama K. and Yoshii A., "Hot-Electron Velocity Characteristics at AlGaAs/GaAs Heterostructures", IEEE Elec. Dev. Lett., Vol. EDL-5, No. 11, Nov. 1984, pp. 464-465

3.53 Das M.B., Kopp W. and Morkoc H., "Determination of Carrier Saturation Velocity in Short-Gate-Length Modulation-Doped FET's", IEEE Elec. Dev. Lett., Vol. EDL-5, No. 11, Nov. 1984, pp. 446-449

3.54 Weiler M.H. and Ayasli Y., "DC and Microwave Models for AlxGa1-xAs/GaAs High Electron Mobility Transistors", IEEE Trans. Elec. Dev., Vol. ED-31, No. 12, Dec. 1984, pp. 1854-1861

3.55 Su S.L., Fischer R., Lyons W.G., Thorne R.E., Kopp W. and Morkoc H., "Modulation-Doped (Al.Ga)As/GaAs FETs with High Transconductance and Electron Velocity", Elec. Lett., Vol. 18, No. 18, Sept. 1982, pp. 794-796

3.56 Widiger D.J., Kizilyalli I.C., Hess K. and Coleman J.J., "Two-Dimensional Numerical Analysis of the High Electron Mobility Transistor", Superlattices and Microstructures, Vol. 1, No. 6, 1985, pp. 465-470

3.57 Al-Mudares M., "Modelling of GaAs-AlGaAs HEMTs", 3rd. GaAs Simulation Workshop, Duisberg, W. Germany, Oct. 1986

3.58 Park K. and Kwack K.D., "A Model for the Current-Voltage Characteristics of MODFET's", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 673-676

3.59 Heliodore F., Lefebvre M., Salmer G. and El Sayed O., "Modeling of Surface Depletion Effects in Submicronic MESFET's", 3rd. GaAs Simulation Workshop, Duisberg, W. Germany, Oct. 1986

3.60 Sze S.M., "Physics of Semiconductor Devices", 2nd. Edition, John Wiley and Sons, 1981, pp. 270-276

3.61 Ohnishi T., Onodera T., Yokoyama N. and Nishi H., "Comparison of the Orientation Effect of SiO_2 and Si_3N_4-Encapsulated GaAs MESFET's", IEEE Elec. Dev. Lett., Vol. EDL-6, No. 4, April 1985, pp. 172-174

3.62 Ozeki M., Kodama K. and Shibatomi A., "Surface Analysis of GaAs MESFETs by gm Frequency-Dispersion Measurement of Transconductance", Fujitsu Sci. Tech. Jnl., Vol. 18, No. 4, Dec. 1982, pp. 475-486

3.63 Blight S.R., Wallis R.H. and Thomas H., "Surface Influence on the Conductance DLTS Spectra of GaAs MESFET's", IEEE Trans. Elec. Dev., Vol. ED-33, No. 10, Oct. 1986, pp. 1447-1453

3.64 Hueschen M., Moll N., Gowen E. and Miller J., "Pulse Doped MODFETs", Proc. IEDM 1984, San Fransisco, pp. 348-351

3.65 Hida H., Miyamoto H., Ohata K., Itoh T., Baba T. and Ogawa M., "Planar AlGaAs/GaAs selectively doped structure with high performances and high stabilities". 1984 GaAs and Related Compounds, Inst. Phys. Conf. Ser. 74, pp. 551-556

3.66 Das M.B. and Roszak M.L., "Design calculations for submicron gate-length AlGaAs/GaAs modulation-doped FET structures using carrier saturation velocity/charge-control model", Sol. St. Elec., Vol. 28, No. 10, 1985, pp. 997-1005

3.67 Wood C.E.C., "Dopant Incorporation, Characteristics, and Behaviour", in "Technology and Physics of Molecular Beam Epitaxy", ed. E.H.C. Parker, Wiley 1986

3.68 Kastalsky A and Hwang J.C.M., "Illumination stimulated persistent channel depletion at selectively doped $Al_{0.3}Ga_{0.7}As$/GaAs interface", Appl. Phys. Lett., Vol. 44(3), Feb. 1984, pp. 333-335

3.69 Hill A.J., "Barrier Height for CHEMT Substrate Current", Internal note, July 1986

3.70 Carrol J.E., Hot electron microwave generators", Edward Arnold Pub., 1970, Chapter 2

3.71 Bandyopadhyay S., Klausmeier-Brown M.E., Maziar C.M., Lundstrom M.S. and Datta S., "Transport parameters in sub-micron devices", in "Simulation of Semiconductor Devices and Processes Vol. 2", eds. K. Board and Owen D.R.J., Pineridge Press, 1986, pp. 217-231

3.72 Cil C.Z. and Tansal S., "A New Model for Modulation-Doped FET's", IEEE Elec. Dev. Lett., Vol. EDL-6, No. 8, Aug. 1985, pp. 434-436

3.73 Laviron M., Delagebeaudeuf D., Delescluse P., Etienne P., Chaplart J. and Linh N.T., "Low noise normally on and normally off two-dimensional electron gas field-effect transistors", Appl. Phys. Lett., Vol. 40, No. 6, March 1982, pp. 530-532

3.74 Best J.S., "The Schottky-barrier height of Au on n-$Ga_{1-x}Al_xAs$ as a function of AlAs content", Appl. Phys. Lett., Vol. 34, Apr. 1979, pp. 522-524

3.75 Dambkes H., personal communication, Oct. 1986

3.76 Hill P.M., "A Comprehensive Analytical Model for III-V Compound Semiconductors", IEEE Trans. Elec. Dev., Vol. ED-32, No. 11, Nov. 1985, pp. 2249 2256

CHAPTER FOUR

HEMT RF MODEL

4.1 INTRODUCTION

In this chapter, a new RF model for HEMTs is presented. With the exception of some parasitic elements (to be discussed), the RF equivalent circuit elements are calculated using the DC model described in Chapter Three. The calculation of S-parameters and several amplifier gains, from the RF model, are also discussed. Comparisons between measured and predicted RF performances of IHEMTs are also presented.

The aim of this work was to develop, using the DC model, a capability to predict the dependences of HEMT RF performance on the structural and material parameters, and bias, of the device. No other models, for either CHEMT or IHEMT, with this objective have appeared in the literature to date.

Physically-based models of this kind have important applications in circuit design. One example of this is in worst-case or yield analyses of analogue MMIC circuit designs.

For a given design, it is important to establish the sensitivity of the overall circuit performance to changes in active device characteristics. It is common to do this by varying (independently) the equivalent circuit elements by a given percentage and calculating, for each combination, the overall circuit performance. A microwave circuit simulation program, such as SUPER COMPACT [4.1], is usually used for these calculations. From this, a worst-case element combination and a yield estimate may be derived.

There are two main limitations to this approach, if the equivalent circuit elements are not derived from a physically-based model. The first is that the equivalent circuit elements do not vary independently in practical devices. For example, if the undoped GaAs layer in an IHEMT were thicker

than the nominal value for a particular device, both the transconductance and gate-source capacitance would be lower than their respective nominal values by a similar proportion. Hence the 'worst-case' combination of equivalent circuit elements may not represent a physically realisable device. Secondly, if the model is not physically-based, it will not be possible to relate acceptable equivalent circuit element tolerances to permissible structural and material parameter variations. Thus it will not be possible to establish whether a given device fabrication process is sufficiently well controlled to produce the MMIC with the required yield.

Throughout this chapter, the discussion refers mainly to the IHEMT. This is simply for convenience and the analysis applies exactly for the CHEMT as well. In the next subsection, the calculation of the RF equivalent circuit elements is presented. In section 4.3 the calculation of S-parameters and device gains are discussed. Section 4.4 contains a comparison between predicted and measured HEMT RF characteristics. Bias dependences of the equivalent circuit elements are also discussed. Finally, section 4.5 contains a discussion and summary of the work. A list of references cited in this chapter appears after the summary.

The RF model presented in this chapter is not completely physically-based as some of the equivalent circuit elements are not calculated (such as the gate resistance), and the output resistance is scaled by an empirically determined constant. This work does however represent a substantial first step on the path to a fully physically-based model.

4.2 RF EQUIVALENT CIRCUIT

4.2.1 Circuit configuration

The complete RF equivalent circuit for the IHEMT is shown in Figure 4.1. Component values which are not calculated from the DC model are indicated by asterisks. These consist of parasitic capacitances and inductances inherent in any measurement system and device layout, used to measure device RF characteristics. In addition, the gate resistance and transconductance time constant, τ, are not calculated in this work. Since each of the elements calculated from the DC model is bias dependent, the circuit is applicable to both large (non-linear) and small signal analyses. Each of the 17 components in the equivalent circuit has a physical origin and each will be discussed in the following two subsections.

In this work it is assumed that the IHEMT gate width, Z_G, is small enough to allow the device to be represented by a lumped element equivalent circuit (Figure 4.1), rather than a distributed element formulation [4.2].

The equivalent circuit in Figure 4.1 contains too many elements to allow each one to be uniquely determined from a set of measured S-parameters [4.3-4.5]. To facilitate comparison with experimental data, a simplified equivalent circuit (Figure 4.2) is derived from the equivalent circuit of Figure 4.1. The simplified equivalent circuit is similar to that used in many FET and HEMT models [4.5-4.8] and is discussed further in Section 4.3.

4.2.2 Calculation of elements from DC model

Source and drain resistances

The values for the source resistance, R_s, and drain resistance, R_d, are calculated (section 3.2.6) in the DC model in order to find the DC solution at a given bias point. These values are used directly in the equivalent circuit.

Figure 4.1. Complete RF equivalent circuit for IHEMT, including parasitics

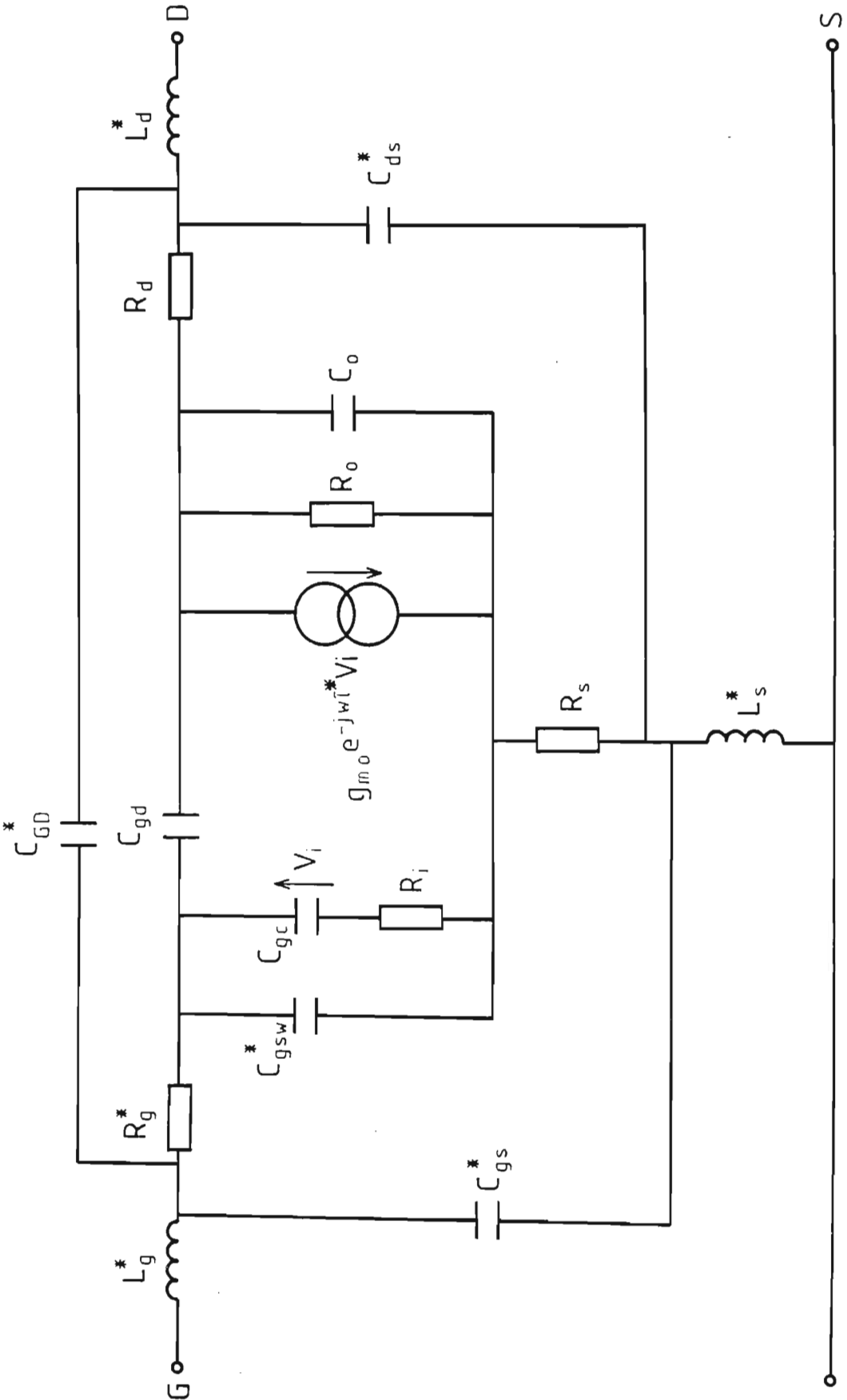
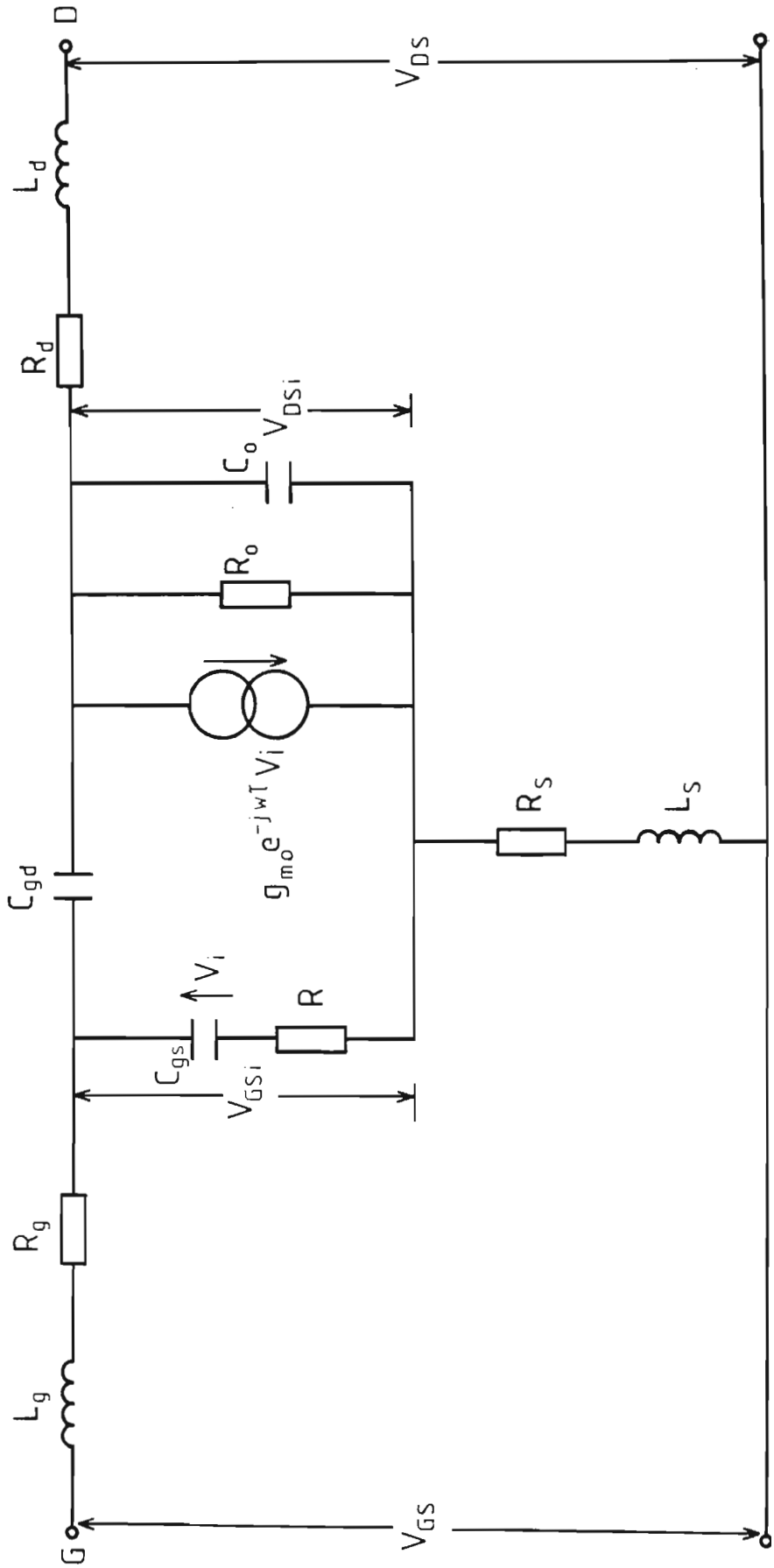


Figure 4.2. Simplified equivalent circuit for IHEMT.



The drain resistance is bias dependent because the length of the access region is decreased as the length of region III is increased (equation 3.34). Since the length of region III increases with increased drain voltage (all other parameters held constant), the drain resistance will decrease with increasing drain bias. In practical devices, the contact resistance term is usually significantly larger than the access region resistance and hence the drain resistance variation is negligible. If this were not the case, however, empirical methods [4.9, 4.10] for determining the source and drain resistance, from measurements on devices at very low applied drain/source bias (around 50 mV), could introduce significant errors in the predicted equivalent circuit elements.

Both source and drain resistances may be frequency dependent because of the frequency dependence of occupied surface state density, n_{ss}, which appears in the expressions for these resistances (equations 3.33 and 3.34). Again, this effect is usually negligible because the contact resistance is the larger component of the total resistance. If the contact resistance were substantially smaller than the resistance of the access regions, R_s and R_d could only be obtained in practical devices from the measured S-parameters.

Intrinsic channel resistance

The intrinsic channel resistance, R_i, is calculated from the effective resistance of Region I of the intrinsic IHEMT. (The term channel is from GaAs MESFET modelling [4.6], and is used here for the IHEMT because the effect of R_i is the same in both devices). Since the intrinsic channel resistance forms a distributed network with C_{gc} (Figure 3.1), in a similar way to the gate resistance, R_g, the RF effective value of R_i is reduced to approximately one third of the DC resistance (reference 4.9, equation 47). Hence R_i is given by:

$$R_i = (V_1 - I_{DS}R_s)/3I_{DS} \tag{4.1}$$

Where V_1 is the potential, calculated in the DC model, at the boundary between regions I and II. The term in parenthesis is the potential across region I and I_{DS} is the current flowing through the region. The intrinsic channel resistance is clearly bias dependent because both V_1 and I_{DS} are bias dependent.

Elements involving two bias points

The next four equivalent circuit elements to be discussed all involve incremental changes in either the applied drain-source or gate-source voltage.

The magnitude of the intrinsic transconductance is defined as:

$$g_{mo} = \partial I_{DS}/\partial V_{Gsi},\ V_{DS} = \text{constant} \tag{4.2}$$

where V_{Gsi} is the effective gate-source voltage across the intrinsic IHEMT. The gate-channel capacitance is calculated from:

$$Cgc = \partial Q/\partial V_{Gsi},\ V_{DS} = \text{constant} \tag{4.3}$$

where ∂Q is the change in total active region charge. Similarly, the output resistance, R_o, and gate-drain capacitance, C_{gd}, are obtained from:

$$R_o = \partial V_{DSi}/\partial I_{DS},\ V_{GS} = \text{constant} \tag{4.4}$$

$$C_{gd} = \partial Q/\partial V_{DSi},\ V_{GS} = \text{constant} \tag{4.5}$$

where V_{DSi} is the effective drain source bias across the intrinsic device.

To calculate these four elements, the DC model is evaluated at the specified bias value at which the equivalent circuit is required (V_{DSA},V_{GSA}). Solutions are then found at two other bias values, (V_{DSA}, V_{GSA} $-\partial VGS$) and ($V_{DSA}-\partial V_{DS}$, V_{GSA}), where ∂V_{DS} and ∂V_{GS} are small increments in applied drain-source and gate source voltage respectively.

At the first bias point (V_{DSA},V_{GSA}) the intrinsic gate-source voltage is given by (Figure 4.2):

$$V_{GSi} = V_{GSA} - I_{DSA}R_s \qquad (4.6a)$$

where I_{DSA} is the drain-source current calculated for this bias point. At the second bias point (V_{DSA}, V_{GSA} $-\partial V_{GS}$), the intrinsic gate-source voltage is given by:

$$V_{GSi} = V_{GSA} - \partial V_{GS} - I_{DSB}R_s \qquad (4.6b)$$

where IDSB is the drain-source current at this second bias point. Combining these two equations, the incremental change in intrinsic gate-source voltage can be found as:

$$\partial V_{GSi} = \partial V_{GS} - R_s\,(I_{DSA}-I_{DSB}) \qquad (4.7)$$

The incremental change in intrinsic drain-source voltage is found using the first (V_{DSA},V_{GSA}) and the third ($V_{DSA}-\partial V_{DS}$,V_{GSA}) bias points. The intrinsic drain source voltage at the first bias point is given by (Figure 4.2):

$$V_{DSi} = V_{DSA} - I_{DSA}\,(Rs + R_{DA}) \qquad (4.8a)$$

where R_{DA} is the drain resistance at this bias point. At the other bias point, V_{DSi} is given by:

$$V_{DSi} = V_{DSA} - \partial V_{DS} - I_{DSC} (R_s + R_{DC}) \qquad (4.8b)$$

where R_{DC} and I_{DSC} are the drain resistance and drain current at the third bias point respectively. From these two equations ∂V_{DSi} is given by:

$$\partial V_{DSI} = \partial V_{DS} - I_{DSA} (R_s + R_{DA}) + I_{DSC} (R_s + R_{DC}) \qquad (4.9)$$

Equations 4.7 and 4.9 can be used directly in the equations 4.2 to 4.6.

Intrinsic Transconductance

The incremental change in drain current for the two bias conditions is given by $I_{DSA} - I_{DSB}$, and substituting this in equation 4.2 and rearranging gives the widely-known formulae:

$$g_{mo} = g_m (1 - R_s g_m)^{-1} \qquad (4.10a)$$

where $g_m = \partial I_{DS} / \partial V_{GS}$ (4.10b)

The transconductance time delay factor, τ, is not calculated in this work. The physical origins of this factor are not well established and therefore, when calculating HEMT RF performance, a nominal value is used.

Gate channel and gate drain capacitances

To calculate the gate-channel capacitance, C_{gc}, and gate-drain capacitance, C_{gd}, an expression for the charge change in the intrinsic device, ∂Q, between two bias values is required. (The term 'gate-channel' capacitance

is used to distinguish this element from the geometric gate-source capacitance, C_{gs}, in Figure 4.1).

To calculate the change in charge, it is helpful to visualize the electron distributions at the two bias values as shown in Figure 4.3. As in Figure 3.1, changes in 2DEG sheet concentration are illustrated as changes in spatial extent of the sheet. In the gate-source access region, no charge change is shown as the physical change here is accounted for in the sidewall capacitance, C_{gsw}, to be discussed later. In Region I, the 2DEG sheet density decreases linearly (eqn. 3.12) from the undepleted value, n_{so}, down to the constant value in regions II and III, n_{sd}. At the end of region III, the sheet density is assumed to revert abruptly back to n_{so}, as required for charge neutrality in this region.

The charge in the 2DEG at the first bias condition, A, is obtained from the 'area' under the edge of the 2DEG (Figure 4.3a):

$$Q_a = qZ_G[\, n_{sda}(L_G+L_{3a}) + L_{1a}(n_{so}-n_{sda})/2] \qquad (4.11)$$

where L_{1a} and L_{3a} are the lengths of regions I and III respectively, n_{sda} is the 2DEG sheet density in regions II and III, Z_G is the gate width and L_G is the gate length.

At the second bias point, either a change in V_{DS} or V_{GS} from the first, the total electron charge in the intrinsic device is given by (Figure 4.3b)

$$Q_b = qZ_G[n_{sdb}(L_G+L_{3a})+L_{1b}(n_{so}-n_{sdb})/2+(L_{3a}-L_{3b})(n_{so}-n_{sdb}) +(N_cY_c+n_{ss})(L_{3a}-L_{3b})] \qquad (4.12)$$

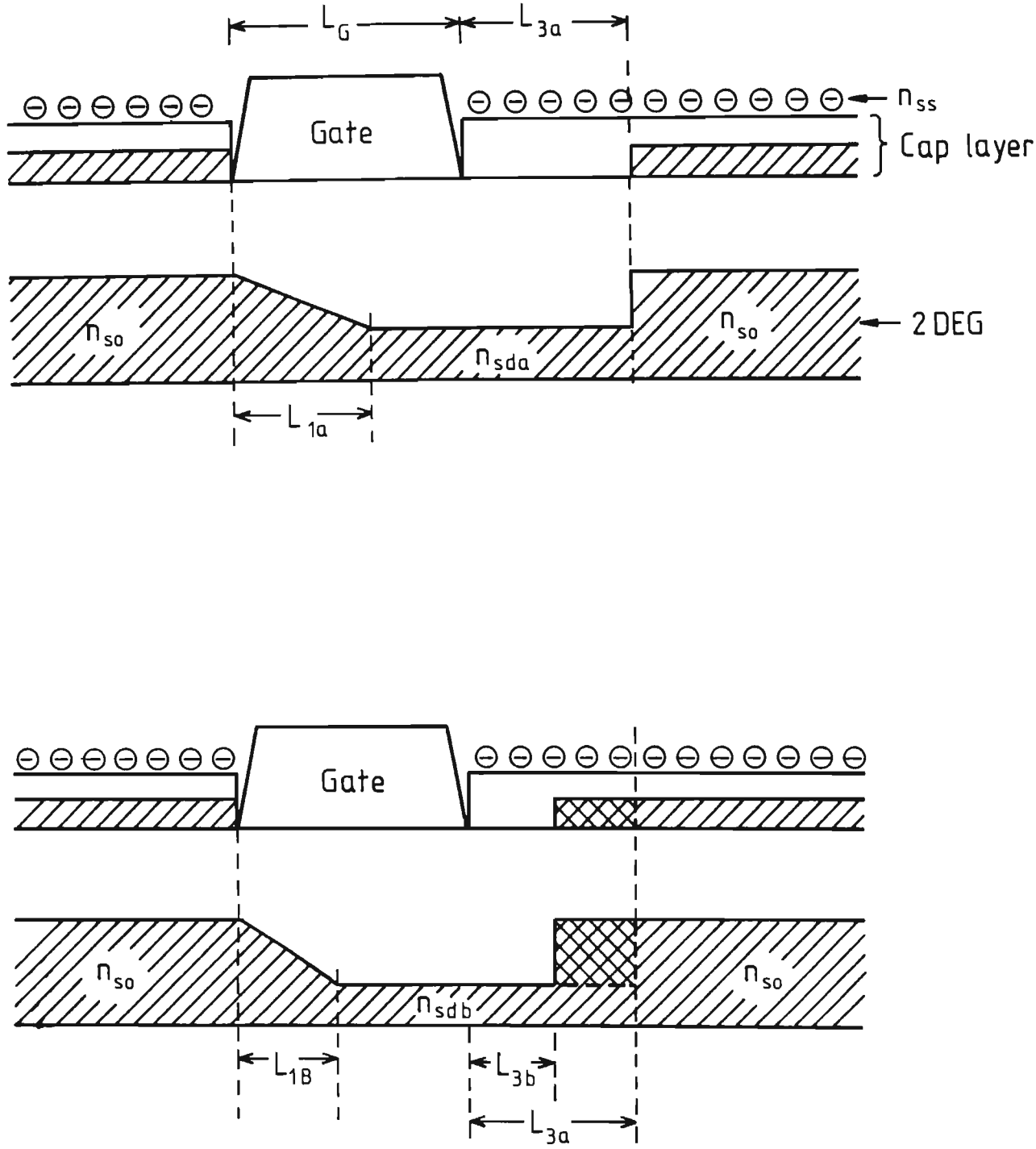


Figure 4.3. Idealized charge distributions in an IHEMT at two different bias values, used for the capacitance calculations.

Note that the same length of device (L_G+L_{3a}) is considered in both cases. The third term in equation 4.12 arises from a 2DEG charge change because of the change in L_3 (lower cross-hatched area in Figure 4.3b) and the fourth term arises from a similar charge change in the cap layer (upper cross-hatched area in Figure 4.3b).

Combining equations 4.11 and 4.12, gives the total charge change as:

$$\partial Q = qZ_G[(n_{sda}-n_{sdb})(L_G+L_{3A})+\tfrac{1}{2}L_{1a}(n_{so}-n_{sda})-\tfrac{1}{2}L_{1b}(n_{so}-n_{sdb}) -(L_{3a}-L_{3b})(n_{so}-n_{sdb}+N_cY_c+n_{ss}) \quad (4.13)$$

which can be used directly in equations 4.3 and 4.5 to calculate the capacitances.

The capacitance calculations described here are based on fundamental considerations of the change in intrinsic device electron charge arising from a change in applied bias. Hence the bias dependence of these capacitances is inherent in their calculation and no particular bias dependence needs to be assumed. In addition, no assumed partition of the total gate capacitance between C_{gc} and C_{gd} (as has been used in some models [4.11, 4.12]), is required in this work.

The charge distributions in Figure 4.3 assume the device is operating in the saturation regime. The analysis is however equally valid for the linear mode of operation, because for this mode L_3 is assumed equal to zero, region I extends under all the gate ($L_1 = L_G$) and the sheet charge density at the drain edge of the gate is found from equation 3.19.

Output resistance

The current change required for equation 4.4 is given by $I_{DSA}-I_{DSC}$ where I_{DSC} is the current calculated at the third bias point ($V_{DSA}-\partial V_{DS}$, V_{GSA}), as discussed previously. Equation 4.4 then calculates the DC output resistance. It has been found [4.13] that the output resistance in GaAs MESFETs at microwave frequencies is typically one third to one quarter of the resistance measured at DC. This has been attributed [4.13] to changes in the occupancy of surface states in the gate-drain region. Since the frequency and bias dependence of n_{ss} is not well established, it was decided in this work to estimate the RF output resistance as one third of the calculated DC value.

Output capacitance

The output capacitance, C_o, arises from coupling between the source and drain conducting regions (the areas with undepleted 2DEG in Figure 4.3), mainly via the substrate [4.14]. The effective spacing between these two regions is $L_G + L_3$, where L_3 is calculated at the required bias. In this work, the output capacitance is calculated using a modified form of the expression derived by Pucel et al. ([4.6], equation 34):

$$C = Z_G (\varepsilon_r+1) \varepsilon_o K(\sqrt{1-k^2})/K(k) \qquad (4.14a)$$

$$\text{where } k = \sqrt{(2L + L_{SD})L_{SD}}/(L+L_{SD}) \qquad (4.14b)$$

$$\text{and} \quad L_{SD} = L_G + L_3$$

The function K( ) is the complete elliptic integral of the first kind. ε_r is the relative dielectric constant of GaAs, ε_o is the permittivity of free space and Z_G is the gate width. These expressions are good approximations [4.6], when L, the length of the source and drain regions, is greater than the spacing L_{SD}.

4.2.3 Gate resistance and parasitic reactances

The remaining elements of the equivalent circuit (Figure 4.1) are not calculated using the DC model because they are dependent on structural details (for example, bonding pad size) which are not specific to the IHEMT itself.

Gate resistance

The effective RF gate resistance, R_g, is approximately one third of the DC gate resistance because of the distributed network formed by R_g with the effective gate-channel capacitance [4.9]. The DC gate resistance is dependent on the gate profile, number of gate fingers and the gate metal elements. Hence it is assumed in this work that the DC resistance is known from measurements on similar gate structures to that used in the device to be simulated.

Parasitic inductances

The parasitic source, gate and drain inductances (L_S,L_G and L_d) are not strictly part of the IHEMT itself. They are included in the equivalent circuit however to facilitate comparisons between measured and predicted RF performance. In this work, in common with most other models, the values of these inductances are obtained from measured S-parameters [4.4, 4.7].

Parasitic capacitances

The drain-source (C_{DS}), gate-source (C_{GS}) and gate-drain (C_{GD}) parasitic capacitances arise because of coupling, via the substrate and air, between the respective bonding pads and metal electrodes forming the device terminals. Each of these capacitances is therefore dependent on the layout

(plan structure) of the device. In this work, the interelectrode capacitances are calculated, for each particular device layout, using the expressions obtained by Pucel et al. [4.6] and those given by Barna [4.15].

The modulation of the cap layer in the source-gate region by the gate-source potential forms a capacitance, which here is termed the gate-sidewall capacitance, C_{gsw}. This capacitance, which is calculated using the depletion approximation, may be large because of the high cap layer dopant density and the relatively low (compared with gate-drain) gate-source voltage, typical in IHEMTs. A further component of this capacitance is the capacitance between the gate electrode and the part of the 2DEG in the gate-source region. This fringing capacitance can be calculated using a similar expression to that derived by Hida et al. [4.16] for the CHEMT.

$$C_F = \varepsilon Z_G \log (2L_{SG}/d) \qquad (4.15)$$

where ε is the dielectric constant of GaAs, L_{SG} is the length of the source gate access region and d is the gate to 2DEG spacing (equation 3.6).

4.3 S-PARAMETERS AND DEVICE GAINS

In this section, the calculation of Y-and S-parameters from the reduced equivalent circuit discussed. These parameters are used to calculate the maximum unilateral gain (MUG), the maximum available gain (MAG) and maximum tunable gain (MTG), which are commonly-used figures of merit for device RF performance.

4.3.1 Simplified equivalent circuit

The simplified equivalent circuit of figure 4.2 is obtained from that of Figure 4.1 by combining C_{GD} with C_{gd}, C_{GS} and C_{gsw} with C_{gc},and C_{DS} with C_o. This simplified equivalent circuit has been widely used [4.5-4.8, 4.17] for both FET and HEMT modelling. In this work, it is used to facilitate comparisons between predicted and measured equivalent circuit elements. The equivalent circuit elements can be found from S-parameters measured over a wide frequency range using an automatic fitting routine [4.17]. Vaitkus [4.3] has shown that the larger the number of elements in an equivalent circuit, the greater is the uncertainty in each element value derived from measured S-parameters. He has further shown however, that the fewer the number of elements, the more obscure are the connections between each element and the elements of the more complete model. Hence Figure 4.2 is a compromise between these two trends.

4.3.2 S- and Y- parameters

The two-part Y-parameter matrix is simply obtained from either equivalent circuit using elementary nodal analysis techniques. S-parameters are more commonly used for microwave measurements and circuit design, and are obtained using the standard equations [4.18]:

$$S_{11} = [(1-Y_{11})\ (1+Y_{22}) + Y_{12}\ Y_{21}]/D \qquad (4.16a)$$

$$S12 = -2Y_{12}/D \qquad (4.16b)$$

$$S_{21} = 2Y_{21}/D \qquad (4.16c)$$

$$S_{22} = [(1+Y_{11})(1-Y_{22}) + Y_{12}Y_{12}]/D \qquad (4.16d)$$

$$D = (1+Y_{11})(1+Y_{22}) - Y_{12}Y_{21} \qquad (4.16e)$$

In general, all of the terms in the S-and Y-matrices are complex quantities. These equations allow the comparison between measured and predicted HEMT characteristics to be made at the S-parameter level.

4.3.3. MUG and MAG

Two commonly used figures of merit for microwave transistors are the maximum unilateral gain (MUG) and the maximum available gain (MAG). Both are readily obtained from the S-parameters. The maximum unilateral gain (in dB) is given by [4.19]:

$$MUG\ (dB) = 10 \log [\,|S_{21}|^2\ /A] \qquad (4.17a)$$

where $A = (1-|S_{11}|^2)\ (1-|S_{22}|^2)$ (4.17b)

The maximum available gain (in dB) is given by [4.19]:

$$MAG\ (dB) = 10 \log\ S_{21}/S_{12})\ (k \pm\sqrt{k^2-1}) \qquad (4.18a)$$

where $K = (1+|D|^2 - |S_{11}|^2 - |S_{22}|^2\ )/(2|S_{12}S_{21}|)$ (4.18b)

$$D = S_{11}S_{22} - S_{12}S_{21} \qquad (4.18c)$$

and $B = 1 + |S_{11}|^2 - |S_{22}|^2 - |D|^2$ (4.18d)

The minus sign is used in (4.18a) when B is greater than zero, otherwise the positive sign is used. Equation 4.18 is only valid when the device is unconditionally stable ($k < 1$). When k is less than unity, the maximum available gain is not defined and the maximum gain is the maximum stable gain (MSG), which is given by [4.19] :

$$\text{MSG (db)} = 10 \log \left| S_{21}/S_{12} \right| \qquad (4.19)$$

4.3.4 Maximum Tunable Gain

In realising practical amplifiers, it is usually necessary to tune the amplifier to obtain the required performance because of tolerances in both active and passive devices. When the influence of reverse transfer gain is not negligible, changes in load impedance seen by the active device (from tuning or drift) cause changes in the input impedance and vice versa. This undesirable effect, calculated as the tunability factor, is primarily a practical consideration of the maximum gain which can be realised. When a device is potentially unstable, the maximum gain is theoretically equal to infinity, but the realisable gain is limited by considerations of stability and tunability.

The tunability factor is given by [4.20] :

$$T = \left| (\partial Y_{IN}/Y_{IN})/(\partial Y_L/Y_L) \right| \qquad (4.20)$$

where Y_{IN} and Y_L are the input and load admittances respectively. Equation 4.20 is clearly a measure of the relative dependence of the input circuit match on changes in the output circuit. In practical amplifiers, a tunability factor of less than 0.3 is advisable [4.20].

The maximum tunable gain is that gain calculated for the load impedance which yields the maximum gain when the tunability factor is 0.3. The MTG is always less than or equal to the MAG and provides a more realistic measure of the realisable circuit gain obtainable from an active device. In this work, the maximum tunable gain is calculated using an algorithm and equations derived from equations in reference 4.20.

4.3.5 Cut-off frequencies

Two other figures of merit often quoted when comparing RF transistor performance are the current-gain cut off frequency, f_T, and the unilateral power-gain cut off frequency, f_{max}. These frequencies can be obtained using the element values from the equivalent circuit of Figure 4.2 as [4.12]:

$$f_T = g_{mo}/(2\pi C_{gs}) \qquad (4.21)$$

$$f_{max} = f_T/\{2\sqrt{(R_i+R_g+R_s)/R_o+2\pi f_T R_g C_{gd}}\} \qquad (4.22)$$

4.4 RF SIMULATION RESULTS

In this section, the results of simulations using the RF model are presented, and compared with experimental data where applicable. The model has been included in the software program with the DC model to facilitate interactive device RF simulation.

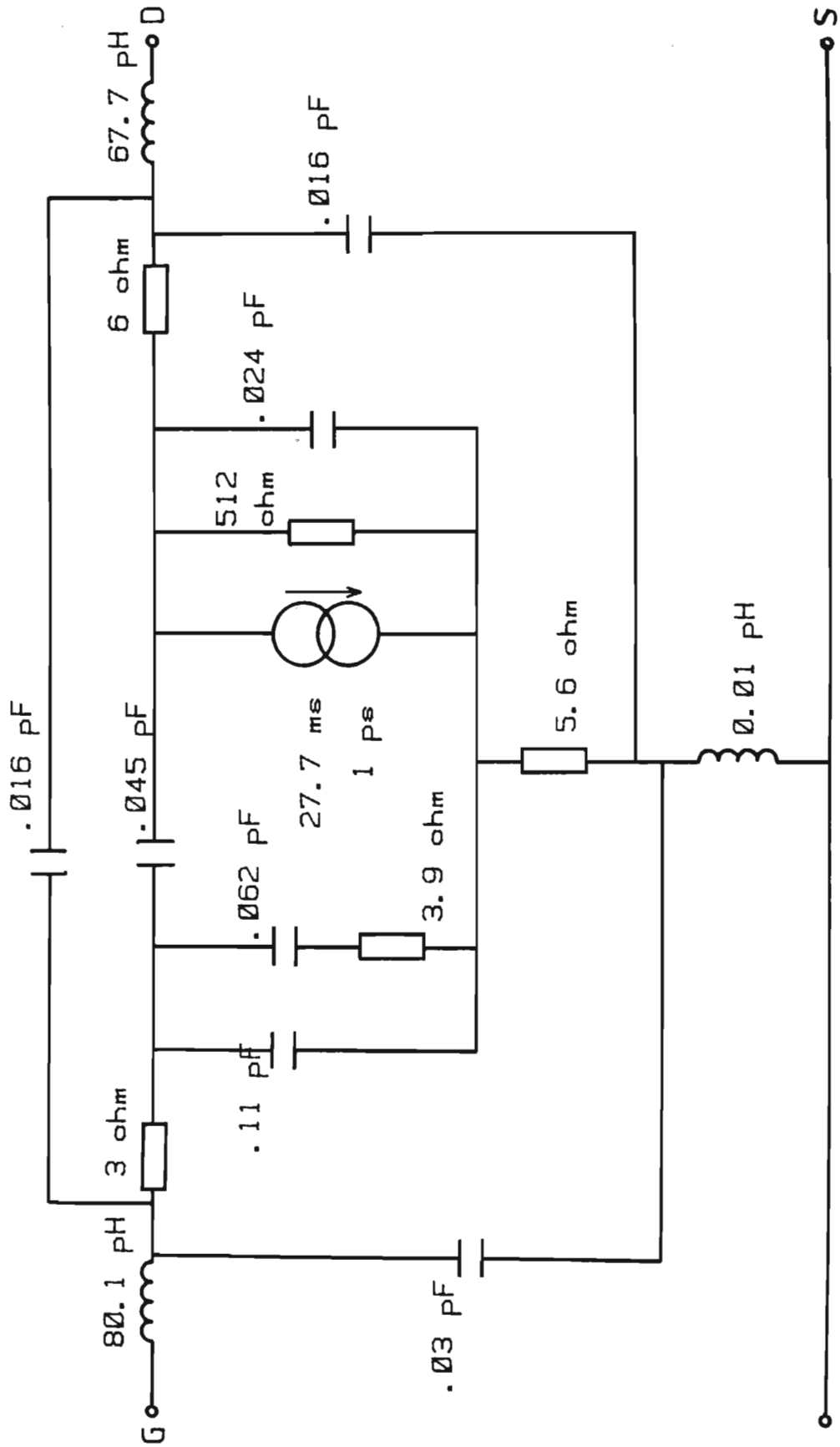
4.4.1 IHEMT S-parameters

In this subsection, a comparison between measured and simulated IHEMT S-parameters is discussed. The device used for this comparison (IHEMT number 5453) is the same one as was used for the comparison between simulated and measured DC I_{DS}-V_{DS} characteristics in Section 3.4.2, and has a gate width of 120 μm.

The S-parameters of the IHEMT were measured from 0.1 to 20 GHz on an HP 8510 connected to a Cascade prober. The applied DC drain-source and gate-source voltages were 2.0 V and +0.5V respectively.

The equivalent circuit elements calculated from the DC model were obtained using exactly the same parameters as used in the DC comparison (Table 3.1). The complete equivalent circuit is shown in Figure 4.4. With the exceptions of the transconductance time delay, gate resistance and the parasitic inductances, the elements were all obtained using the procedures discussed in Section 4.2. The inductance values were obtained using an automatic circuit fitting routine [4.17] and a value of 1 ps was assumed for the transconductance time delay. The gate resistance was obtained from DC measurement.

Figure 4.4. Calculated equivalent circuit for IHEMT number 5453



Using a reduced equivalent circuit (Figure 4.2), the S-parameters from 0.1 to 20 GHz were computed. These are compared with the measured S-parameters in the graphs of Figure 4.5. The agreement is good, particularly since some of the parasitic elements can be only approximately determined. The discrepancy between measured and simulated S_{11} at low frequencies is typical of a device with gate leakage current. Subsequent DC measurements confirmed a low (less than 0.5 mA) gate current in this particular device. The effect of this current was simulated (not shown) by placing a large resistance across C_{gs}, which reduced the difference between the measured and simulated values of S_{11}.

The predicted MUG, MAG and MTG for this device from 5 to 30 GHz is shown in Figure 4.6. The characteristic change in slope from MSG (conditional stability) at low frequencies to MAG (unconditional stability) at higher frequencies is clearly evident. The MTG is between 2 and 3 dB lower than the MAG/MSG curve across the frequency range and lacks the slope change of the latter curve.

The coplanar pad geometry of this device (photograph in Chapter 2) introduces significant parasitic capacitances, which degrade its RF performance. In order to determine the possible performance of this type of IHEMT in an MMIC, the simulation was repeated with the pad-pad capacitances removed and the inductance values halved. In addition, the gate-cap layer capacitance (part of C_{gsw}) was removed to simulate a device with a thin cap layer, which should have higher RF gain. The magnitude of S_{21}, MUG, MAG and MTG, predicted up to 50 GHz, is plotted in Figure 4.7. Useful circuit gain up to 50 GHz, may be achievable with such a device. The predicted f_T

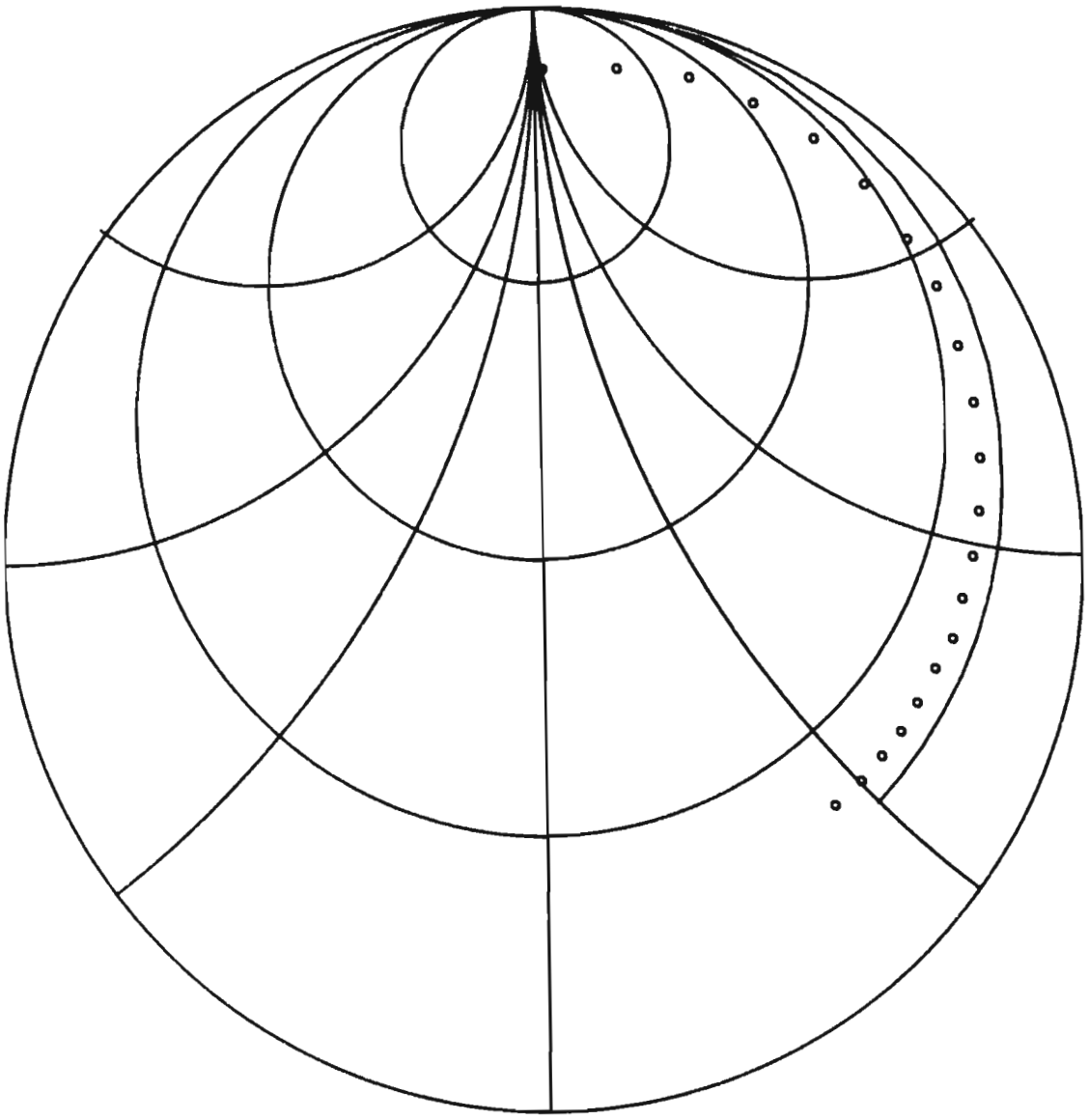


Figure 4.5a. Simulated S_{11} from 0.1 to 20 GHz. Experimental values are shown as circles.

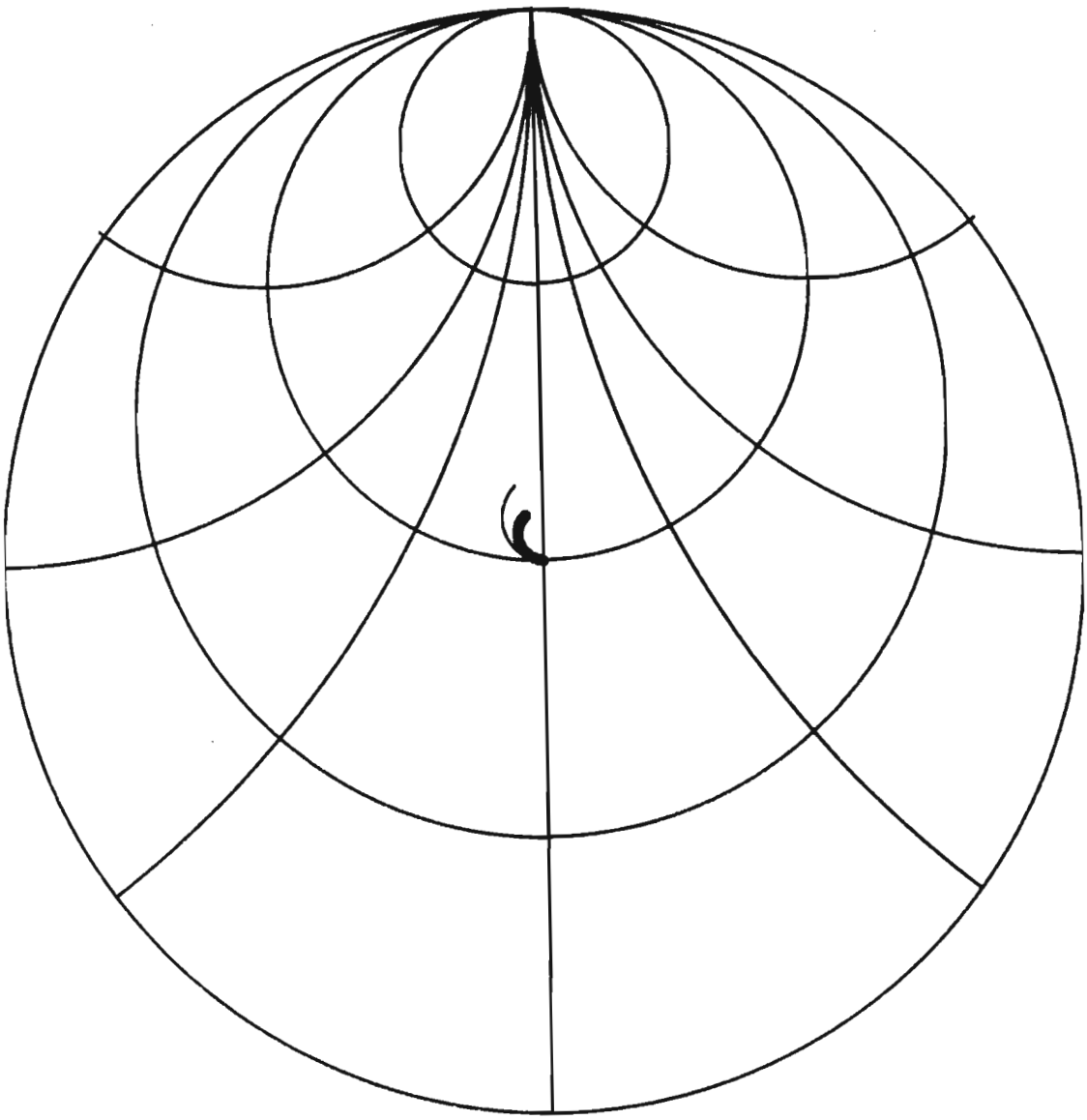


Figure 4.5b. Simulated S_{12} from 0.1 to 20 GHz. Experimental values are shown as circles.

Figure 4.5c. Simulated S_{21} magnitudes from 0.1 to 20 GHz. Experimental values are plotted as circles.

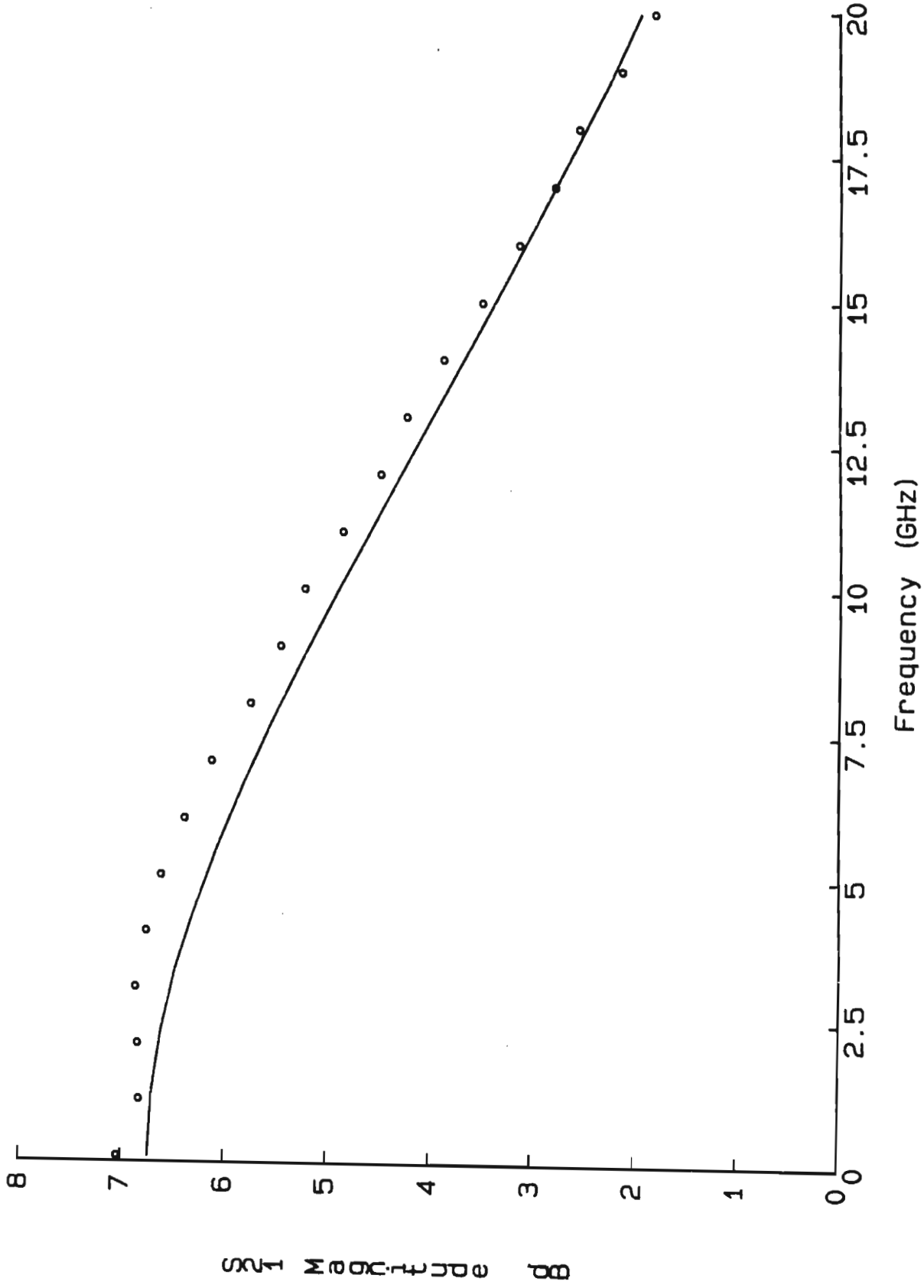
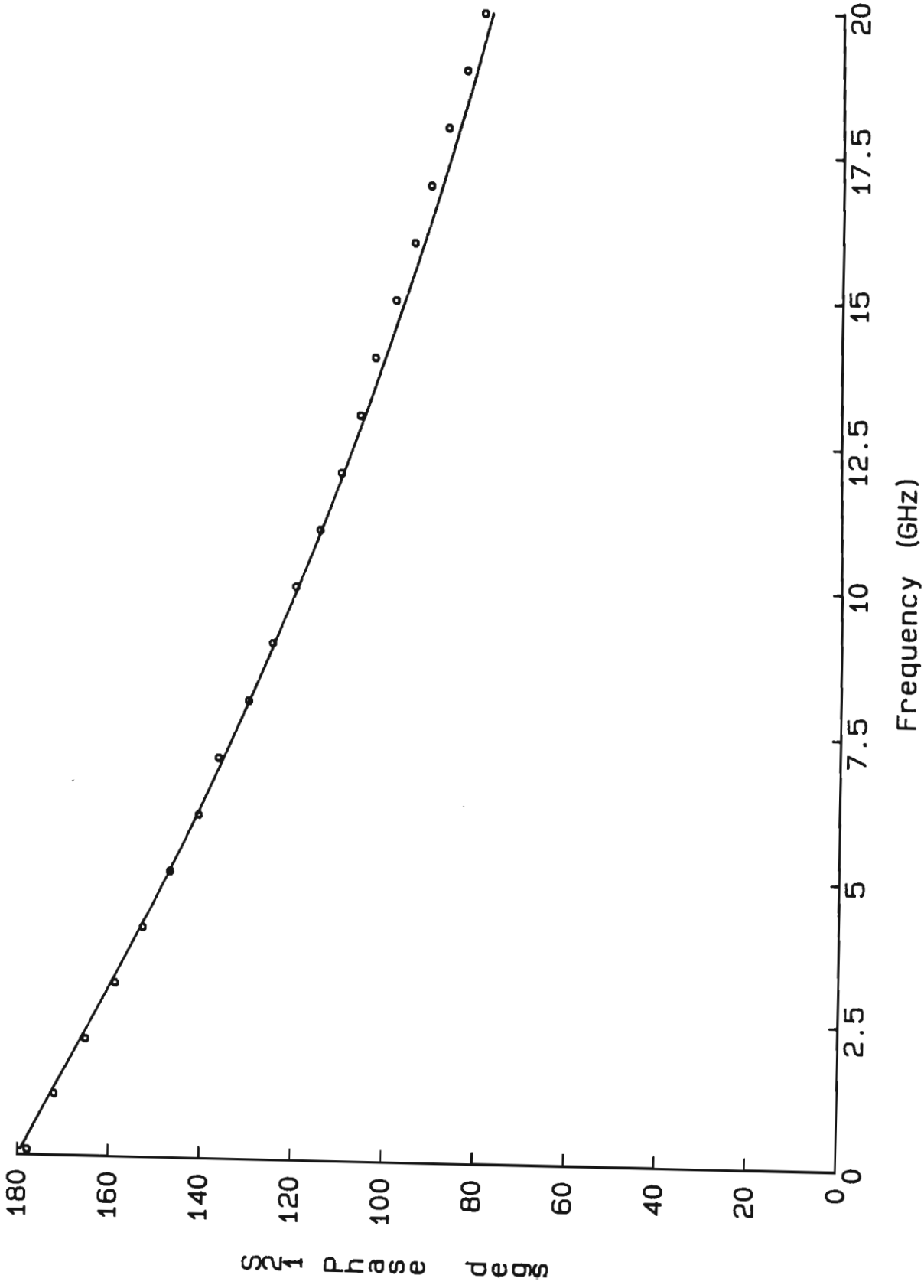


Figure 4.5d. Simulated S_{21} phase from 0.5 to 20 Ghz. Experimental values are plotted as circles.



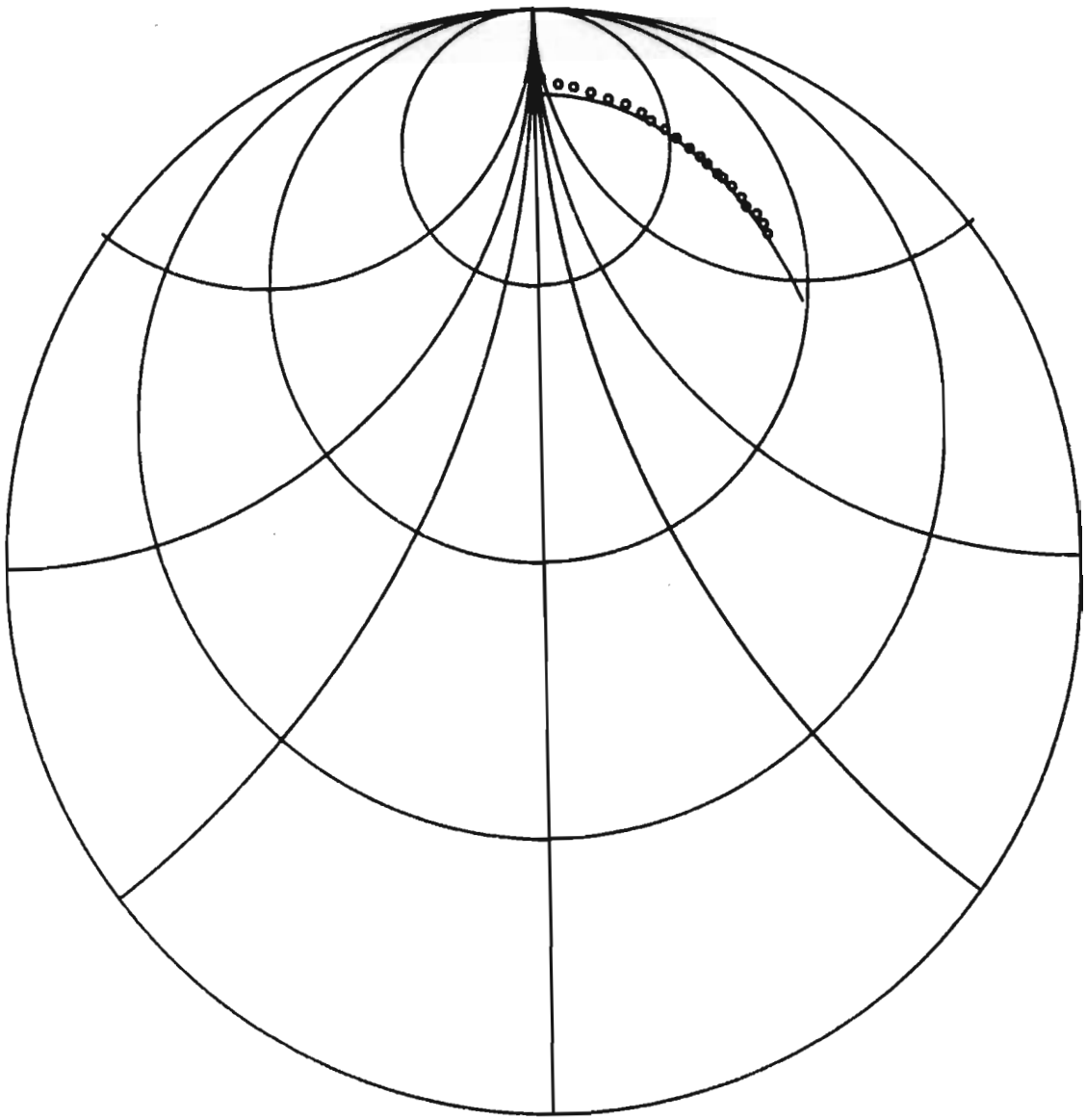


Figure 4.5e. Simulated S_{22} from 0.5 to 20 GHz. Experimental values are plotted as circles.

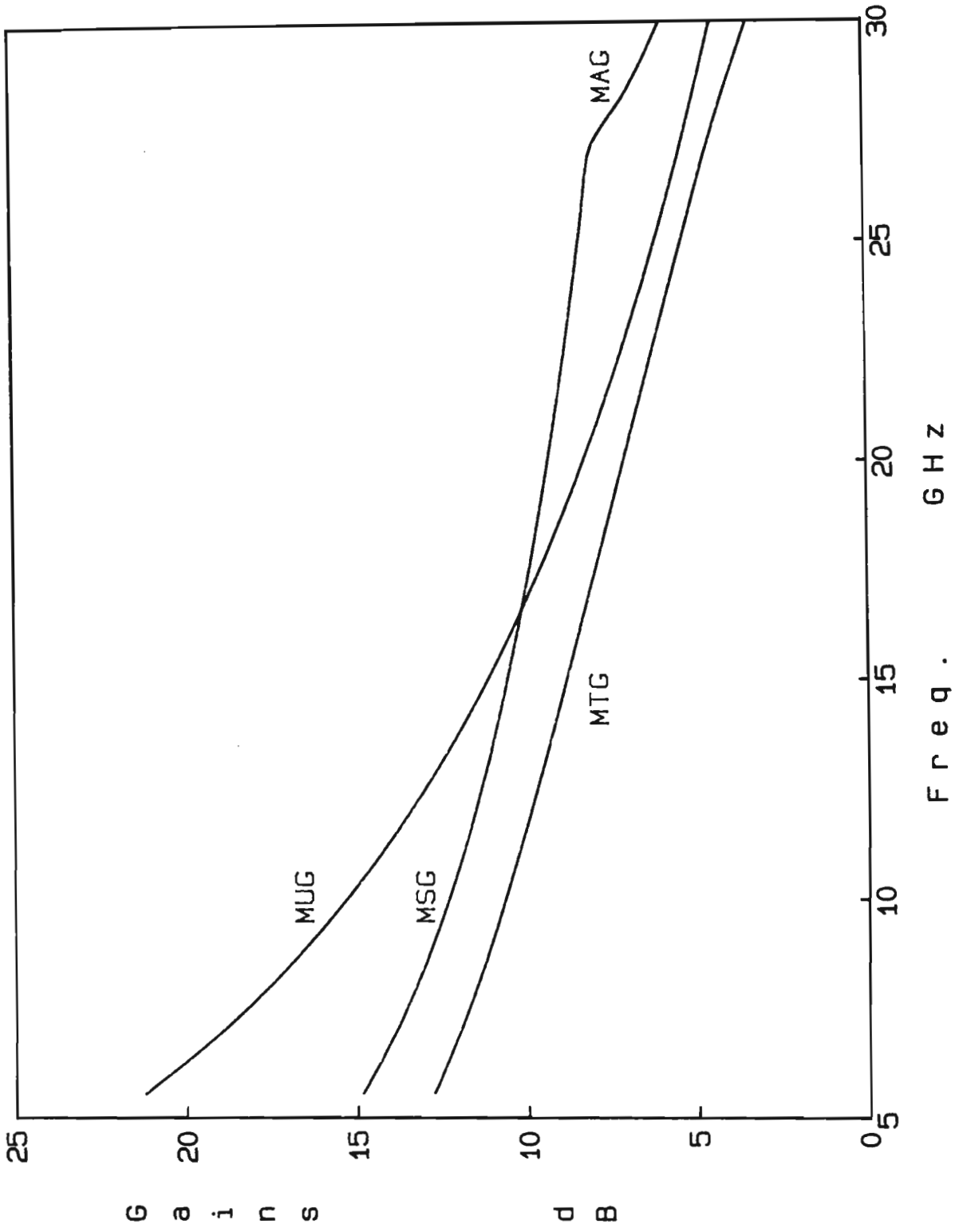


Figure 4.6. Predicted MUG, MAG/MSG and MTG for device 5453, from 0.5 to 30 GHz.

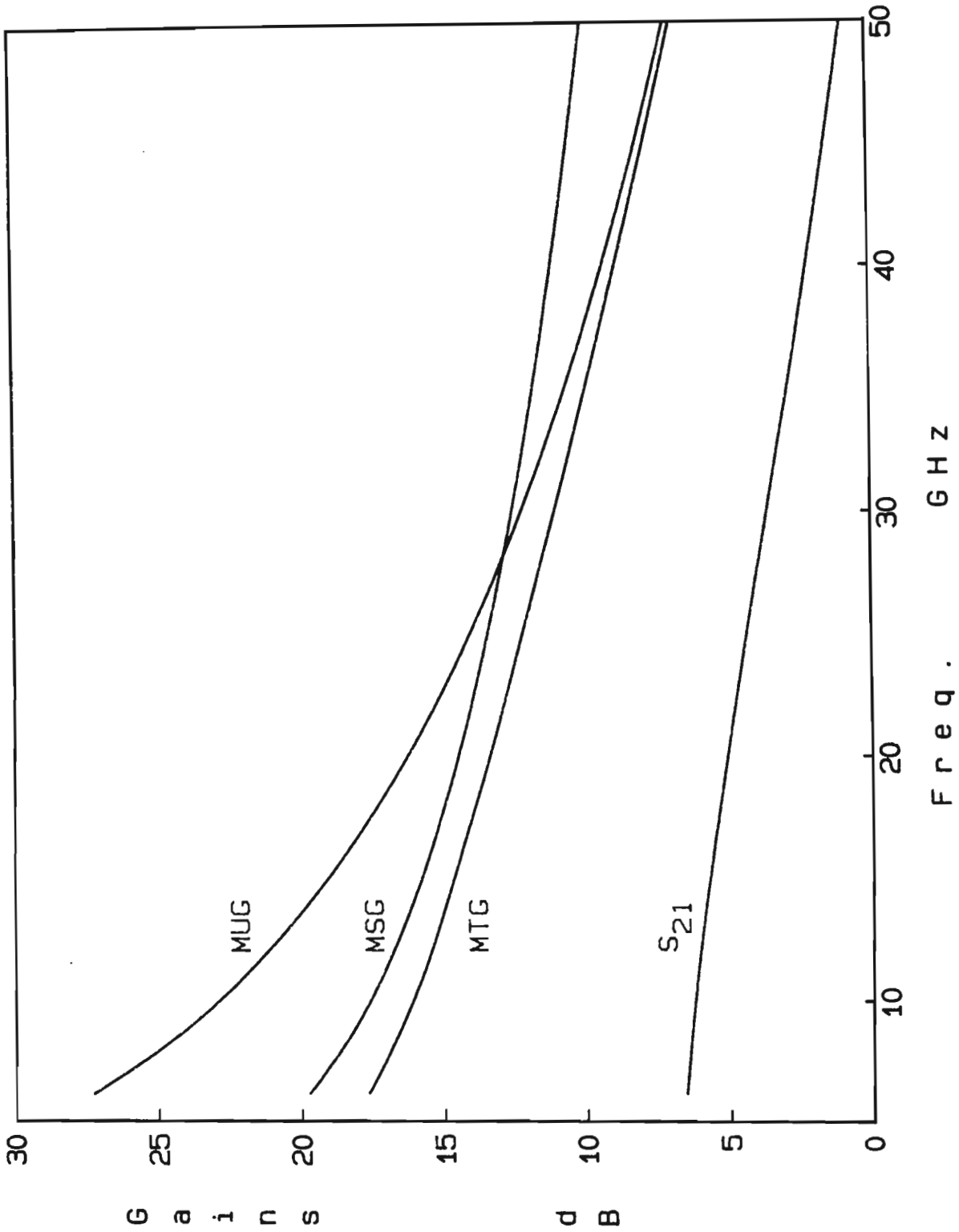


Figure 4.7. Predicted MUG, MAG and MTG for device 5453 with reduced parasitics, from 0.5 to 50 GHz. The magnitude of S_{21} is also plotted.

of this device is 42 GHz, compared with 22 GHz for the device with all the parasitics included.

4.4.2 Bias dependence of elements

The correct prediction of the bias dependence of RF equivalent circuit elements of HEMTs is essential for the development of a large signal model for these devices. Experimental results of the bias dependence of CHEMT elements have been published [4.7, 4.21] but no large signal model has appeared to date [4.7].

In this subsection a comparison between the simulated and experimentally determined bias dependence of four key elements in the equivalent circuit (C_{gs}, C_{gd}, g_{mo} and R_o) is given.

The experimental element values cannot be measured directly but must be obtained from fits to the measured S-parameters. Since no unique equivalent circuit can be obtained from a given set of S-parameters [4.3], the resulting comparison is strictly a measure of the accuracy of both the IHEMT model and the circuit fitting routine (assuming no errors in the measured parameters). However, the qualitative trends will be well represented. In this work, the automatic fitting routine developed at GEC [4.17] for the MESFET, which uses the same equivalent circuit as the model (Figure 4.2), was used.

The IHEMT used for this comparison is the same device as used in the other comparisons (device number 5453). The S-parameters were measured over the frequency range 0.1 to 20 GHz and at 9 different bias values (V_{DS} at IV, 2V and 3V; V_{GS} at +0.1V, +0.3V, +0.5V). The element values were

predicted from the model for the same gate bias values, but with drain bias values from 0.5V to 3V in 0.5V steps.

Figure 4.8 shows the predicted and fitted values for gate-drain capacitance, C_{gd}. The qualitative trends are the same in both sets of curves but the predicted curves are offset by around 6 fF, which can be attributed to an over-estimation of the parasitic gate-drain capacitance in the model. This effect is reflected in the difference between the predicted and measured S_{12} curves in Figure 4.4b. The coincident values on the fitted curves at V_{DS} = 2V for the 0.5V and 0.3V V_{GS} curves are an artifact of the fitting routine which rounds the calculated C_{gd} values to the nearest femtofarad. At drain bias values well beyond the knee voltage, C_{gd} decreases with increasing gate voltage, whereas at lower drain bias the trend is reversed. This crossover in both fitted and modelled curves is similar to that measured by Arnold et al. for the CHEMT [4.7, 4.21]. More measured values for the IHEMT would be required to determine the drain bias at which crossover occurs, but it is clearly near the knee voltage. The gate-drain capacitance increases rapidly with decreasing drain bias below the knee voltage ([4.7], and predicted curves curves in Figure 4.8).

The dependence of gate-source capacitance, C_{gs}, is shown in Figure 4.9. The model correctly predicts a decrease in capacitance with decreasing gate bias and an increase in capacitance with increasing drain bias. These dependences are the same as in CHEMTs [4.21]. The quantitative agreement is reasonable (maximum difference between predicted and fitted values is 13 percent of fitted value, at V_{DS} = 3V and V_{GS} = 0.5V), but the predicted dependences on both gate and drain bias are smaller than fitted.

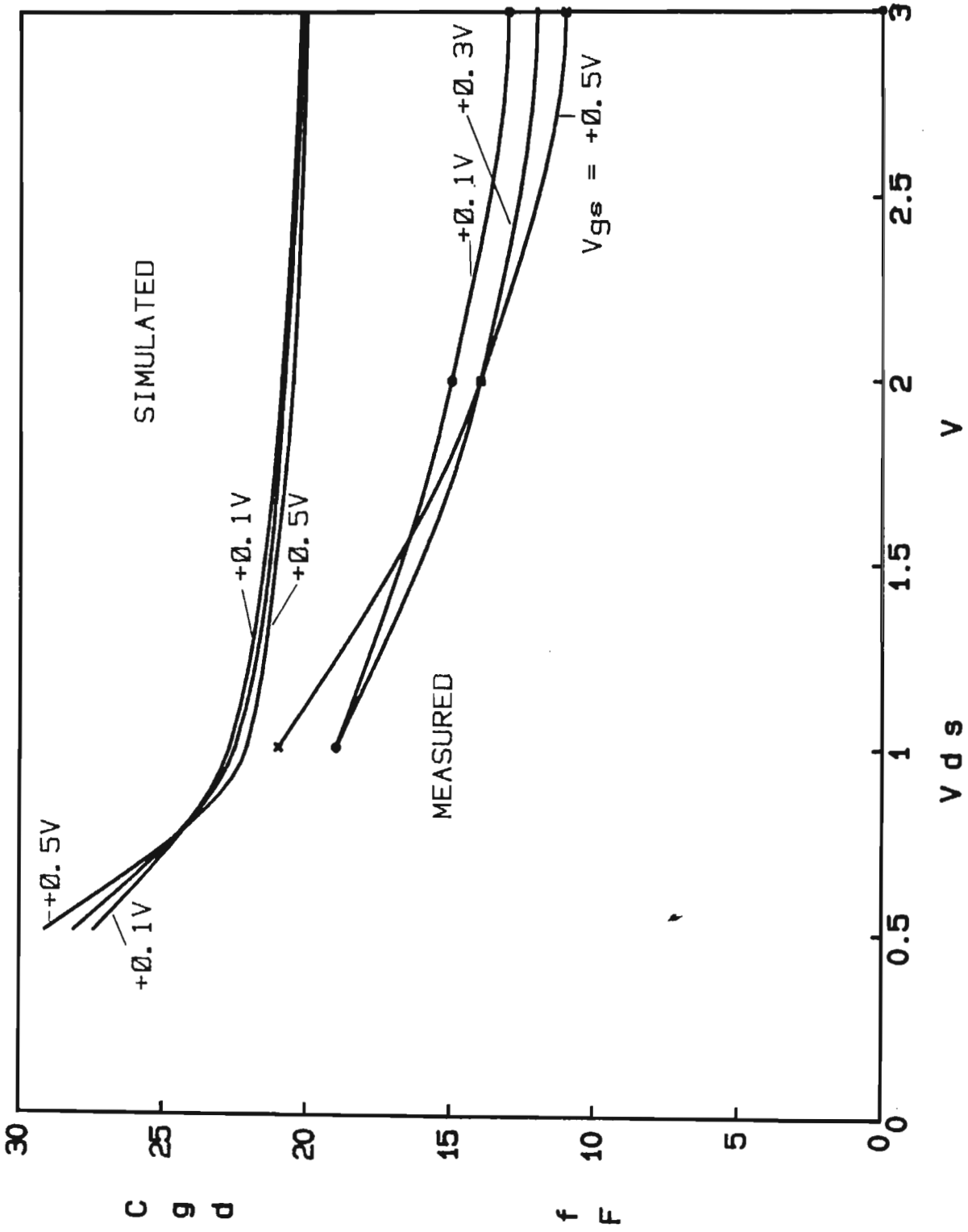


Figure 4.8. Simulated and measured bias dependence of gate-drain capacitance.

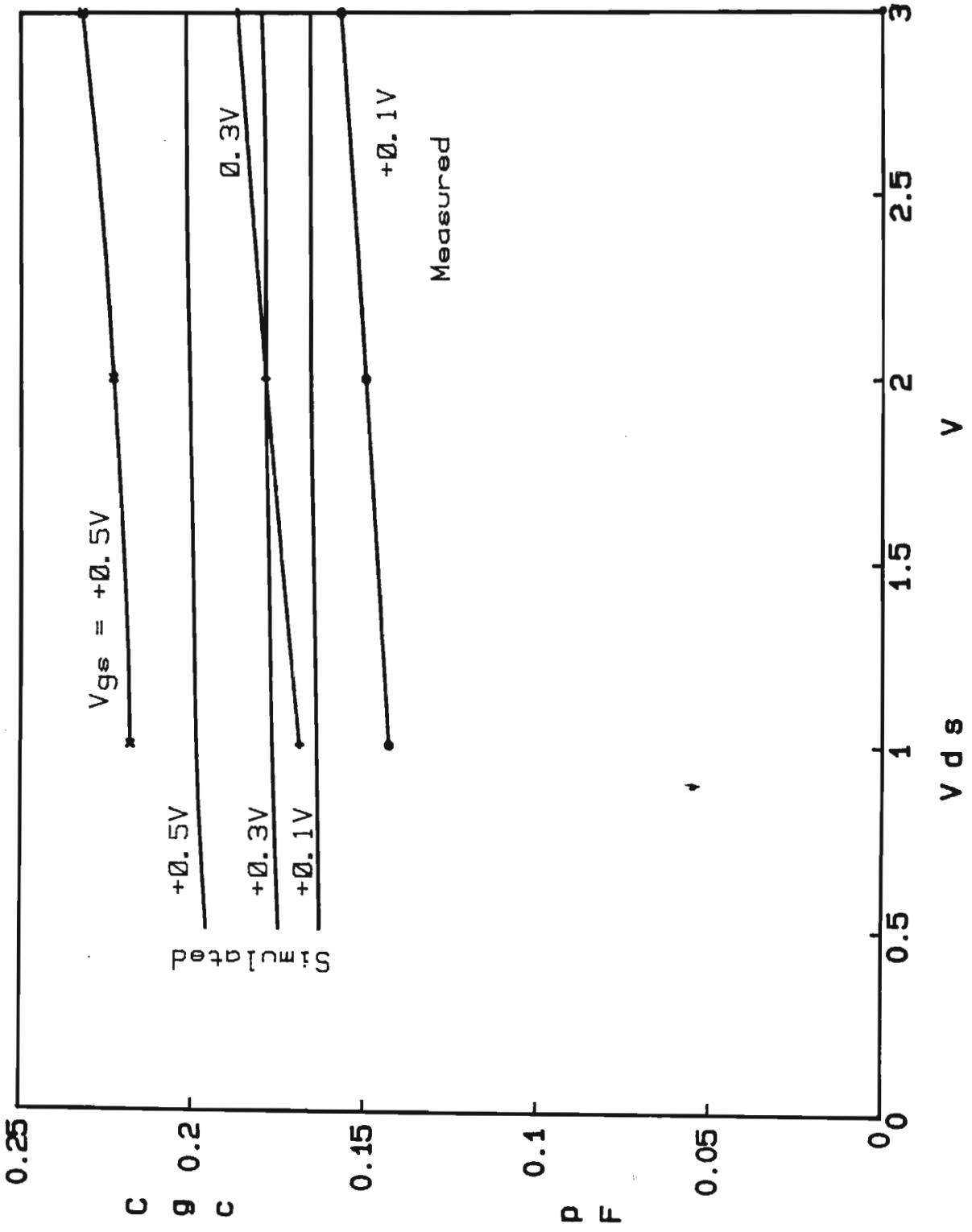


Figure 4.9. Simulated and measured bias dependence of gate-channel capacitance.

As noted by Arnold et al. [4.21], neither dependence is predicted by the often-used approximation for C_{gs}:

$$C_{gs} = \varepsilon Z L_g / d \qquad (4.23)$$

which is constant for a given device geometry (except near the threshold voltage in CHEMT where the well-widening increases the effective distance from gate to 2DEG).

The comparison between predicted and fitted intrinsic transconductance, g_{mo}, is shown in Figure 4.10. The variation in predicted g_{mo} with changing gate bias is less than the fitted variation. This underestimation of transconductance compression in the model has been discussed in Chapter 3. The fit to this circuit is strongly dependent on the accuracy of the DC I_{DS}-V_{DS} simulation (Figure 3.6), which would be improved by allowing parameters other than the recess depth to be varied as discussed in Chapter 3. This comparison does suggest however, that the DC and RF transconductances do not differ greatly in practical devices.

Figure 4.11 contains a comparison between the fitted and predicted values of the output resistance, R_o. The agreement between these curves is qualitative only; the predicted curves correctly show an increase in output resistance with either increasing drain bias or decreasing gate bias, but the slope of these curves is clearly incorrect. Arnold et al. [4.7] also measured an increasing output resistance with drain bias, but also found that beyond a particular drain bias (in their case, 3V for 1 μm gate length CHEMTs) the resistance remained almost constant.

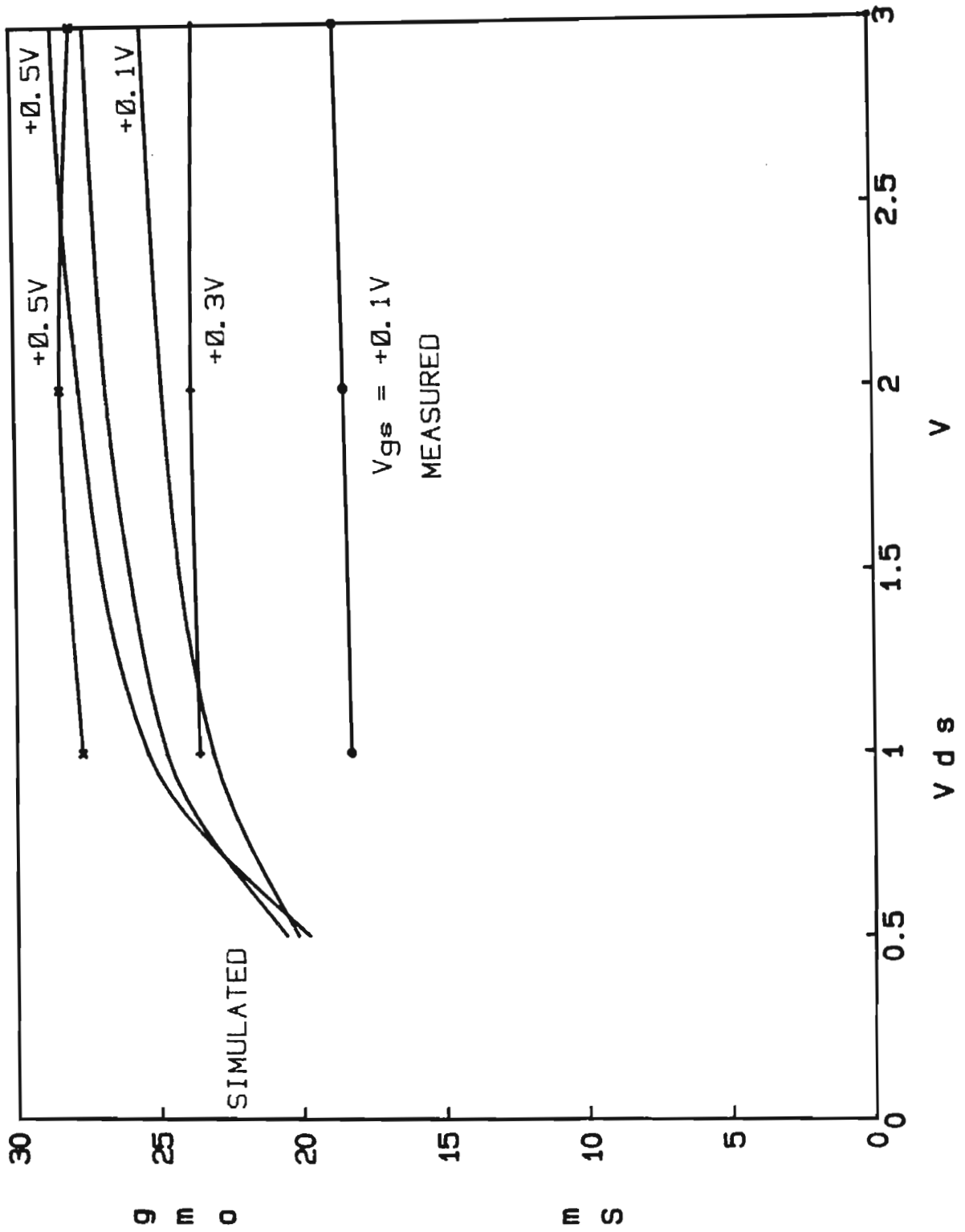


Figure 4.10. Simulated and measured bias dependence of intrinsic transconductance.

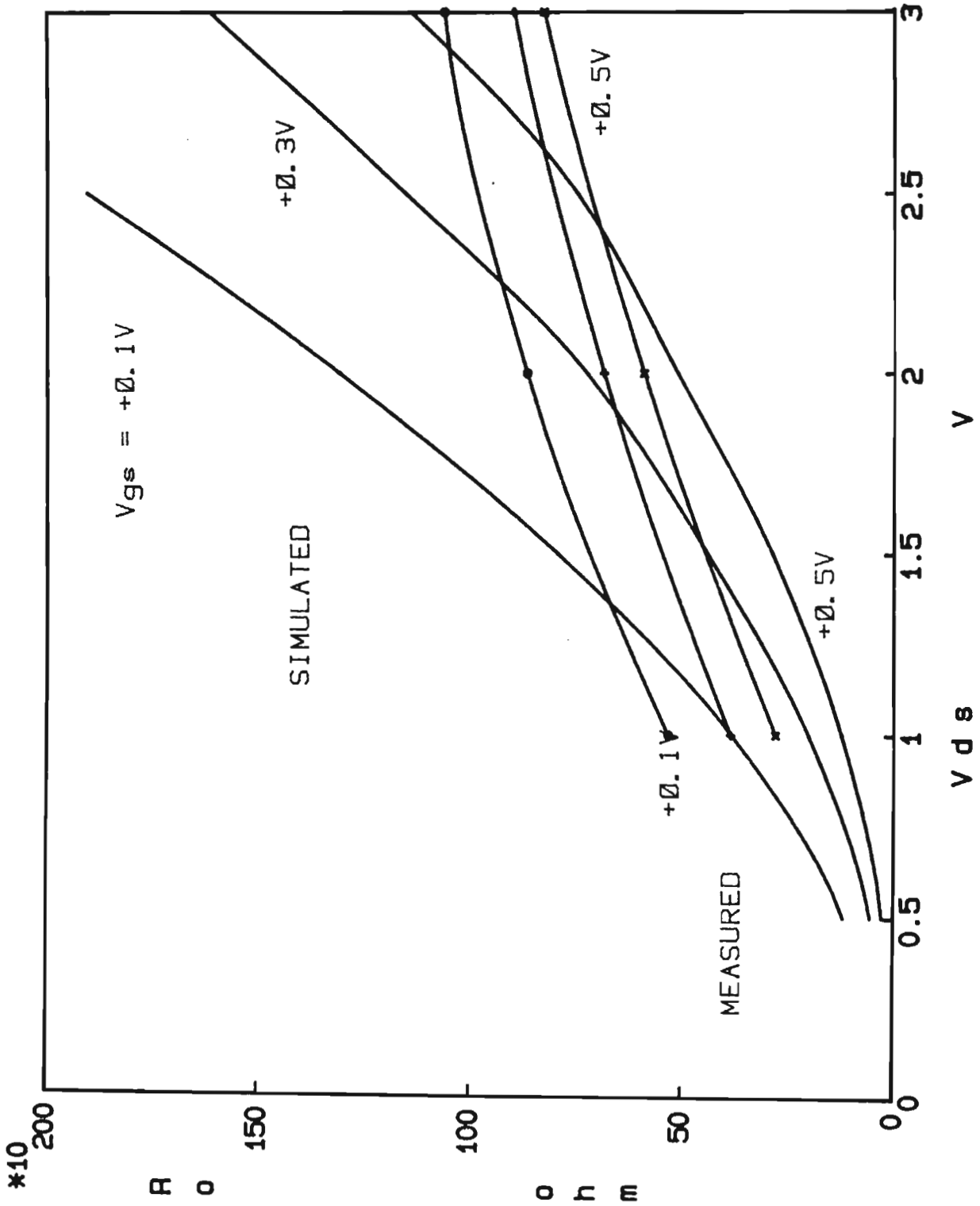


Figure 4.11. Simulated and measured bias dependence of output resistance.

One possible source of the differences between the fitted and predicted R_o values is the influence of surface charge. As was noted earlier, little quantitative information exists on the dependence of occupied surface state density in HEMTs (or FETs) on either bias or frequency. Consequently, predicted values are obtained from simulations with a fixed surface state density, which is almost certainly non-physical. Pucel [4.22] has stated that little progress has been made in solving the problem of the disparity between DC and RF output resistance in MESFETs, and this problem clearly also exists for HEMTs. Once a quantitative understanding of surface state dependences on bias and frequency is obtained, a clear analysis of any limitations in this work will be possible.

The fitted transconductance time delay factor, τ, was found to increase linearly with increasing drain bias, for a given gate voltage. The same dependence has been found by Curtice and Ettenberg [4.23] in MESFETs, who attributed this dependence to a decrease in saturated drift velocity with increasing drain bias. However, it is not clear what physical effect may give rise to such a decrease. Although τ is not calculated in this work, an alternative qualitative explanation can be obtained from the DC IHEMT model. For a given gate and surface state density, the length of region III increase as the drain bias increases. Hence the transit time for an electron travelling through the intrinsic device at the saturated drift velocity also increases. The need for a bias dependent saturated drift velocity arises from the (incorrect) assumption that the transit length is equal to the gate length (Appendix A in ref. 4.23), which is clearly not bias dependent. It is the author's intention to include a calculation of τ in the IHEMT model in the future.

4.4.3 Milimeter wave predictions

There is continuing interest in the development of GaAs MMICs with higher and higher operating frequencies. The HEMT (either CHEMT or IHEMT) is receiving widespread attention for millimeter wave applications because of the generally higher transconductances and lower noise figures of these devices, compared with those of MESFETs. The IHEMT model has been used to predict the potential circuit gain of devices with 75x0.25 μm gates, in the range 10 to 100 GHz.

As device gate length is reduced, parasitic elements become increasingly important to overall device performance. The prediction is however based on a structure which is consistent with contemporary fabrication technology. A gate to 2DEG spacing of 50 nm and an interelectrode spacing of 0.375 μm is assumed. A contact resistance of 0.1 ohm.mm is assumed; a value which has been obtained in practical IHEMTs [4.24] and CHEMTs [4.25]. It is further assumed that the device is part of an MMIC and so is not degraded by large parasitic pad capacitances, as in the 0.5 μm devices discussed in the last section.

The calculated maximum tunable gain is shown in Figure 4.12. This figure is similar to one presented at the WOCSEMMAD conference [4.26]. Also shown are several experimental points for the circuit gain achieved with 0.25 μm FETs and CHEMTs, as published in the open literature [4.27-4.36]. The agreement between the forecast and measured results is better than could have been expected.

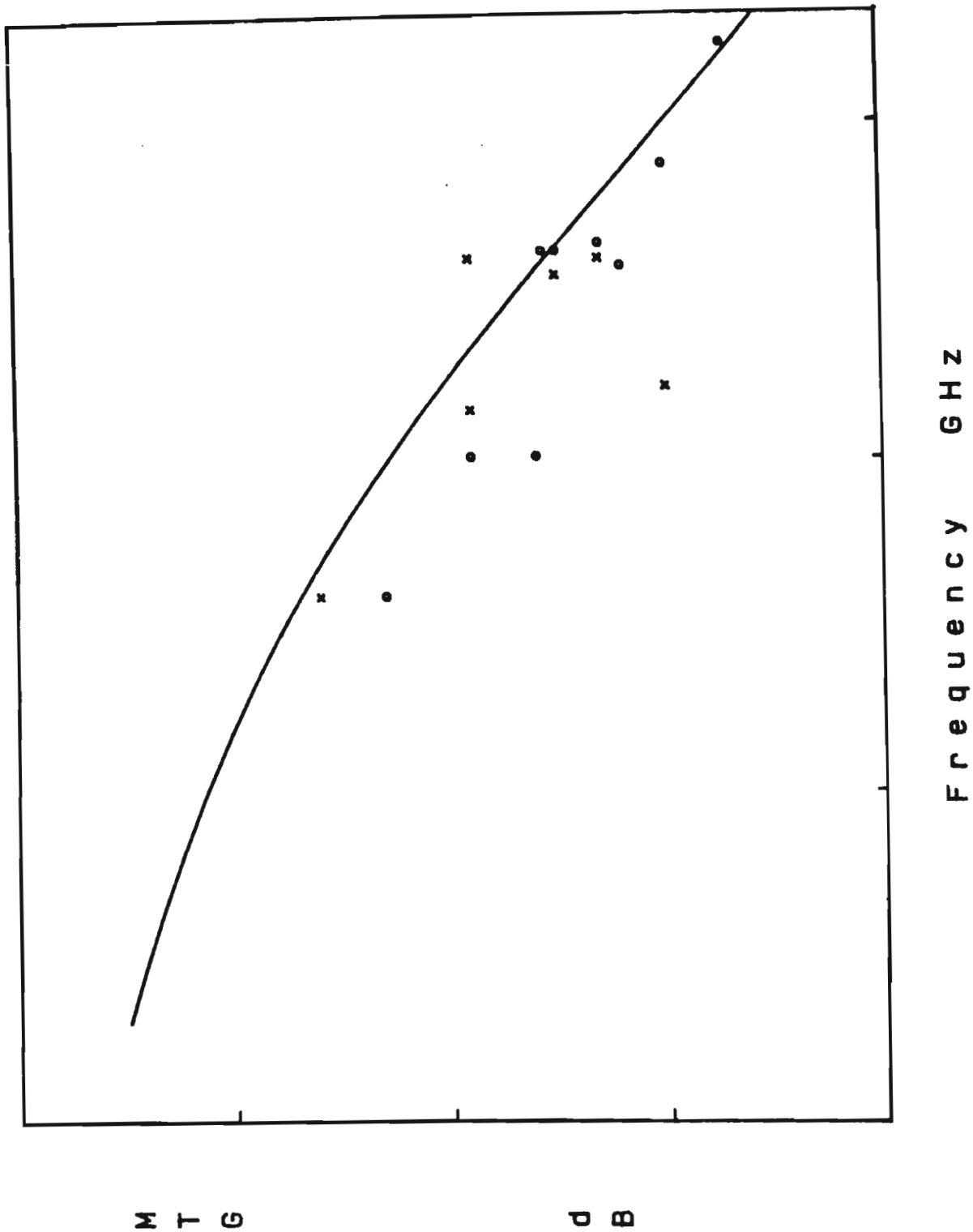


Figure 4.12. Calculated maximum tunable gain for a 0.25 micron gate length device. Experimental circuit gains from 0.25 micron FET (x) and HEMT (o) amplifiers are plotted for comparison.

A key feature of this prediction is that a saturated drift velocity of 10^7 cm/s was used, which implicitly assumes that no overshoot occurs. As the gate length is reduced, the length of region III becomes an increasingly larger proportion of the gate length (at a given drain bias) which means that the length of the transit region does not decrease as rapidly as the gate length. Consequently, overshoot effects are expected to be less important than predicted in models which assume no extension of the active region beyond the drain edge of the gate.

The good agreement obtained in Figure 4.12 cannot be taken as conclusive evidence that overshoot is negligible, for the same reasons as discussed in Chapter 3 for the DC simulations. The detailed structures of the CHEMTs yielding the experimental data in Figure 4.12 are not known to the author (the 94 GHz result [4.27] was published after the predicted curve) and hence the predictions can only be considered as an indication that the velocity is not substantially greater than 10^7 cm/s. Calculations assuming a velocity of 1.7×10^7 cm/s [4.25] yielded an expected circuit gain of 9 dB at 100 GHz, considerably greater than achieved experimentally to date.

It is interesting to note that some of the 0.25 μm FET results are comparable with the published HEMT results, and at 60 GHz the best FET result is more than 2dB higher than the HEMT result [4.28, 4.29]. While HEMT amplifier development is clearly in its infancy, these results indicate that for gain alone, HEMTs may not prove to be markedly superior to optimised FETs. However, the lower noise figure obtainable in HEMTs offers a distinct advantage over FETs.

Maximum tunable gain, although not widely used, has been found in this work to provide useful estimates of achievable circuit gains. In general, the calculated MTG is around 3 dB lower than the calculated MAG/MSG curve.

4.5 CONCLUSION

In this chapter, a new RF model for HEMTs has been discussed. The majority of the equivalent circuit elements are obtained from the DC model presented in Chapter 3. The calculation of S-parameters and amplifier gains have also been discussed, and comparisons between the simulated and measured RF performance of HEMTs have been made.

The agreement between the predicted and measured RF performance at both the S-parameter and equivalent circuit element levels has been shown to be reasonable, particularly since the only adjustable parameter used was the recess depth in the DC part of the model. The major discrepancy between measured and predicted elements was found in the output resistance, and is thought to be due to frequency dependent occupation of surface states in the gate drain region. A similar disparity exists in all FET and HEMT models and is clearly an area for future investigation.

Noise in HEMT devices has not been considered in this work but the model presented here should provide a useful starting point for noise modelling in future.

CHAPTER 4 REFERENCES

4.1 SUPER-COMPACT, Linear Circuit Analysis and Optimization, Compact Engineering, Paterson, NJ.

4.2 Heinrich W., "Limits of FET Modelling by Lumped Elements", Elec. Lett., Vol. 22, No. 12, June 1986, pp. 630-632

4.3 Vaitkus R.L., "Uncertainty in the values of GaAs MESFET equivalent circuit elements extracted from measured two-port scattering parameters", Proc. 1983 IEEE/Cornell Conf. on High Speed Semiconductor Devices and Circuits, pp. 301-308

4.4 Curtice W.R. and Camnisa R.L., "Self-Consistent GaAs FET Models for Amplifier Design and Device Diagnostics", IEEE Trans. Microwave Theory Tech., Vol. MTT-32, No. 12, Dec. 1984, pp. 1573-1578

4.5 Kondoh H., "An Accurate FET Modelling from Measured S-parameters", 1986 IEEE MTT-S Symposium Digest, pp. 377-380

4.6 Pucel R.A., Haus H.A. and Statz H., "Signal and Noise Properties of Gallium Arsenide Microwave Field-Effect Transistors", Advances in Electronics and Electron Physics, Vol. 38, New York: Academic Press, 1975, pp. 195-265

4.7 Arnold D.J., Fischer R., Kopp W.F., Henderson T.J. and Morkoc H., "Microwave Characterisation of (Al,Ga)As/GaAs Modulation-Doped FET's: Bias Dependence of Small-Signal Parameters", IEEE Trans. Elec. Dev., Vol. ED-31, No. 10, Oct. 1984, pp. 1399-1402

4.8 Hikosaka K., Hirachi Y. and Abe M., "Microwave Power Double-Heterojunction HEMT's", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 583-589

4.9 Fukui H., "Determination of the Basic Device Parameters of a GaAs MESFET", Bell Sys. Tech. Jnl., Vol. 58, No. 3, March 1979, pp. 771-797

4.10 Hower P.L. and Bechtel N.G., "Current Saturation and Small-Signal Characteristics of GaAs Field-Effect Transistors", IEEE Trans. Elec. Dev., Vol. ED-20, No. 3, Mar. 1973, pp. 213-220

4.11 Ketterson A.A. and Morkoc H., "GaAs/AlGaAs and InGaAs/AlGaAs MODFET Inverter Simulations", IEEE Trans. Elec. Dev., Vol. ED-33, No. 11, Nov. 1986, pp. 1626-1634

4.12 Yaeger H.R. and Dutton R.W., "Circuit Simulation Models for the High Electron Mobility Transistor", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 682-691

4.13 Ladbrooke P.H., "The Possible Dependence of Drain Conductance upon Surface Charge in GaAs MESFET", Internal Note, June 1985

4.14 Das M.B., "A High Aspect Ratio Design Approach to Millimeter-Wave HEMT Structures", IEEE Trans. Elec. Dev., Vol. ED-32, No. 1, Jan. 1985, pp. 11-17

4.15 Barna A., "VHSIC, Technologies and Tradeoffs", John Wiley and Sons, 1981, pp. 2-13

4.16 Hida H., Ohata K., Suzuki Y. and Toyoshima H., "A New Low-Noise AlGaAs/GaAs 2DEG FET with a Surface Undoped Layer", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 601-607

4.17 Bridge J.P. and Caddick C.J., "Automatic Circuit Fitting to Measured MESFET Scattering Parameters", submitted to IEEE Trans. Micro. Theory and Tech.

4.18 Vendelin G.D., "Design of Amplifier and Oscillators by the S-parameter Method", John Wiley and Sons, 1982

4.19 Anderson R.W., "S-parameter techniques for faster more accurate network design", Hewlett-Packard Aplication Note 95-1, from Hewlett-Packard Jnl., Vol. 18, No. 6, Feb. 1967

4.20 Abrie P.L.D., in "The Design of Impedance-Matching Networks for Radio-Frequency and Microwave Amplifiers", Artech House, 1985

4.21 Arnold D., Kopp W., Fischer R., Klem J. and Morkoc H., "Bias Dependence of Capacitances in Modulation-Doped FET's at 4 GHz", IEEE Elec. Dev. Lett., Vol. EDL-5, No. 4, April 1984, pp. 123-125

4.22 Pucel R., "MMICs, Modelling, and CAD- Where do we go from here?", Proc. 16th European Microwave Conf., Dublin, Ireland, Sept. 1986, pp. 61-70

4.23 Curtice W.R. and Ettenberg M., "A Nonlinear GaAs FET Model for Use in the Design of Output Circuits for Power Amplifiers", IEEE Trans. Micro. Theory and Tech., Vol. MTT-33, No. 12, Dec. 1985, pp. 1383-1394

4.24 Kinoshita H., Nishi S., Akiyama M. and Kaminishi K., "High-speed Low-power Ring Oscillator Using Inverted-Structure Modulation-doped GaAs/n-AlGaAs Field-Effect Transistors", Jap. J. Appl. Phys., Vol. 24, No. 8, Aug. 1985, pp. 1061-1064

4.25 Jones W.L. and Eastman L.F., "High-Performance AlGaAs/GaAs MODFET's with Improved Ohmic Contacts", IEEE Trans. Elec. Dev., Vol. ED-33, No. 5, May 1986, pp. 712-716

4.26 Hill A.J. and Ladbrooke P.H., presented at WOCSEMMAD 1986, San Fransisco, Feb. 1986

4.27 Smith P.M., Chao P.C., Duh K.H.G., Lester L.F. and Lee B.R., "94 GHz transistor amplification using an HEMT", Elec. Lett., Vol. 22, No. 15, July 1986, pp. 780-781

4.28 Sholley M. and Nichols A., "60 and 70 GHz (HEMT) Amplifiers", 1986 IEEE MTT-S Digest, pp. 463-465

4.29 Chao P.C., Smith P.M., Duh K.H.G. and Hwang J.C.M., "60 GHz GaAs Low-Noise MESFETs by Molecular Beam Epitaxy", Paper IVA-8, 44th. Ann. Dev. Res. Conf., 1986

4.30 Duh K.H.G., Chao P.C., Smith P.M., Lester L.F., Lee B.R., "60 GHz Low-noise high-electron-mobility transistors", Elec. Letts., Vol. 22, No. 12, June 1986, pp. 647-649

4.31 Berenz J., Nakano K., Hsu T-I., Goel J., "HEMT 60 Ghz Amplifier", Elec. Lett., Vol. 21, No. 22, Oct. 1985, pp. 1028

4.32 Chao P.C., Palmateer S.C., Smith P.M., Mishra U.K., Duh K.H.G. and Hwang J.C.M., "Millimeter-Wave Low-Noise High Electron Mobility Transistors", IEEE Elec. Dev. Lett., Vol. EDL-6, No. 10, Oct. 1985, pp. 531-533

4.33 Berenz J.J., "High Electron Mobility Transistors (HEMT)", MTT-S, July 1984, pp. 43-52

4.34 Kim B., Tserng H.Q. and Shih H.D., "Millimeter-Wave GaAs FET's Prepared by MBE", IEEE Elec. Dev. Lett., Vol. EDL-6, No. 1, Jan. 1985, pp. 1-2

4.35 Watkins E.T., Schellenberg J.M., Hackett L.H., Yamasaki H. and Feng M., "A 60 GHz GaAs FET Amplifier", 1983, IEEE MTT-S Digest, June 1983, pp. 145-147

CHAPTER FIVE

CONCLUSION

CONCLUSION

In this thesis, a new physically based DC model for HEMTs has been presented. The model consists of a set of analytic equations which are derived from fundamental principles and which must be solved simultaneously to obtain a solution for the currents and fields in the device. This model addresses several physical effects not included in other analytic HEMT models, such as the effects of surface charge, extension of the active area beyond the drain edge of the gate, and the additional drain current component arising from real space transfer of hot carriers out of the 2DEG (substrate current). The inclusion of these effects has increased the model complexity, but the computation time of the software incorporating the model is at least two orders of magnitude lower than that of typical 2D or Monte Carlo simulators. Therefore, the model can readily be incorporated in circuit simulation software in the future.

An RF model, which uses a lumped equivalent circuit to represent the HEMT, has also been presented. The majority of the equivalent circuit elements and their bias dependence are obtained from the DC model. The RF model is therefore inherently physically based, and no assumptions about the element's dependence on structure or bias are made.

Using the DC and RF models in a simulation program, comparisons between predicted and measured device performance were made. Good agreement between experiment and simulation at both DC and RF have been obtained, with the exception of the RF output resistance. The output resistance remains the most difficult RF equivalent circuit element to predict in all field effect transistors, and the HEMT has proved no exception. More information on the quantitative dependence of surface state occupation on bias, frequency and material processing is required before this disparity can be resolved.

An important result of this work is that good agreement between measured and predicted device performance, at both DC and RF, was obtained using a saturated electron drift velocity of 10^7 cm/s. This value is significantly lower than used in many HEMT models, where its value is chosen for a best fit with measured performance. The inclusion of substrate current, which may be a significant fraction of the total drain current, has led to good agreement with a lower value of drift velocity.

The assessment of the accuracy of any model, and the validity of the approximations in it, can only finally be determined by an extensive series of experiments in which accurate measurements are made on devices which can be repeatably produced. The experimentally determined dependences of electrical characteristics on single changes in device structural and material parameters and bias, will determine the suitability of each approximation used in the model. These procedures require many man-years of effort. This process is complicated at RF by the need to derive equivalent circuits from S-parameters. In general no unique circuit, with its element bias dependences, can be derived from a set of S-parameters. it is the author's view that the model presented here provides a sound basis for undertaking further coupled experimental and theoretical studies in the future.

Throughout this work, the IHEMT has been emphasised, and several potential advantages of this device over the CHEMT have been discussed. Whether these advantages are ever realised in practical devices or not, an understanding of the IHEMT is vital to the future modelling of multi-2DEG devices, which are currently emerging as contenders for microwave power devices. The IHEMT may well prove to be more predictable and reproducible because the barrier confining the 2DEG from the substrate is compositional, defined by the conduction band discontinuity.

IHEMTs with 0.5 micron gate length fabricated during the course of this work have shown comparable performance with CHEMT devices. Transconductances of up to 300 mS/mm at gate-drain breakdown voltages of more than 18V have been measured, the latter performance greatly exceeding that of comparable CHEMTs. Maximum tunable gain calculations have shown that these devices, with parasitic pad capacitances removed and a thinner cap layer, will provide useful amplifier gain up to millimeter-wave frequencies.

For HEMT modelling to progress further, there is a need for more quantitative data on some material properties, for example, the dependence of AlGaAs saturated electron drift velocity on aluminium mole fraction, and the Schottky-barrier height of various gate metals on AlGaAs and highly-doped GaAs. It is hoped that experimental programs, of comparable size to those used in determining the conduction band discontinuity dependence on mole fraction, will be initiated for these, and other, important issues.

APPENDIX A

HEMT REVIEW PAPER

This paper was published in the GEC Journal of Research, Vol. 4., No. 1, 1986.

HIGH ELECTRON MOBILITY TRANSISTORS (HEMTS) – A REVIEW

A. J. HILL and P. H. LADBROOKE

The last few years have seen the emergence of field-effect transistors which exploit the superior transport properties of electrons moving along the heterojunction interface between two compound semiconductor materials. Various acronyms have been coined for these devices (HEMT, TEGFET, MODFET, SDHT), all of which describe either the technology employed in creating the structure, or the resultant electronic properties. The term HEMT – for High Electron Mobility Transistor – provides the clue that these particular transistors exhibit high gain and high speed, among the highest of any semiconductor devices known today. The importance being attached internationally to HEMT can be gauged by the fact that digital integrated circuits using this technology will almost certainly form the basis of the fifth generation supercomputers. HEMT also shows much promise in monolithic microwave integrated circuits (MMICs), demonstrably outperforming the GaAs MESFET with respect to gain, low noise and frequency response. This paper reviews the principle, and the technology, of the High Electron Mobility Transistor.

1. INTRODUCTION

Field-effect transistors in which the primary conduction path is a two-dimensional electron gas have been known for a long time. The most easily understood and widely used is the silicon enhancement mode MOSFET, in which a sheet of electron charge carriers is attracted to the silicon/silicon dioxide interface to form the conduction path between source and drain. In recent years, a similar device has been implemented in compound semiconductors, with the difference that the wide bandgap insulator (silicon dioxide) is replaced by a wide bandgap, crystalline semiconductor alloy so that, among other things, the motion of electrons is not impeded by scattering at a rough interface. The device is therefore fast.

Several acronyms are used for these transistors including HEMT (High Electron Mobility Transistor), MODFET (MOdulation-Doped Field Effect Transistor), TEGFET or 2DEGFET (Two-dimensional Electron Gas FET) and SDHT (Selectively Doped Heterojunction Transistor). HEMT, the acronym used by the developers of the first such device, will be used in this paper.

Frequency dividers with maximum operating frequencies of 5.5 GHz at 300K and 10.1 GHz at 77K, using 1.0 μm dual-gate HEMTs, have been demonstrated, the latter representing the highest frequency ever reported for static dividers of any technology. For analogue applications, 0.35 micron gate length devices with room temperature noise figures of 1.3 dB and associated gains of 9 dB at 15 GHz have been fabricated. These figures are considerably better than those for conventional gallium arsenide (GaAs) MESFETs of similar geometry.

In this paper, the principles of HEMT operation and performance are reviewed. Devices based on gallium arsenide are discussed since the vast majority of HEMTs produced to date have used this material. The basic principles, however, are applicable to other III–V compounds. The Inverted HEMT (IHEMT), which so far has received less attention than the HEMT, is emphasized because of its promise, particularly as a large signal device.

Alternatives to aluminium gallium arsenide (AlGaAs), which has generally been used as the larger-bandgap material in GaAs HEMTs, are also presented.

A comprehensive reference list is included.

2. FUNDAMENTALS OF OPERATION

The HEMT is in many ways similar to an n-MOSFET, but consists of layers of compound semiconductors rather than silicon (fig. 1). The bulk silicon is replaced by a layer of undoped GaAs and the silicon dioxide is replaced by a layer of n-doped AlGaAs. Advanced techniques of crystal growth (notably molecular beam epitaxy, MBE,) are required to produce the layers with an atomically smooth heterojunction interface between them. It is a result of solid state physics that the electrons contributed by the n-doping in the AlGaAs are free to move through the entire crystal until such time as they fall into the lowest energy states allowed to them. In the HEMT, these lowest energy states are to be found just to the GaAs side of the heterojunction interface, with the outcome that all the electrons accumulate there in a thin sheet. Subsequently, they are free to move only in the two-dimensional plane of the interface, and the term 'two-dimensional electron gas' (2DEG) is used to describe the conduction electrons' state or condition.

The spatial separation of conduction electrons from their parent donor impurities reduces the rate at which the electrons are scattered by the ionized impurities, so that the momentum gained from an externally applied field is randomized less quickly. To the solid state device engineer, this is manifest as a preservation of coherence of electron velocity along

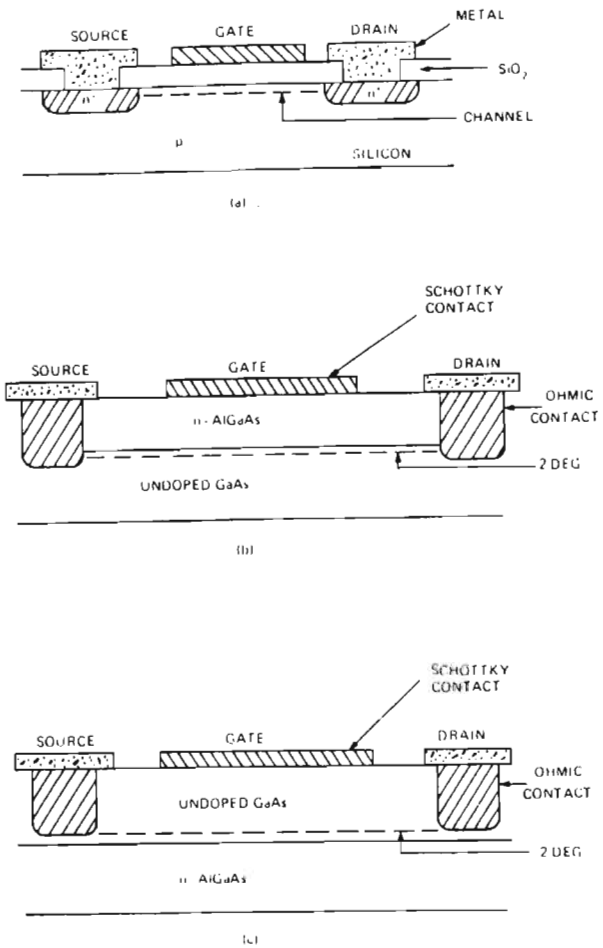


Fig. 1. (a) The silicon MOSFET
(b) Conventional HEMT
(c) The Inverted HEMT (IHEMT) structure

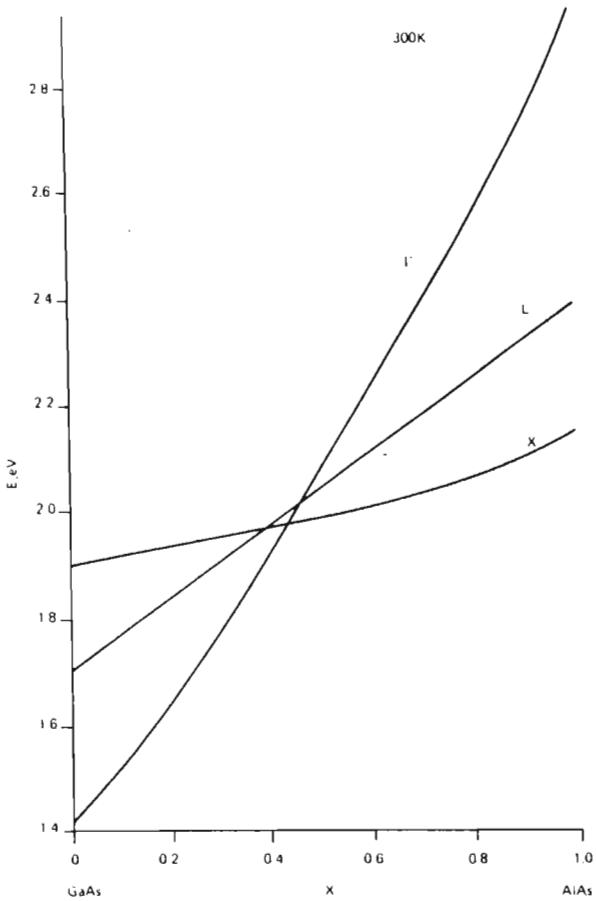


Fig. 2. Energies of the Γ, L and X conduction band valleys in the $Al_xGa_{1-x}As$ system as a function of aluminium mole fraction, x. Energies are relative to the top of the valence band at K = (0, 0, 0). From ref. 4.

the direction of the applied field, that is, an enhanced electron mobility, to which the ultimate speed of the transistor is related.

The ability to obtain a large sheet concentration of conduction electrons without incurring the penalty of a commensurately large impurity scattering rate is one of the major advantages of the HEMT structure over the standard MESFET.

Drain-source conduction in both HEMT and MESFET devices is modulated by the voltage applied to the Schottky-barrier gate electrode placed between the source and drain contacts.

Details of the device operation outlined above are presented in subsequent sections of the paper.

3. HETEROJUNCTIONS AND MODULATION DOPING

The larger-bandgap material used with GaAs to form the heterojunction must be closely lattice-matched to the GaAs to prevent the upper layer growing in a polycrystalline form. AlGaAs is lattice-matched to within 0.1 per cent of GaAs and the growth of extremely abrupt heterojunctions by molecular beam epitaxy is possible using these two materials[1–3].

The bandgap of $Al_xGa_{1-x}As$ depends on the Al mole fraction, x. The energies of the conduction band minima measured from the top of the valence band (K = 0,0,0) are shown in fig. 2[4]. The band diagram of a typical GaAs/n-$Al_xGa_{1-x}As$ heterojunction is shown in fig. 3. The sum of the conduction- and valence-band discontinuities is equal to the energy gap discontinuity determined using fig. 2.

If only the $Al_xGa_{1-x}As$ (AlGaAs) is doped ('modulation doped heterojunction')[5] and the conduction band edge in the GaAs is lower in energy than the donor energy level in the AlGaAs, electrons which have sufficient energy to overcome the donor binding energy will transfer from the AlGaAs to the GaAs[6, 7]. This process satisfies the equilibrium requirement of a constant Fermi energy through the heterojunction. The electron transfer from AlGaAs to GaAs causes strong electric fields (of the order 10^5 V/cm) perpendicular to the interface which in turn causes bending of the energy bands near the interface. The band-bending results in the formation of a quasi-triangular potential well approximately 10 nm wide to which the transferred electrons are confined, forming a two-dimensional electron gas (2DEG).

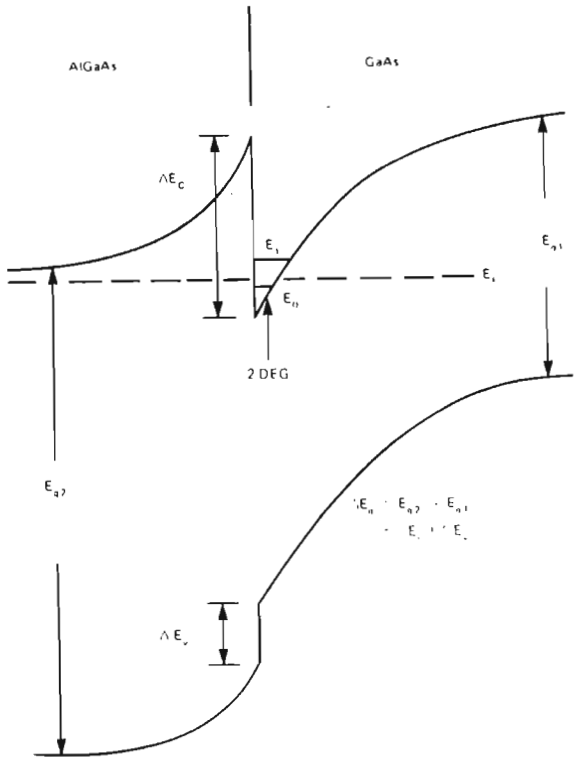


Fig. 3. Band diagram of a GaAs/n-$Al_xGa_{1-x}As$ heterojunction showing the 2-dimensional gas of free electrons formed on the GaAs side of the interface by electrons originating from the n-doping in the $Al_xGa_{1-x}As$. E_0 and E_1 are the energies of the sub-bands in the potential well. E_{g1} and E_{g2} are the GaAs and AlGaAs energy bandgaps respectively. ΔE_c and ΔE_v are the conduction and valence band discontinuities respectively.

The electron energies in the quantum well are increased because of their quantum-mechanical confinement and form discrete energy subbands. In a typical HEMT heterojunction, the Fermi level is higher in energy than the lowest energy sub-band, E_0, which necessitates the use of Fermi-Dirac rather than Maxwell-Boltzmann statistics in calculating the sheet carrier concentration, n_{so}, in the degenerate 2DEG. A thin layer of the AlGaAs adjacent to the hetero-interface (spacer layer) is left undoped in HEMT devices to separate further the electrons in the 2DEG from the ionized impurity scattering centres, so enhancing the mobility of the 2DEG electrons[8]. The number of electrons confined in the 2DEG is a function of the thickness of this layer, d_{sp}, the donor density in the doped AlGaAs, N_D, and the Al mole fraction. The relationships between these quantities and the resultant sheet carrier concentration are complex and are discussed in more detail in Section 5.

A two-dimensional hole gas (2DHG) can be formed in an analagous manner by doping the AlGaAs with an acceptor rather than a donor impurity.

The number of carriers in the 2DEG is strongly dependent on the conduction band discontinuity, ΔE_c, at the heterojunction interface. While the total energy gap difference, ΔE_g, across the interface can be deduced from the Al mole fraction, x, using the relationships of fig. 2, the ratios of the conduction- and valence-band discontinuities to the total energy gap discontinuity are not well established. In 1974, Dingle et al.[9] deduced a conduction band discontinuity of 88 per cent of the total energy gap discontinuity (and hence $\Delta E_v = 0.12\Delta E_g$) from optical absorption spectra in rectangular GaAs quantum wells confined by AlGaAs layers with a mole fraction of 0.2. These ratios have become known as 'Dingle's Rule' and have been applied by a large number of authors to AlGaAs/GaAs heterojunctions with widely differing aluminium mole fractions. There appears to be no experimental justification for the assumption that $\Delta E_c/\Delta E_g$ should be constant for all mole fractions, particularly since AlGaAs becomes an indirect semiconductor for mole fractions greater than 0.45. Recently, papers questioning the accuracy of 'Dingle's Rule' have appeared; Waldrop et al.[10] found $\Delta E_v/\Delta E_g = 0.3$ for AlAs on GaAs, Miller et al.[11] found $\Delta E_c/\Delta E_g = 0.57$ for $x = 0.3$ and Batey et al.[12] deduced $\Delta E_v/\Delta E_g = 0.4$ for $x = 0.38$ and 0.6. In all cases, ΔE_g refers to the direct energy gap difference. More work in determining the correct ratios for all Al mole fractions is required as almost all papers on HEMT structures have assumed 'Dingle's Rule' to be valid. Unfortunately, not all measurement techniques produce reliable results[13] and different theoretical models of heterojunction line-ups predict widely different discontinuities[13–17]. Accurate modelling and optimization of HEMT devices will depend on the satisfactory resolution of this issue.

4. TRANSPORT IN HEMT STRUCTURES

4.1. Low Field Mobility in 2DEGs

In 1978, Dingle and co-workers at Bell Labs. grew the first modulation doped superlattices by MBE and observed mobility enhancement at low temperatures[5]. In doped bulk GaAs, the low temperature Hall mobility is dominated by ionized impurity scattering and is proportional to $T^{1.5}$. The 77K mobility of bulk material with low compensation ratio is typically $1.5 \times 10^5\,cm^2/Vs$ for a donor density of $10^{14}\,cm^{-3}$ and reduces to $8000\,cm^2/Vs$ for a doping density of $10^{17}\,cm^{-3}$ because of ionized impurity scattering[18]. Donor densities of $10^{17}\,cm^{-3}$ are typical of GaAs MESFETs with low on-resistances. Modulation doping, by greatly reducing ionized impurity scattering, allows a similar number of carriers in HEMT devices while maintaining high mobility. The background impurity concentration in a practical HEMT is approximately $10^{14}\,cm^{-3}$, but the large density of electrons in the 2DEG (typically $10^{12}\,cm^{-2}$) results in carrier screening which enhances the mobility. Mobilities as high as $1.95 \times 10^5\,cm^2/Vs$ have been achieved in some devices at 77K[19].

Below a temperature of 100K, optical phonon scattering is negligible, but above this temperature, it is the dominant scattering mechanism in both 2DEGs and bulk GaAs[20, 21]. The 2DEG electron mobility for polar optical scattering is proportional to T^{-k} where k has been calculated as 2.03[22] and values

between 1.77 and 2.4 have been measured[20, 23, 24]. The 300K mobility is limited to approximately 9000 cm^2/Vs for both 2DEGs and pure bulk GaAs. The dependence of electron mobility on temperature for a typical HEMT device is shown in fig. 4.

At lower temperatures, the Hall mobility is limited by several other scattering processes, the most important of which is Coulombic scattering[6, 25] because of the close proximity of the 2DEG to donor ions in the AlGaAs. The wave function of electrons in the 2DEG extends from the GaAs into the AlGaAs and the resultant electrostatic interaction with the donor ions scatters the electrons[26]. For a constant sheet carrier concentration in the 2DEG, the Coulombic scattering rate is inversely proportional to z^2 where z is the distance from the donor ions to the peak of the electron wave function in the 2DEG[19, 22, 27]. The incorporation of a spacer layer of undoped AlGaAs increases z, and the resultant increase in mobility demonstrates the dominance of Coulombic scattering, at least for spacer layers less than approximately 10 nm wide in typical HEMTs. The first work on spacer layers yielded a maximum 77K mobility twice as large as the previous best reported for a single heterojunction[8]. Some researchers found that the mobility increases monotonically as the spacer layer width is increased up to 20 nm[19, 27]; others[6, 8, 22, 26] have found that the mobility increases up to a certain width (in one case, 7.5 nm) and then decreases as the spacer width is increased further, while a recent paper found little correlation between mobility and spacer width[28]. The differences in these results are due to the influence of other mechanisms, in particular the reduction of the sheet carrier concentration with increasing spacer layer thickness, which reduces the effect of carrier screening[6, 19, 22, 26]. The Coulombic scattering rate also increases as the ionized donor concentration in the AlGaAs increases, but the dependence of the measured mobility on the doping concentration, N_D, is complicated by the dependence of the sheet carrier concentration on N_D, Al mole fraction, acceptor concentration and temperature due to incomplete donor ionization, especially at low temperatures[29, 30].

The peak mobility obtainable in HEMT devices is limited by scattering by impurities incorporated in the nominally undoped GaAs near the heterojunction as a result of imperfect growth techniques[19]. This impurity scattering is reduced by carrier screening and improved growth methods. The rapid increases in the record mobilities reported by different laboratories during the early years of HEMT development were usually due to improved growth conditions such as cleaner MBE systems and optimum growth temperatures which yielded cleaner and smoother interfaces[31, 32]. Optimal MBE growth techniques may both reduce interface roughness, which scatters carriers, and also reduce the concentration of interface charge caused by charged interface states or deep states in the undoped AlGaAs, which also act as scattering centres[19, 33].

The maximum achievable mobility in HEMT structures is limited by piezoelectric and acoustic deformation potential scattering and this ultimate limit has been calculated as 6.5×10^6 cm^2/Vs at 1K for a sheet carrier concentration of 4×10^{11} cm^{-2}[33]. Interface alloy scattering, caused by the interaction of the electron wave function with potential fluctuations due to uneven aluminium distribution in the AlGaAs, plays a minor role in limiting the low temperature mobility[6, 21].

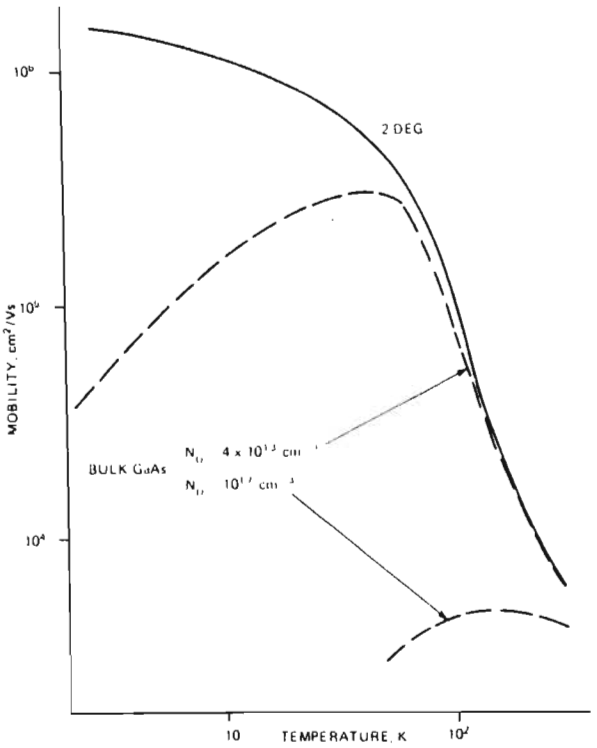


Fig. 4. Low-field electron mobility in a 2-dimensional electron gas, as a function of lattice temperature, compared with the mobility in bulk GaAs at two different doping levels[69].

The role of carrier screening in reducing the effect of the scattering mechanisms has been exclusively investigated in both theoretical studies[25, 34], and practical measurements. The Hall mobility is proportional to n_s^k, where measured values of k vary from sample to sample in the range 0.45 to 2.0[32, 35–37]. The value of k tends to 0.5 for higher mobility samples, in good agreement with theoretical predictions[25, 34]. Above a sheet concentration of approximately 7×10^{11} cm^{-2}, this relationship is no longer valid[19] because inter-sub-band scattering in the 2DEG becomes important and because thin spacer layers and large donor densities, which may degrade the interface, are required to obtain large sheet concentrations[38].

As the sheet carrier concentration is increased, the Fermi energy (and hence the average electron kinetic energy) in the degenerate 2DEG increases. This further enhances the mobility because impurity scattering is less effective for higher-energy electrons[33, 39].

The spacer layer width, d_{sp}, required to maximize the electron mobility is a compromise between reducing Coulombic scattering (increasing d_{sp}) and reducing the effect of all the scattering mechanisms by

increasing the sheet carrier concentration (decreasing d_{sp}). The previously mentioned differences in measured mobilities as a function of spacer width reflect the differing relative importance of the various scattering processes from sample to sample. In designing practical HEMTs and IHEMTs, mobility is not the only criterion determining the optimum device structure. These other criteria are discussed in Section 5.

The dependence of mobility on the aluminium mole fraction, x, is also a result of several interdependent factors. Early workers[32] found the mobility increased monotonically as the mole fraction was reduced, a finding which was attributed to increasing AlGaAs quality for lower mole fractions. Other researchers found that the mobility increased as the mole fraction increased, up to 0.3[40] or 0.4[26, 41]. As the mole fraction is increased, the conduction band discontinuity also increases which causes the electron wave function to extend less far into the AlGaAs, which reduces the interaction with donor ions (Coulombic scattering). Clearly, the relative importance of this effect will depend on the spacer layer width and other parameters. Other device characteristics, such as the sheet carrier concentration, also vary with the aluminium mole fraction and must be considered in choosing the optimum mole fraction for a particular application.

4.2. 2DEG Transport in moderate electric fields

The extremely high mobilities, characteristic of 2DEGs under small applied electric fields at cryogenic temperatures, degrade as the field strength is increased to the levels present in FET devices subject to typical operating voltages. Electrons in the 2DEG rapidly gain energy and become hot because of their high mobility at low field strengths. Above a critical field in the range 50–500 V/cm[26, 42, 43], however, the mobility decreases as a result of spontaneous phonon emission[42, 44] by the hot electrons. The energy separation between the first two sub-bands in the 2DEG is typically 40 meV and high mobility electrons in the lowest energy sub-band quickly gain sufficient energy from the field to be scattered into the lower-mobility second sub-band[42]. The higher the initial mobility in the lower sub-band, the faster the mobility decreases as a function of the applied electric field[26] (fig. 5). Even in moderate electric fields, the dominant scattering mechanism is polar optical phonon scattering rather than Coulombic scattering which dominates the cryogenic Hall mobility. Differences in the low-field mobility at 77K of samples with different spacer widths become negligible as the electric field is increased because of the shift in dominant scattering process[42]. At higher temperatures, where polar optical phonon scattering is dominant, the mobility remains constant as the field is increased[26] (fig. 5).

Experimentally derived velocity-field characteristics of 2DEGs show reasonable agreement with Monte-Carlo simulations for undoped GaAs[45, 46]. In fields greater than 2 kV/cm, the electrons can be considered to be subject to three-dimensional scattering effects in the same manner as bulk GaAs[42, 47]. The saturated drift velocity of the electrons is the same as in bulk GaAs[45, 48]. In 2DEGs, the peak velocity occurs at a field strength of approximately 2 kV/cm and increases from 2×10^7 to 3×10^7 cm/s as the temperature decreases from 300K to 77K[42, 46].

The enhanced peak electron velocity of 2DEGs at room temperature, further increased at cryogenic temperatures, is one reason for the faster operating speeds of HEMT devices. Because the electric field in the channel or 2DEG under the gate in both HEMTs and MESFETs is greater than the fields at which the Hall mobility is measured, it is the velocity-field characteristic rather than the Hall mobility which primarily determines the device speed. The enhanced mobility increases the speed of the HEMT by reducing access resistances, (especially the source resistance), which degrade the transconductance of the device.

In FET devices with gate lengths less than one micron, velocity overshoot effects[49] become important in determining device speed. Electrons accelerated by the electric field in the channel or 2DEG may attain a velocity greater than the static peak velocity before being scattered into a higher-energy sub-band (fig. 6). This dynamic over-velocity has an effect on the switching speed of the FET since it increases the average electron velocity. The amount of overshoot increases as the carrier low-field mobility increases and hence the enhanced mobility

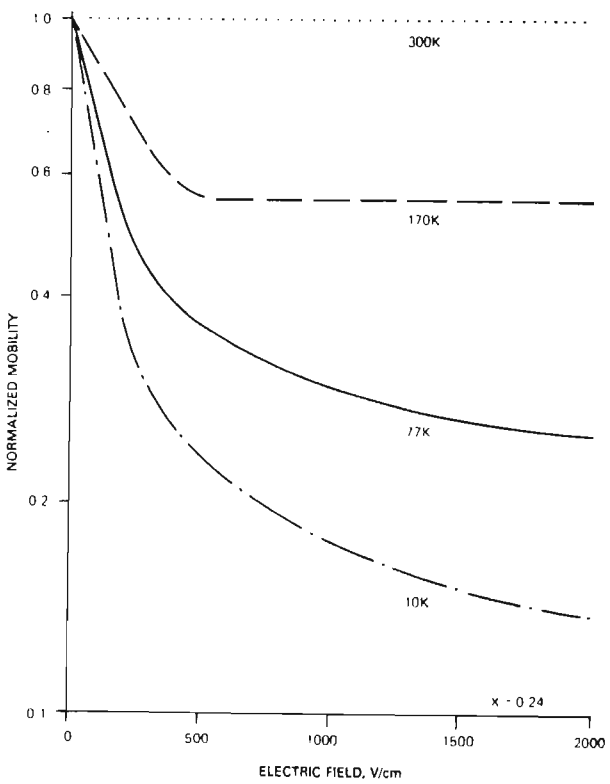


Fig. 5. Dependence of electron mobility in a 2-dimensional electron gas on electric field strength[26].

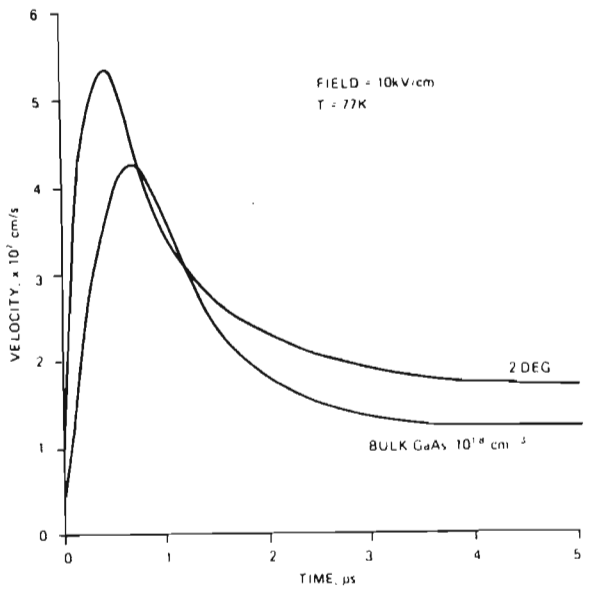


Fig. 6. Time response of electrons in a 2-DEG and 10^{18} cm^{-3} n-doped GaAs to a 10 kV/cm step function in electric field for a lattice temperature of 77K. From Monte Carlo simulation[45].

of HEMTs, particularly at cryogenic temperatures, means that this effect is more important in HEMTs than MESFETs and so contributes to the speed advantage of HEMTs[50]. Overshoot may increase the average electron velocity at 77K to 4×10^7 cm/s over a distance of 0.4 micron[51].

4.3. Transport in AlGaAs

Electrons in the doped AlGaAs have low mobility at cryogenic temperatures because of the ionized impurities present in the material, and at 300K the mobility is typically less than 1200 cm^2/Vs[52]. Electrons in the 2DEG may gain sufficient energy from the field to overcome the conduction band edge discontinuity and transfer back into the AlGaAs (real space transfer)[53]. It is desirable that the carriers in the AlGaAs have a low mobility since they form a component of current analagous to substrate current in MESFETs.

The mobility in the AlGaAs is a function of the crystal quality, which is dependent on the growth conditions. The mobility decreases as the donor concentration and compensation ratio increase. The mobility also decreases as the mole fraction is increased up to 0.45[40, 52] because the intervalley scattering rate increases as the energy differences between the Γ and L/X valley minima decrease (fig. 2). For larger mole fractions, the mobility remains small because of the low mobility of electrons in the L and X valleys. A typical AlGaAs sample with an aluminium mole fraction of 0.3 and donor concentrations of 10^{17} cm^{-3} would exhibit electron mobilities of between 500 and 1000 cm^2/Vs at both 300K and 77K[40, 52].

The saturated drift velocity in AlGaAs is also dependent on the aluminium mole fraction but the exact relationship is not well established. Recent measurements[48] found the saturated drift velocity in MBE-grown AlGaAs to be 7×10^6 and 3×10^6 cm/s for mole fractions of 0.24 and 0.3 respectively. Earlier work[54] yielded a value of 4.8×10^6 cm/s for a mole fraction of 0.45 in LPE-grown material.

5. HEMT DEVICES

5.1. Fabrication

A typical sequence of layers for a HEMT may consist of one micron of undoped GaAs grown directly on the substrate, followed by 40 nm of $Al_{0.3}Ga_{0.7}As$, of which the first 5 nm is left undoped and the rest silicon-doped to 1×10^{18} cm^{-3}, followed by an n^+GaAs capping layer several hundred angstroms thick. The capping layer prevents oxidation of the AlGaAs layer. At a growth rate of 1 μm/hr, the optimum MBE growth temperatures for GaAs and AlGaAs are 600°C and 700°C respectively[1–3].

Device fabrication after layer growth is similar to that of conventional MESFETs. For discrete devices, a mesa is etched, ohmic metal deposited and alloyed, gate metal deposited and gates defined. The ohmic metal is usually a AuGe/Ni/Au alloy and the alloying stage is critical in order to obtain low resistance contacts to the 2DEG.

For HEMT integrated circuits, a boron or proton implant to isolate devices is preferable to a mesa etch, so as to keep the process planar. The source and drain resistances may be reduced by using a self-aligned silicon implant to form the source and drain[55, 56]. A rapid thermal anneal is used to activate the Si implant to avoid mobility degradation caused by long anneal times[57, 59].

Both enhancement (normally-off) and depletion (normally-on) devices may be made by suitable selection of gate metal (to adjust the barrier height) and AlGaAs thickness and doping density. For a depletion device, the AlGaAs thickness is chosen so that the depletion region from the Schottky-barrier just overlaps the one caused by electron transfer to the 2DEG. The AlGaAs must be fully depleted to prevent the formation of a parasitic MESFET in parallel with the HEMT. For an enhancement device, the AlGaAs is made thinner or the Schottky-barrier height increased so that the 2DEG and the AlGaAs are both depleted at zero bias. Enhancement and depletion devices can be made on the same substrate by either etching a gate recess for enhancement mode devices[60] or using different metals for the two gates[61], or leaving the depletion devices ungated in logic circuits[62, 63]. The free surface potential of AlGaAs with an aluminium mole fraction of 0.23 is approximately 0.33 eV, whereas the barrier height of a gold gate on the same material is approximately 1.0 eV[63]. By the correct choice of AlGaAs thickness and doping density, gated devices will be normally-off, and ungated devices, normally-on. Thus enhancement-depletion (E-D) logic using ungated devices as depletion loads can be implemented. This technology is simpler than MOS where an extra implant or diffusion is required to implement depletion devices.

The simple planar process used for HEMT fabrication makes this technology suitable for VLSI circuits. Low power dissipation of HEMT logic and high thermal conductivity of GaAs are two further advantages of this technology over silicon-based technology for VLSI. The radiation hardness of HEMTs is superior to that of MOSFETs because of the lower density of interfacial deep states compared with the density at the Si-SiO_2 interface in MOSFETs[64]. Diodes, capacitors and resistors may also be fabricated without additional processing steps which means that analogue integrated circuits with HEMTs are also viable.

Recently, the successful fabrication of GaAs HEMTs on Si[65] and Ge[66] substrates has been demonstrated. These substrates have the advantage of lower cost than GaAs substrates.

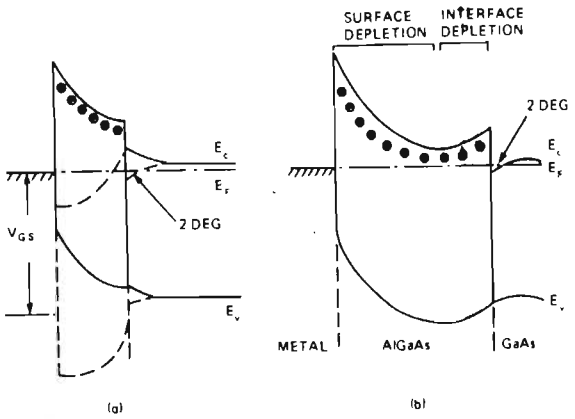


Fig. 7. Band diagrams of
(a) Enhancement-mode HEMT and
(b) Depletion-mode HEMT, for zero gate bias (——) and positive gate bias (– – – –).

5.2. HEMT Charge Control and I-V characteristics

The charge control of HEMTs is similar to that in MOSFETs[67], but HEMT performance is enhanced because of the larger mobility and velocity of the 2DEG electrons[68] and because the relative permittivity of AlGaAs is larger than SiO_2, which increases the transconductance of the HEMT for the same thickness of AlGaAs or SiO_2[69].

The band diagrams of an enhancement and depletion mode HEMT under zero bias are shown in fig. 7. In the enhancement device, the space charge region of the gate Schottky barrier extends into the GaAs and the resultant bandbending precludes the formation of a 2DEG under zero-bias conditions. The donors in the AlGaAs do not cause the formation of a 2DEG but instead act to control the surface potential of the GaAs at the heterojunction[37, 70]. A positive bias causes downward bandbending and electron accumulation in a 2DEG at the interface (– – – in fig. 7). In the depletion device, a 2DEG exists at zero bias and a negative gate potential is required to deplete the 2DEG.

The dependence of sheet carrier concentration on the gate-source voltage is linear[71, 72] below a critical sheet concentration. Above this value, the concentration saturates and further increase in gate potential leads to the formation of a neutral region in the AlGaAs where the donor level crosses the quasi-Fermi energy level. This region will allow parallel conduction in the AlGaAs, forming a parasitic MESFET. The minimum gate voltage at which a neutral region forms is effectively the threshold voltage of the parasitic MESFET and sets an upper limit on the useful gate voltage which may be applied to the gate[73, 74]. A second departure from linearity occurs for gate voltages near the HEMT threshold voltage because the quasi-triangular well widens and hence the energy of the electrons decreases[69]. This change also causes the transconductance to decrease.

The drain-source current can be approximated by

$$I_{DS} = qn_s vw \qquad (1)$$

where w is the channel width and v is the average electron drift velocity in the channel, including velocity overshoot effects. The sheet concentration dependence on gate voltage is given by

$$n_s = \frac{\varepsilon}{qd_e}(V_{GS} - V_T) \qquad (2)$$

where V_T is the threshold voltage and d_e is the AlGaAs thickness plus approximately 8 nm[75, 76]. These simple approximations are very similar to those used in modelling MOSFETs. The increased 2DEG mobility at cryogenic temperatures yields larger transconductances in HEMTs by increasing the average electron drift velocity, whereas in bulk MOSFETs and MESFETs, there is no such increase. A further performance advantage over these devices occurs because there is a sharper pinchoff characteristic which means that less gate voltage swing is used in traversing the low transconductance region near pinchoff.

Details of analytical and numerical modelling of HEMT devices will be found in references 70 and 76 to 78.

5.3. HEMT Optimization

Early HEMTs were optimized to yield the maximum possible Hall mobility with little regard to the effect of the large spacer layer width on the sheet carrier concentration. Later, when it was realized that the high mobility plays only a secondary role in determining device switching speed, spacer layer widths were reduced from approximately 15 down to 2 nm devices for logic applications. The saturated drift velocity of electrons in the 2DEG, unlike the mobility, is independent of the spacer width. Thin spacer layers yield large sheet carrier concentrations which in turn yield enhanced transconductance and drain current per unit gate width. The spacer layer cannot be removed completely because the increased donor-electron interaction would reduce the mobility to the point where the enhanced average electron velocity in the HEMT would be degraded[79].

For logic circuits, the HEMT should be optimized

for maximum transconductance and drain current to charge and discharge load capacitances as fast as possible. This requires using the thinnest possible doped AlGaAs layer and hence the highest possible dopant concentration to obtain the required sheet carrier concentration in the 2DEG. Although both transconductance and gate-source capacitance increase as the AlGaAs thickness is reduced, the interconnection capacitance in logic circuits remains constant and hence an overall speed improvement results.

The extrinsic transconductance, g_m, of the HEMT is reduced from the intrinsic transconductance, g_{mo}, as

$$g_m = \frac{g_{mo}}{1 + g_{mo}R_s} \quad (3)$$

where R_s is the parasitic source resistance. To maximize the transconductance therefore, the source resistance must be minimized. The problem of making a low resistance ohmic contact through the AlGaAs to the 2DEG is the most important factor hindering even greater performance in HEMTs and is a major reason favouring the further development of IHEMTs in the future.

Little work has been reported on the optimization of HEMT devices for analogue applications. The noise figure can be optimized by minimizing the ratio of the gate-source capacitance to the square root of the transconductance[80]. This can most easily be achieved by reducing the parasitic source resistance. Increasing the aluminium mole fraction increases the conduction band-edge discontinuity and hence reduces the number of hot electrons injected into the AlGaAs. Since generation-recombination noise is generated in the AlGaAs (see section 7), this may also decrease the noise figures of HEMTs.

5.4. Practical performance

One micron gate length devices have been fabricated with maximum transconductance and drain currents of 250 mS/mm and 300 mA/mm at 300K respectively[81]. At 77K the transconductance increased to 400 mS/mm. The largest transconductance achieved in HEMTs was 450 and 570 mS/mm at 300K and 77K respectively, with a gate length of 0.33 micron and an AlGaAs buffer layer between the 2DEG and the substrate[82].

HEMT logic circuits have received more attention than analogue amplifiers and many ring oscillator and frequency divider results have been published. A 19-stage direct-coupled FET Logic (DCFL) ring oscillator with 0.4 micron gate length drivers has been demonstrated with a gate delay time of 11ps at 77K[83], the fastest switching speed ever attained in any technology. The static divider with an operating frequency of 10.1 GHz at 77K is the fastest static frequency divider of any technology to date. These results are more than twice as fast as the fastest silicon dividers. Low power HEMT D-type flip-flops with driver gate lengths of 0.6 μm have been fabricated which have the same maximum operating frequency as 0.6 μm MESFET flip-flops (5.9 GHz), but with only one fifth as much power consumption[84]. HEMT logic at 77K has power-delay products comparable to those of Josephson junctions at 4.2K.

A 4k-bit static RAM, with an access time of 2ns at 77K has been fabricated by Fujitsu. This is faster than any silicon SRAM ever produced and demonstrates the great potential of HEMT logic for high speed LSI circuits.

For analogue applications, 0.35 μm gate length HEMTs with noise figures of 1.3, 1.5 and 2.7 dB, with associated gains of 9, 10.5 and 5.9 dB at 15, 18 and 34 GHz respectively, have been reported[80]. A HEMT with an equivalent noise temperature of 10.5K at a lattice temperature of 12.5K, at 8.5 GHz has also been fabricated[82]. The material factor in one form of the Fukui noise figure equation is typically 2.5 for a GaAs MESFET and approximately 1.6 in HEMTs[80].

The low power, high speed and low noise performance of HEMTs ensures that these devices will play an increasingly important role both in high frequency analogue and in digital circuits in the future.

5.5. P-HEMT

HEMT devices where conduction between source and drain occurs in a two-dimensional hole gas (2DHG) have received little attention because of the significantly lower hole mobility in GaAs. However, the development of a P-HEMT offers the possibility of producing ultra-low power, complementary logic circuits, similar to CMOS.

The hole mobility of 2DHG has been shown to have a similar temperature dependence to a 2DEG, but reduced by a factor of 1.5[85]. The saturated drift velocity of holes in a 2DHG has been estimated as 5×10^6 cm/s at 77K and P-HEMTs with a transconductance of 35 mS/mm for a 1.5 micron gate length device have been demonstrated[86]. It is estimated that this figure could increase to 100 mS/mm for an optimized device, which would be acceptable for a complementary logic circuit.

6. INVERTED HEMT

The Inverted HEMT (fig. 1), IHEMT, has been largely neglected in favour of the standard HEMT because of the higher mobilities which have been obtained in the latter device. However, the IHEMT structure inherently overcomes several of the problems of the HEMT structure and also has potential advantages as a large signal device. Much of the foregoing discussion on transport and HEMT characteristics applies implicitly to the IHEMT and will not be repeated here.

6.1. Fabrication

Initial investigation into mobility enhancement in inverted heterojunctions grown by MBE failed to produce a 2DEG[6], although previously this had been achieved in LPE-grown material[87]. Further work, however, showed that, although a high

mobility 2DEG could be obtained in standard heterojunctions grown by MBE at substrate temperatures in the range 600–680°C, the inverted junction only yielded high mobility when grown at temperatures in a narrow range around 700°C[31, 35, 38].

The inverted heterojunction is extremely sensitive to the quality of the AlGaAs. The highest mobilities ever reported for an inverted heterojunction with a 10 nm spacer layer, are 42000 and 70000 cm^2/Vs at 77K and 5K respectively[88], which are considerably lower than those obtained in standard heterojunctions. The lower mobilities in inverted heterojunctions have been attributed to interface roughness[89], impurity buildup at the interface and silicon diffusion into the spacer layer and GaAs, because of the accumulation of silicon at the upper surface of the doped AlGaAs (surface segregation)[88, 90]. The use of a three period undoped superlattice of GaAs and AlGaAs as a replacement for the spacer layer has been shown to yield enhanced 2DEG mobilities at the inverted hetero-interface, as high as 106,000 and 256,000 cm^2/Vs at 77K and 10K respectively[91]. The multiple interfaces of the superlattices getter impurities, relieve lattice mismatch strain, smooth the interface and act as a diffusion stop to silicon donors. The lower mobilities obtained in inverted heterojunctions do not preclude the fabrication of high performance IHEMTs since increases in the Hall mobility above 30000 cm^2/Vs yield minimal increases in the device cut-off frequency[70]. The disadvantage of lower mobility in the IHEMT is minimal in comparison to the inherent structural advantages of the device. Theoretical predictions[92] indicate that IHEMTs should exhibit similar transconductances to HEMTs.

For device fabrication, there are two considerations. In the conventional HEMT, it may be necessary to make a Schottky gate onto AlGaAs which oxidizes very rapidly and, secondly, ohmic contacts have to be made through the AlGaAs to contact the 2DEG. In the IHEMT, the uppermost layer is GaAs and hence these problems reduce to those encountered in the well-established MESFET technology. Making repeatable low resistance ohmic contacts to HEMTs is one of the challenges remaining in HEMT development. The more stable GaAs surface may result in higher yield and enhanced reproducibility in IHEMT devices, considerations which are vital in the development of HEMT or IHEMT integrated circuits in the future[92].

6.2. Small-signal characteristics

Charge control in IHEMTs is very similar to that in HEMTs. In the HEMT, undesirable drain conductance occurs in the saturation characteristics as a result of hot electron injection from the 2DEG into the GaAs buffer and substrate. This problem can be overcome by the growth of a thick, undoped AlGaAs buffer layer under the GaAs layer containing the 2DEG[82]. In the IHEMT, however, the doped AlGaAs is between the 2DEG and the substrate and hence electron confinement is inherent in the structure[93]. In order to prevent the formation of a parasitic MESFET in the IHEMT AlGaAs layer, the AlGaAs must be sufficiently thin so that it is fully depleted[92, 93] by charge transfer to the 2DEG. IHEMTs designed by the authors and fabricated at Hirst Research Centre have yielded extrinsic transconductances up to 180 mS/mm, one of the highest values reported for the IHEMT to date.

6.3. Large signal Characteristics

The available output power in FETs is commonly estimated as

$$P = 0.125\, I_{dss}\, V_{dsb} \qquad (4)$$

where I_{dss} is the zero gate-voltage saturation current and V_{dsb} is the drain-source breakdown voltage. In short gate devices, the saturation current is determined by the average electron velocity in the device. The higher peak velocity and velocity overshoot enhancement in both HEMT and IHEMT will increase this current over that in conventional MESFETs. The physical process governing the breakdown voltage at, or near, channel pinch-off, is avalanche breakdown at the drain edge of the gate. The fixed ionic charge in the channel of a pinched-off MESFET (fig. 8(a)) induces a non-uniform charge distribution on the lower surface of the gate metal. The resultant electric field is also non-uniform and is largest at the drain edge of the gate. When the field at the drain end of the gate exceeds a critical value, avalanche breakdown occurs. In an IHEMT, the fixed ionic charge (from the ionized donors) is further away from the gate (fig. 8(b)) and will induce a more uniform charge distribution along the gate electrode. Consequently, the device should withstand greater drain potentials[94]. In the standard HEMT, the fixed ions are not displaced from the gate and should exhibit similar breakdown characteristics to the MESFET.

IHEMTs have been fabricated at Hirst Research Centre with 15V drain-source breakdown voltage, the highest value reported. This result suggests the IHEMT may have potential as a large signal device and further practical and theoretical investigations are currently being undertaken by the authors[95]. Among other things, this work will endeavour to clarify the importance of changing surface-state charge with device operating conditions in suppressing, or partially suppressing, the electric field singularity at the drain edge of the gate in both HEMTs and MESFETs.

7. AlGaAs, DIFFICULTIES AND ALTERNATIVES

The use of AlGaAs as the larger-bandgap material in heterojunctions on GaAs introduces several undesirable effects to HEMT or IHEMT performance. The effects are mainly related to the large number of deep states present in the AlGaAs and the high donor activation energy at mole fractions typical in heterojunction devices.

Aluminium mole fractions between 0.2 and 0.4 are

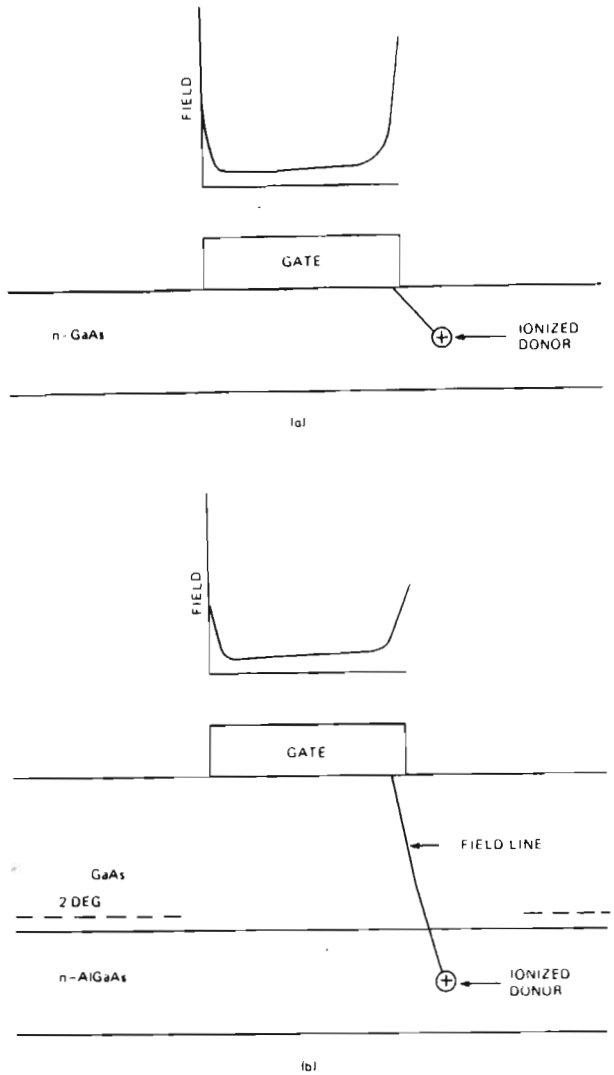


Fig. 8. Schematic of electron charge distribution along the gate electrode (a) in a MESFET (b) in an IHEMT

typical in HEMTs because the 2DEG mobility is largest in this range[41]. Larger mole fractions, while yielding larger conduction band-edge discontinuities, result in poor mobility because of the decreasing interface quality[32]. Unfortunately, many of the undesirable effects caused by AlGaAs are most pronounced for mole fractions between 0.2 and 0.45. The preferred *n*-type dopant for AlGaAs is silicon because it is less amphoteric than germanium and because there is less surface segregation than with tin. The substitution of other dopants for silicon does not alter the trends observed with silicon[52].

The high concentration of deep states and the large donor activation energy in AlGaAs results in a lower number of free electrons for a given doping density than would be obtained in bulk GaAs. Donor atoms in AlGaAs occupy two donor levels, one deep and one shallow, and even at room temperature there may be incomplete ionization of these donors[51, 52]. As the temperature is decreased to approximately 150K, there is an exponential freeze-out of carriers to the deep donors and for still lower temperatures, freeze-out to the shallow donors. This results in an undesirable decrease in 2DEG sheet carrier concentration and shift in threshold voltage as the temperature is decreased[96, 97]. The current-voltage characteristics and transconductance are strong functions of the number of occupied deep states in the AlGaAs, particularly at low temperatures. The transconductance at cryogenic temperatures may be greater or less than at room temperature, depending on the density of occupied deep states[98, 99]. It is desirable to find an alternative to AlGaAs that has a lower concentration of available deep states to reduce these unwanted characteristics. The ratio of shallow to deep donor concentrations is dependent on the growth conditions[52], dopant density and aluminium mole fraction. Donor activation energies decrease as the doping level is increased but the ratio of shallow to deep donors also decreases[52]. The dominant donor activation energy increases monotonically from less than 10meV for mole fractions less than 0.23 to a maximum of 160meV at a mole fraction of 0.48[52].

Another difficulty with AlGaAs is that at low temperatures, carriers may be photo-excited from deep states in the material and increase the free carrier concentration. The excited carriers are persistent because there is a large energy barrier to recapture of 0.2 eV[50] or 0.3 eV[97], because of the large lattice relaxation which occurs. This persistent photoconductivity (PPC) effect causes instability in heterojunction devices, especially at low temperatures. The excited electrons may transfer into the GaAs which may increase the 2DEG concentration[24] by a factor of three above the dark value, or may set up a parallel conduction path in the AlGaAs. Under applied electric fields in the dark, 2DEG electrons may gain sufficient energy to tunnel into deep states in the AlGaAs, which results in collapse of the I-V characteristics of the device[97, 100].

The deep state responsible for persistent photoconductivity (PPC centre) is thought to be caused by a donor-vacancy (DX) complex such as a silicon atom at a gallium site together with an arsenic vacancy[101], although other centres may also play a role[102]. The concentration of PPC centres is proportional to the donor density and the two may be of the same order of magnitude[52, 101]. The PPC centre concentration is negligible in GaAs and AlAs and has a maximum for a mole fraction of approximately 0.32[52]. The concentration of centres decreases as the MBE growth temperature of the AlGaAs is increased to the optimum of 700°C[29, 97].

Deep states in AlGaAs have also been shown to be responsible for generation-recombination noise in HEMTs[103, 104]. The deep states in the AlGaAs can either trap hot electrons directly or act as recombination centres. The resultant noise is caused by the random fluctuations in the number of free carriers.

A replacement for bulk AlGaAs with a low donor ionization energy and low density of deep states is required to enhance HEMT performance. A promising replacement is a superlattice consisting of thin layers of AlAs and doped-GaAs[96]. Both HEMTs

and IHEMTs have been fabricated using this superlattice[105, 106, 107] and PPC and carrier freeze-out effects are virtually eliminated. The AlAs layers are made thin (approximately 2 nm), so that the electrons are not confined to their parent n-GaAs layers and are free to move through the superlattice and into the adjacent bulk GaAs to form a 2DEG, in the same manner as HEMTs with bulk AlGaAs. The bandgap of the superlattice is a function of both the n-GaAs and AlAs layer widths. Since the density of deep states is greatly reduced, there is nearly complete donor ionization even at 77K. These superlattices appear likely to become important in the future since low donor ionization energies (less than 10meV) can be obtained together with large equivalent bandgaps. Thus large conduction band-edge discontinuities to limit real space transfer can be obtained, without incurring the limitations of bulk AlGaAs. The IHEMTs fabricated at Hirst Research Centre and discussed in Section 6 incorporate a superlattice of this type.

CONCLUSION

The High Electron Mobility Transistor (HEMT) is reminiscent of an nMOSFET in that the electrons which constitute the conducting path, or channel, reside as a thin sheet at the interface between two dissimilar materials. Unlike the SiO_2/Si insulator-semiconductor interface, however, which forms the basis of the MOSFET, the heterojunction interface upon which the HEMT is based consists of two lattice matched compound semiconductors, with the result that extremely high electron mobility, relative to the MOSFET, is obtained. Device and circuit engineers exploit this property to achieve higher gain, lower noise, and higher switching speed and/or frequency response transistors and integrated circuits than any other FET technology can offer. To attest to this fact, one needs only to look at the fifth generation supercomputer currently being developed by a major Japanese corporation: it uses entirely HEMT technology.

So far as analogue applications are concerned, HEMT may well in time displace the GaAs MESFET as the active device in monolithic microwave integrated circuits (MMICs): it has already done so in some discrete applications. Some idea of the potential of high electron mobility transistors can be gained from two examples. First, in the field of broadband travelling wave MMIC amplifiers, the Hirst Research Centre's DC-12 GHz amplifier using GaAs MESFETs delivers 6 dB of gain over the band: HEMTs would give more like 12 dB gain from the same chip. Secondly, it has now been shown that cooled low-noise amplifiers using discrete HEMTs are superior to GaAs MESFET amplifiers in radio astronomy telescope receivers (and it is only a few years since the MESFET in its turn, displaced the parametric amplifier). Furthermore, initial studies indicate that HEMT is a leading contender in the quest to realize a transistor for operation at around 100 GHz, which is a most exciting possibility.

The literature on HEMT is extensive, and expanding all the time. This paper has attempted to review the salient features of the physics and technology involved, and to provide a reference resource on the subject; given the pace of developments, it will inevitably and quickly become out of date.

ACKNOWLEDGEMENT

A. J. Hill wishes to acknowledge the financial support of the University of Natal, Durban, South Africa.

REFERENCES

1 DRUMMOND, T. J., MORKOC, H. and CHO, A. Y., 'Molecular beam epitaxy growth of (Al,Ga)As/GaAs heterostructures', J. Crystal Growth, *56*, p. 449–454, 1982.

2 GOSSARD, A. C., PETROFF, P. M., WIEGMANN, W., DINGLE, R. and SAVAGE, A., 'Epitaxial structures with alternate-atomic-layer composition modulation', Appl. Phys. Lett., *29*, 6, p. 323–325, September 1976.

3 MILLER, D. L., 'MBE materials considerations for HEMT circuits', Proc. IEEE GaAs IC Symposium, Boston, p. 37–40, October 1984.

4 BHATTACHARYA, P. K., DAS, U. and LUDOWISE, M. J., 'Transport properties of n-type metalorganic chemical-vapor-deposited $Al_xGa_{1-x}As$ ($0 < x < 0.6$)', Phys. Rev. B., *29*, 12, p. 6623–6631, June 1984.

5 DINGLE, R., STORMER, H., GOSSARD, A. C. and WEIGMANN, W., 'Electron mobilities in modulation doped semiconductor heterojunction superlattices', Appl. Phys. Lett., *33*, 7, p. 665–667, 1978.

6 DRUMMOND, T. J., MORKOC, H. and CHO, A. Y., 'Dependence of electron mobility on spatial separation of electrons and donors in $Al_xGa_{1-x}As$/GaAs heterostructures', J. Appl. Phys., **52**, 3, p. 1380–1386, March 1981.

7 STORMER, H. L., DINGLE, R., GOSSARD, A. C., WIEGMANN, W. and STURGE, M. D., 'Two-dimensional electron gas at a semiconductor-semiconductor interface', Solid State Comm., *29*, p. 705–709, 1979.

8 WITKOWSKI, L. C., DRUMMOND, T. J., STANCHAK, C. M. and MORKOC, H., 'High mobilities in $Al_xGa_{1-x}As$-GaAs heterojunctions', Appl. Phys. Lett., *37*, 11, p. 1033–1035, December 1980.

9 DINGLE, R., WIEGMANN, W., and HENRY, C. H., 'Quantum states of confined carriers in very thin $Al_xGa_{1-x}As$-GaAs-$Al_xGa_{1-x}As$ heterostructures', Phys. Rev. Lett., *33*, 14, p. 827–830, 30 September 1984.

10 WALDROP, J. R., KOWALCZYK, S. P., GRANT, R. W., KRAUT, E. A. and MILLER, D. L., 'XPS measurement of GaAs-AlAs heterojunction band discontinuities: Growth sequence dependence', J. Vac. Sci. Technol., *19*, 3, p. 573–575, September 1981.

11 MILLER, R. C., KLEINMAN, D. A. and GOSSARD, A. C., 'Energy-gap discontinuities and effective masses for GaAs-$Al_xGa_{1-x}As$ quantum wells', Phys. Rev. B., *29*, 12, p. 7085–7087, June 1984.

12 BATEY, J., WRIGHT, S. L. and DIMARIA, D. J., 'Energy band-gap discontinuities in GaAs:(Al,Ga)As heterojunctions', J. Appl. Phys., *57*, 2, p. 484–487, January 1985.

13 KROEMER, H. 'Barrier control and measurements: Abrupt semiconductor heterojunctions', J. Vac. Sci. Technol. B. *2*, 3, p. 433–439, July 1984.

14 BAUER, R. S., ZURCHER, P. and SANG, H. W., 'Inequality of semiconductor heterojunction conduction-band-edge discontinuity and electron affinity difference', Appl. Phys. Lett., *43*, 7, p. 663–665, October 1983.

15 KROEMER, H. 'Critique of two recent theories of heterojunction line-ups', IEEE Electron Devices., EDL-*4*, 2, p. 25–26, February 1983.

16 NUSSBAUM, A., 'Heterojunction discontinuities: The current position', IEEE Electron Device Letters, EDL-5, 11, p. 499–501, November 1984.

17 HARRISON, W. A., 'Elementary theory of heterojunctions', J. Vac. Sci. Technol., *14*, 4, p. 1016–1021, July 1977.

18 WALUKIEWICZ, W., LAGOWSKI, J. and GATOS, H. C., 'Electron mobility in n-type GaAs at 77K; determination of the compensation ratio', J. Appl. Phys., *53*, 1, p. 769–770, January 1982.

19 HIYAMIZU, S., SAITO, J., NANBU, K. and ISHIKAWA, T., 'Improved electron mobility higher than 10^6 cm^2/Vs in selectively doped GaAs/N-AlGaAs heterostructures grown by MBE', Japan J. Appl. Phys., *22*, 1, p. L609–611, October 1983.

20 HIYAMIZU, S., FUJII, T., MIMURA, T., NANBU, K., SAITO, J. and HASHIMOTO, H., 'The effect of growth temperature on the mobility of two-dimensional electron gas in selectively doped GaAs/N-AlGaAs heterostructures grown by MBE', Japan J. Appl. Phys., *20*, 6, p. L455–458, June 1981.

21 DRUMMOND, T. J., KOPP, W., MORKOC, H., HESS, K., CHO, A. Y. and STREETMAN, B. G. 'Effect of background doping on the electron mobility of (Al,Ga)As/GaAs heterostructures', J. Appl. Phys., *52*, 9, p. 5689–5690, September 1981.

22 DRUMMOND, T. J., MORKOC, H., HESS, K. and CHO, A. Y., 'Experimental and theoretical electron mobility of modulation doped $Al_xGa_{1-x}As$/GaAs heterostructures grown by molecular beam epitaxy', J. Appl. Phys., *52*, 8, p. 5231–5234, August 1981.

23 MENDEZ, E. E., PRICE, P. J. and HEIBLUM, M. 'Temperature dependence of the electron mobility in GaAs-GaAlAs heterostructures', Appl. Phys. Lett., *45*, 3, p. 294–296, August 1984.

24 DRUMMOND, T. J., KOPP, W., FISCHER, R., MORKOC, H., THORNE, R. E. and CHO, A. Y., 'Photoconductivity effects in extremely high mobility modulation doped (Al,Ga)As/GaAs heterostructures', J. Appl. Phys., *53*, 2, p. 1238–1240, 1982.

25 MORI, S. and ANDO, T., 'Electronic properties of a semiconductor superlattice II. Low temperature mobility perpendicular to the superlattice', J. Phys. Soc. Japan, *48*, 3, p. 865–873, March 1980.

26 KEEVER, M., KOPP, W., DRUMMOND, T. J., MORKOC, H. and HESS, K., 'Current transport in modulation-doped $Al_xGa_{1-x}As$/GaAs heterojunction structures at moderate field strengths', Japan J. Appl. Phys., *21*, 10, p. 1489–1495, October 1982.

27 TAKEDA, Y., KAMEI, H. and SASAKI, A., 'Mobility calculation of two-dimensional electron gas in GaAs/AlGaAs heterostructure at 4.2K', Electronics Letters, *18*, 7, p. 309–311, April 1982.

28 HWANG, J. C. M., KASTALSKY, A., STORMER, H. L. and KERAMIDAS, V. G., 'Transport properties of selectively doped GaAs-(AlGa)As heterostructures grown by molecular beam epitaxy', Appl. Phys. Lett., *44*, 8, p. 802–804, April 1984.

29 KLEM, J., MASSELINK, W. T., ARNOLD, D., FISCHER, R., DRUMMOND, T. J., MORKOC, H., LEE, K. and SHUR, M. S., 'Persistent photoconductivity in (Al,Ga)As/GaAs modulation doped structures: dependence on structure and growth temperature', J. Appl. Phys., *54*, 9, p. 5214–5217, September 1983.

30 ANDO, T., 'Self-consistent results for a GaAs/$Al_xGa_{1-x}As$ heterojunction. II. Low temperature mobility', J. Phys. Soc. Japan, *51*, 12, p. 3900–3907, December 1982.

31 WANG, W. I., WOOD, C. E. C. and EASTMAN, L. F., 'Extremely high electron mobilities in modulation-doped GaAs-$Al_xGa_{1-x}As$ heterojunction superlattices', *17*, 1, p. 36–37, January 1981.

32 STORMER, H. L., GOSSARD, A. C., WIEGMANN, W. and BALDWIN, K. 'Dependence of electron mobility in modulation-doped GaAs-(AlGa)As heterojunction interfaces on electron density and Al concentration', Appl. Phys. Lett., *39*, 11, p. 912–914, December 1981.

33 LEE, K., SHUR, M. S., DRUMMOND, T. J. and MORKOC, H. 'Low field mobility of 2-D electron gas in modulation doped $Al_xGa_{1-x}As$/GaAs layers', J. Appl. Phys., *54*, 11, p. 6432–6438, November 1983.

34 PRICE, P. J., 'Two-dimensional electron transport in semiconductor layers II: Screening', J. Vac. Sci. Technol., *19*, 3, p. 599–603, September 1981.

35 MORKOC, H. DRUMMOND, T. J., THORNE, R. E. and KOPP, W. 'Mobility enhancement in inverted $Al_xGa_{1-x}As$/GaAs modulation doped structures and its dependence on donor-electron separation', Japan J. App. Phys., *20*, 12, p. L913–916, December 1981.

36 TSUI, D. C., GOSSARD, A. C., KAMINISKY, G. and WIEGMANN, W. 'Transport properties of GaAs-$Al_xGa_{1-x}As$/GaAs heterojunction field-effect transistors', Appl. Phys. Lett., *39*, 9, p. 712–714, November 1981.

37 MIMURA, T. 'The present status of modulation-doped and insulated-gate field-effect transistors in III–V semiconductors', Surface Sci., *113*, p. 454–463, 1982.

38 ISHIKAWA, T., HIYAMIZU, S., MIMURA, T., SAITO, J. and HASHIMOTO, H. 'The effect of annealing on the electrical properties of selectively doped GaAs/N-AlGaAs heterojunction structures grown by MBE', Japan J. Appl. Phys., *20*, 11, p. L814–816, November 1981.

39 WALLIS, R. H., 'Effect of free carrier screening on the electron mobility of GaAs: A study by field-effect measurements', Physica 117B and 118B, North-Holland Publishing Co., p. 756–758, 1983.

40 SAITO, J., NANBU, K., ISHIKAWA, T. and HIYAMIZU, S. 'Dependence of the mobility and the concentration of two-dimensional electron gas in selectively doped GaAs/N-$Al_xGa_{1-x}As$ heterostructure on the AlAs mole fraction', Japan J. Appl. Phys., *22*, 2, p. L79–81, February 1983.

41 DRUMMOND, T. J., KOPP, W., FISCHER, R. and MORKOC, H. 'Influence of AlAs mole fraction on the electron mobility of (Al,Ga)As/GaAs heterostructures', J. Appl. Phys., *53*, 2, p. 1028–1029, February 1982.

42 INOUE, K., HIYAMIZU, S., INAYAMA, M. and INUISHI, Y. 'Analyses of 2D electron transport at a GaAs/AlGaAs interface', Proc. 14th. Solid State Devices, Tokyo 1982, in Jap. J. Appl. Phys. Suppl. 22-1, p. 357–363, 1983.

43 SHAH, J., PINCZUK, A., STORMER, H. L., GOSSARD, A. C. and WIEGMANN, W. 'Electric field induced heating of high mobility electrons in modulation-doped GaAs-AlGaAs heterostructures', Appl. Phys. Lett., *41*, 1, p. 55–57, January 1983.

44 MORKOC, H. 'Current transport in modulation doped (Al,Ga)As/GaAs heterostructures: Applications to high speed FET's', IEEE Electron Device Letters, EDL-2, 10, p. 260–262, October 1981.

45 TOMIZAWA, M., YOKOYAMA, K. and YOSHII, A. 'Hot-electron velocity characteristics at AlGaAs/GaAs heterostructures', IEEE Electron Device Letters, EDL-5, 11, p. 464–465, November 1984.

46 DRUMMOND, T. J., KOPP, W., MORKOC, H. and KEEVER, M., 'Transport in modulation doped ($Al_xGa_{1-x}As$/GaAs): Correlations with Monte Carlo calculations (GaAs)', Appl. Phys. Lett., 41, p. 277–279, 1982.

47 INOUE, M., INAYAMA, M., HIYAMIZU, S. and INUISHI, Y. 'Parallel electron transport and field effects of electron distributions in selectively-doped GaAs/n-AlGaAs', Japan J. Appl. Phys., *22*, 4, p. L213–215, April 1983.

48 HIRANO, M., TAKANASHI, Y. and SUGETA, T. 'Current-voltage characteristics of an AlGaAs/GaAs heterostructure FET for high gate voltages', IEEE Electron Device Letters, EDL-5, 11, p. 496–499, November 1984.

49 CARNEZ, B., CAPPY, A., KASZYNSKI, A., CONSTANT, E. and SALMER, G. 'Modelling of a submicrometre gate field-effect transistor including effects of nonstationary electron dynamics', J. Appl. Phys., *51*, 1, p. 784–790, 1980.

50 SOARES, R., GRAFFEUIL, J. and OBREGON, J. 'GaAs/AlGaAs two-dimensional electron gas FET: Microwave and high speed logic applications' from 'Applications of GaAs MESFETs', Artech House, p. 461–485, 1983.

51 EASTMAN, L. F. 'Use of molecular beam epitaxy in research and development of selected high speed compound devices', J. Vac. Sci. Technol., Bl. 2, p. 131–134, April–June 1983.

52 CHAND, N., HENDERSON, T. J., KLEM, J., MASSELINK, W. T., FISCHER, R., CHANG, Y-C and MORKOC, H. 'Comprehensive analysis of Si-doped $Al_xGa_{1-x}As$ (x = 0 to 1): Theory and experiments', Phys. Rev. B., *30*, 8, p. 4481–4492, October 1984.

53 HESS, K., MORKOC, H., SCHICHIJO, H. and STREETMAN, B. G. 'Negative differential resistance through real-space transfer', Appl. Phys. Lett., *35*, 6, p. 469–471, September 1979.

54 HILL, G. and ROBSON, P. N., 'Hot electron transport in the $Ga_{1-x}Al_xAs$ system', Proc. 3rd Int. Conf. on Hot Carriers in Semiconductors', France, p. 335–342, July 1981.

55 CIRILLO, N. C., ABROKAWAH, J. K. and SHUR, M. S., 'Self-aligned modulation-doped (Al,Ga)As/GaAs field-effect transistors', IEEE Electron Device Letters, EDL-*5*, 4, p. 129–131, April 1984.

56 CHERN, J. G. J. and OLDHAM, W. G., 'Determining specific contact resistivity from contact end resistance measurements', IEEE Electron Device Letters, EDL-*5*, 5, p. 178–180, May 1984.

57 ISHIKAWA, T., HIYAMIZU, S., MIMURA, T., SAITO, J. and HASHIMOTO, H., 'The effect of annealing on the electrical properties of selectively doped GaAs/N-AlGaAs heterojunction structures grown by MBE', Japan J. Appl. Phys., *20*, 11, p. L814–816, November 1981.

58 LEE, H. WICKS, G. and EASTMAN, L. F., 'High temperature annealing of modulation doped GaAs/AlGaAs heterostructures for FET applications', Proc. IEEE/Cornell Conference, p. 204–208, August 1984.

59 PEARAH, P., HENDERSON, T., KLEM, J., MORKOC, H., NILSSON, B., WU, O., SWANSON, A. W. and CH'EN, D. R., 'Rapid thermal annealing of modulation-doped $Al_xGa_{1-x}As$/GaAs heterostructures for device applications', J. Appl. Phys., *46*, 6, p. 1851–1855, September 1984.

60 SU, S. L., FISCHER, R., LYONS, W. G., THORNE, R. E., KOPP, W. and MORKOC, H., 'Modulation-doped (Al,Ga)As/GaAs FETs with high transconductance and electron velocity', Electronics Letters, *18*, 18, p. 794–796, September 1982.

61 TAKANASHI, Y., HIRANO, M. and SUGETA, T., 'Control of threshold voltage of AlGaAs/GaAs 2DEG FET's through heat treatment', IEEE Electron Device Letters, EDL-*5*, 7, p. 241–243, July 1984.

62 DELAGEBEAUDEUF, D., LAVIRON, M., DELESCLUSE, P., TUNG, P. N., CHAPLART, J. and LINH, N. T., 'Planar enhancement mode two-dimensional electron gas FET associated with a low AlGaAs surface potential', Electronic Letters, *18*, 2, p. 103–105, January 1982.

63 TUNG, P. N., DELAGEBEAUDEUF, D., LAVIRON, M., DELESCLUSE, P., CHAPLART, J. and LINH, N. T., 'High-speed two-dimensional electron-gas FET logic', Electronics Letters, *18*, 2, p. 109–110, January 1982.

64 TSUI, D. C., GOSSARD, A. C. and DOLAN, G. J., 'Radiation effects on modulation-doped GaAs-$Al_xGa_{1-x}As$ heterostructures', Appl. Phys. Lett., *42*, 2, p. 180–182, January 1983.

65 FISCHER, R., HENDERSON, T., KLEM, J., MASSELINK, W. T., KOPP, W. and MORKOC, H. 'Characteristics of GaAs/AlGaAs MODFETs grown directly on (100) silicon', Electronics Letters, *20*, 22, p. 945–947, October 1984.

66 FISCHER, R., KLEM, J., HENDERSON, T., MASSELINK, W. T., KOPP, W. and MORKOC, H., 'GaAs/AlGaAs MODFET's grown on (100) Ge', IEEE Electron Device Letters, EDL-*5*, 11, p. 456–457, November 1984.

67 PIERRET, R. F. and LUNDSTROM, M. S., 'Correspondence between MOS and modulation-doped structures', IEEE Trans. Elec. Dev., ED-*31*, 3, p. 383–385, March 1984.

68 DRUMMOND, T. J., SU, S. L., KOPP, W., FISCHER, R., THORNE, R. E., MORKOC, H., LEE, K. and SHUR, M. S., 'High velocity N-on and N-off modulation doped GaAs/$Al_xGa_{1-x}As$ FETs', Int. Elec. Dev. Meet., 1982, p. 586–589.

69 SOLOMON, P. M. and MORKOC, H., 'Modulation-doped GaAs/AlGaAs heterojunction field-effect transistors (MODFETs) Ultrahigh-speed device for supercomputers', IEEE Trans. Elec. Dev., ED-*31*, 8, p. 1015–1027, August 1984.

70 MIMURA, T., JOSHIN, K. and KURODA, S., 'Device modelling of HEMTs', Fujitsu Sci. Tech. Jnl., p. 243–279, September 1983.

71 HIRAKAWA, K., SAKAKI, H. and YOSHINO, J., 'Concentration of electrons in selectively doped GaAlAs/GaAs heterojunction and its dependence on spacerlayer thickness and gate electric field', Appl. Phys. Lett., *45*, 3, p. 253–255, August 1984.

72 VINTER, B., 'Sub-bands and charge control in a two-dimensional gas field-effect transistor', Appl. Phys. Lett., *44*, 3, p. 307–309, February 1984.

73 LEE, S. J., CROWELL, C. R., and LEE, C. P., 'Optimization of HEMTs in ultra high speed GaAs integrated circuits', IEEE Int. Elec. Dev. Meet., p. 103–106, 1983.

74 LEE, K., SPUR, M., DRUMMOND, T. J. and MORKOC, H., 'A unified method for characterizing (Al,Ga)As/GaAs MODFETs including parasitic MESFET conduction in the (Al,Ga)As', Proc. IEEE Cornell Conf. p. 177–186, August 1983.

75 DAMBKES, H. and HEIME, K., 'Field-effect transistors with a two-dimensional electron gas as channel' from 'Two-dimensional systems, Heterostructures and Superlattices', Eds. Bauer, G., Kuchar, F. and Heinrich, H., Springer Series in Solid-State Sciences 53, Springer-Verlag, 1984.

76 LEE, K., SHUR, M. S., DRUMMOND, T. J. and MORKOC, H., 'Current-voltage and capacitance-voltage characteristics of modulation-doped field-effect transistors', IEEE Trans. Elec. Dev., ED-*30*, 3, p. 207–212, March 1983.

77 LEE, K., SHUR, M. S., DRUMMOND, T. J. and MORKOC, H., 'Electron density of the two-dimensional electron gas in modulation doped layers', J. Appl. Phys., *54*, p. 2093–2096, 1983.

78 LUNDSTROM, M. S. and SCHUELKE, R. J., 'Numerical analysis of heterostructure semiconductor devices', IEEE Trans. Elec. Dev., ED-*30*, 9, p. 1151–1159, September 1983.

79 DRUMMOND, T. J., FISCHER, R., SU, S. L., LYONS, W. G. and MORKOC, H., 'Characteristics of modulation-doped $Al_xGa_{1-x}Al$/GaAs field-effect transistors: Effect of donor-electron separation', Appl. Phys. Lett., *42*, 3, p. 262–264, February 1983.

80 BERENZ, J. J., NAKANO, K. and WELLER, K. P., 'Low noise high electron mobility transistors', IEEE Monolithic Circuits Symposium, Digest of papers, p. 83–86, May 1984.

81 LEE, K., SHUR, M. S., DRUMMOND, T. J., SU, S. L., LYONS, W. G., FISCHER, R. and MORKOC, H., 'Design and fabrication of high transconductance modulation-doped (Al,Ga)As/GaAs FETs', J. Vac. Sci. Technol., JVST B1, p. 186–189, 1983.

82 TASKER, P. J., CAMNITZ, L. H., LUNARDI, L. M., LEE, H., MAKI, P. A., ENQUIST, P. and EASTMAN, L. F., 'Progress toward performance limits for MODFET and HBT devices', Proc. GaAs Conf. Biarritz, France, 1984.

83 PEI, S. S., SHAH, N. J., HENDEL, R. H., TU, C. W. and DINGLE, R., 'Ultra high speed integrated circuits with selectively doped heterostructure transistors', Proc. 1984 IEEE GaAs IC Symposium, p. 129–132.

84 KIEHL, R. A., FEUER, M. D., HENDEL, R. H., HWANG, J. C. M., KERAMIDAS, V. G., ALLYN, C. L. and DINGLE, R., 'Selectively doped heterostructure frequency dividers', IEEE Electron Device Letters, EDL-*4*, 10, p. 377–379, October 1983.

85 STORMER, H. L., GOSSARD, A. C., WIEGMANN, W., BLONDEL, R. and BALDWIN, K., 'Temperature dependence of the mobility of two-dimensional hole systems in modulation-doped GaAs-(AlGa)As', Appl. Phys. Lett., *44*, 1, p. 139–141, January 1984.

86 TIWARI, S. and WANG, W. I., 'p-Channel MODFET's using GaAlAs/GaAs two-dimensional hole gas', IEEE Electron Device Letters, EDL-*5*, 8, p. 333–335, August 1984.

87 TSUI, D. C. and LOGAN, R. A., 'Observation of two-dimensional electrons in LPE-grown GaAs-$Al_xGa_{1-x}As$ heterojunctions', Appl. Phys. Lett., *35*, 2, p. 99–101, July 1979.

88 SASA, S., SAITO, J., NANBU, K., ISHIKAWA, T. and HIYAMIZU, S., 'Improved 2DEG mobility in inverted GaAs/n-AlGaAs heterostructures grown by MBE', Japan J. Appl. Phys., *23*, 8, p. L573–575, August 1984.

89 FISCHER, R., MASSELINK, W. T., SUN, Y. L., DRUMMOND, T. J., CHANG, Y. C., KLEIN, M. V., MORKOC, H. and ANDERSON, E., 'Improvement of the inverted GaAs/AlGaAs hetero-interfaces', J. Vac. Sci. Technol., B2, 2, p. 170–174, April 1984.

90 INOUE, K. and SAKAKI, H., 'A new highly-conductive (AlGa)As/GaAs/(AlGa)As selectively doped double-heterojunction field-effect transistor (SD-DH-FET)', Japan J. Appl. Phys., *23*, 2, p. L61–63, February 1984.

91 DRUMMOND, T. J., KLEM, J., ARNOLD, D., FISCHER, R., THORNE, R. E., LYONS, W. G. and MORKOC, H., 'Use of a superlattice to enhance the interface properties between two bulk heterolayers', Appl. Phys. Lett., *42*, 7, p. 615–617, April 1983.

92 LEE, K., SHUR, M., DRUMMOND, T. J. and MORKOC, H., 'Charge control model of inverted GaAs-AlGaAs modulation doped FET's, (IMODFET's), J. Vac. Sci. Technol., B2, 2, p. 113–116, April 1984.

93 THORNE, R. E., FISCHER, R., SU, S. L., KOPP, W., DRUMMOND, T. J. and MORKOC, H., 'Performance of inverted structure modulation-doped Schottky barrier field effect transistors', Japan J. Appl. Phys. *21*, 4, p. L223–224, April 1982.

94 LADBROOKE, P. H., 'On the large signal characteristics of modulation-doped transistors at room temperature', Unpublished work, Cornell University 1982.

95 HILL, A. J. and LADBROOKE, P. H., 'Inverted HEMT, prospects and uncertainty', IEE Colloquium "IC's above 1 GHz", p. 5.1–5.4, March 1985.

96 BABA, T., MIZUTANI, T. and OGAWA, M., 'Elimination of persistent photoconductivity and improvement in Si activation coefficient by spacial separation from Ga and Si in Al-Ga-As:Si solid system', Japan J. Appl. Phys., *22*, 10, p. L627–629, October 1983.

97 DRUMMOND, T. J., FISCHER, R. J., KOPP, W. F., MORKOC, H., LEE, K. and SHUR, M. S., 'Bias dependence and light sensitivity of (Al,Ga)As/GaAs MODFET's at 77K', IEEE Trans. Elec. Dev., ED-*30*, 12, p. 1806–1811, December 1983.

98 CHI, J. Y., HOLMSTROM, R. P. and SALERNO, J. P., 'Effect of traps on low-temperature high electron mobility transistor characteristics', IEEE Electron Device Letters, EDL-*5*, 9, p. 381–384, September 1984.

99 CHI, J. Y., HOLSTROM, R. P. and SALERNO, J. P., 'AlGaAs defect characterization in high electron mobility transistors by thermally stimulated drain conductance', IEEE Electron Device Letters, EDL-*5*, 11, p. 476–478, November 1984.

100 FISCHER, R., DRUMMOND, T. J., KLEM, J., KOPP, W., HENDERSON, T. S., PERRACHIONE D. and MORKOC, H., 'On the collapse of drain I–V characteristics in modulation-doped FET's at cryogenic temperatures', IEEE Trans. Elec. Dev., ED-*31*, 8, p. 1028–1032, August 1984.

101 YAMANAKA, K., NARITSUKA, S., MANNOH, M., YUASA, T., NOMURA, Y., MIHARA, M., and ISHII, M., 'Influence of growth conditions and alloy composition on deep electron traps of n-$Al_xGa_{1-x}As$ grown by MBE', J. Vac. Sci. Technol., B2, 2, p. 229–232, April 1984.

102 NATHAN, M. I., HEIBLUM, M., KLEM, J. and MORKOC, H., 'Persistent photoconductivity in AlGaAs-GaAs heterostructures', J. Vac. Sci. Technol., B2, 2, p. 167–169, April–June 1984.

103 LORECK, K., DAMBKES, H., HEIME, K. and PLOOG, K., 'Deep level analysis in (AlGa)As-GaAs MODFETs by means of low-frequency noise measurements', IEEE Int. Elec. Dev. Meet., p. 107–110, 1983.

104 LORECK, L., DAMBKES, H., HEIME, K., PLOOG, K., and WEIMANN, G., 'Deep level analysis in (AlGa)As-GaAs 2D electron gas devices by means of low-frequency noise measurements', IEEE Electron Device Letters, EDL-*5*, 1, p. 9–11, January 1984.

105 FISCHER, R., MASSELINK, W. T., KLEM, J., HENDERSON, T. and MORKOC, H., 'Elimination of drain I/V collapse in MODFETs through the use of thin n-GaAs/AlGaAs superlattice', Electronics Letters, *20*, 18, p. 743–744, August 1984.

106 BABA, T., MIZUTANI, T., OGAWA, M. and OHATA, K., 'High performance (AlAs/n-GaAs superlattice)/GaAs 2DEGFETs with stabilized threshold voltage', Japan J. Appl. Phys., Vol 23, No 8, p. L654–656, August 1984, and Proc. 42nd Dev. Res. Conf., Paper IIA-1, June 1984.

107 ARNOLD, D., HENDERSON, T. J., KLEM, J., FISCHER, R., KOPP, W., KETTERSON, A., MASSELINK, W. T. and MORKOC H., 'High performance inverted and large current double interface modulation-doped field-effect transistors with the bulk (Al,Ga)As replaced by a superlattice at the inverted interface', Appl. Phys. Lett., *45*, 8, p. 902–904, October 1984.

APPENDIX B

APPENDIX B

B1 FIELD SOLUTION IN REGION II OF INTRINSIC DEVICE (Section 3.2.4)

Poisson's equation gives (equation 3.22a)

$$\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} = \frac{\rho}{\varepsilon} \qquad (B.1)$$

where F_x and F_y are the longitudinal and transverse (from 2DEG towards gate) electric field components respectively. The volume charge density in the charge sheet is obtained from the sheet charge as:

$$\rho = \frac{q\ (n_{so} - n_{sd})}{Y_a} \qquad (B.2)$$

It is not possible to find a complete analytic solution for the field components in the charge sheet and gate-sheet area. To find a solution, therefore, the assumption is made that the longitudinal field along the charge sheet is a function of x-co-ordinate only:

$$\frac{\partial F_x}{\partial x} = \frac{\rho}{\varepsilon} f(x) \qquad (B.3)$$

The validity of this approach can be seen from the Fy-dependence From (B.1) and (B.3):

$$\frac{\partial F_y}{\partial y} = \frac{\rho}{\varepsilon} (1-f(x)) \qquad (B.4)$$

and integrating over the thickness of the charge sheet, Ya, yields

$$F_y(x) = \frac{\rho}{\varepsilon} (1-f(x))\ Y_a \qquad (B.5a)$$

which from (B.2), is equivalent to

$$F_y(x) = \frac{q}{\varepsilon} (n_{so}-n_{sd})\ [1-f(x)] \qquad (B.5b)$$

This equation simply indicates that the transverse field component arising from the sheet charge $q(n_{so}-n_{sd})$ is a function of horizontal distance along the charge sheet. Equation (B.5) can be rewritten as:

$$F_y\ (x) = \left(\frac{\rho}{\varepsilon} - \frac{\partial F_x}{\partial x}\right) Y_a \qquad (B.6)$$

Equation B.6 indicates that the transverse field, F_y, arising from an element of volume charge is determined by the charge within the element and the charge in longitudinal field through the element. This statement is simply a form of Gauss's Law.

Integrating F_y over the sheet charge to gate electrode spacing, d, must yield the gate potential. This is one boundary condition for the fields in region II. F_y remains constant between the sheet charge and gate because no charge exists in this region. A further boundary condition is that the 2DEG potential at the source end of region II is equal to V_l.

From (B.3), the dependences of longitudinal field and potential along the charge sheet are given by:

$$F_x(x) = \frac{\rho}{\varepsilon} \int f(x)\, dx \qquad (B.7)$$

and $$V(x) = \frac{\rho}{\varepsilon} \iint f(x)\, dx^2 \qquad (B.8)$$

Hence by Kirchoff's voltage law:

$$V_G + V_1 + \frac{\rho}{\varepsilon} \iint f(x)\, dx^2 = - \frac{\rho}{\varepsilon} (1-f(x))\, Y_a\, d \qquad (B.9)$$

Rearranging and differentiating twice yields:

$$\frac{d^2}{dx^2} f(x) - (Y_a.d)^{-1} f(x) = 0 \qquad (B.10)$$

This differential equation has the standard solution [B.1]:

$$f(x) = A \cosh k_x + B \sinh kx \qquad (B.11a)$$

where $$k = (\sqrt{Y_a.d})^{-1} \qquad (B.11b)$$

At x = 0, the boundary between regions I and II, the field must be entirely one-dimensional as in region I. From (B.5), f(0) must equal zero and hence A in (B.11a) also equals zero. Equation (B.7) can now be integrated and applying the boundary condition $F = -F_s$ at the interface with region I, F_x is found:

$$F_x(x) = - F_s \cosh (xk) \qquad (B.12)$$

which is equation (3.23).

The negative sign in (B.12) occurs because the field arising from the positive charge is in the -x direction, that is, towards the source.

B.2 FIELD SOLUTION IN REGION III (Section 3.2.5)

The solution in region III is simpler than in region II because of the assumption of a uniform distribution of surface charge which means that F_y is independent of x. The solution method follows the same approach as in region II but f(x) is found to be constant.

The y-directed field is given by (equation 3.27)

$$F_y = \frac{-qn_{ss}}{\varepsilon} \tag{B.13}$$

Using (B.5) again, $F_y = \frac{\rho}{\varepsilon} (1-fx)) Y_a$ (B.14)

From (B.13) and (B.14), f(x) is obtained:

$$f(x) = 1 + \frac{qn_{ss}}{\rho Ya} \tag{B.15}$$

From (B.7), and performing the integration yields

$$F_x(x) = \frac{\rho}{\varepsilon} \left(1+ \frac{qN_{ss}}{\rho Ya}\right) x + \text{constant} \tag{B.16}$$

At x = 0, the field must equal the peak field in region II, F_p. The first term can be simplified using (B.2) to yield

$$F_x(x) = \frac{q}{\varepsilon} \left[N_D - \frac{n_{sd}}{Y_a} + \frac{n_{ss}}{Y_a}\right] x + F_p \tag{B.17}$$

which is equation (3.28).

B.3 EVALUATING INTEGRALS FOR SUBSTRATE CURRENT (Section 3.3)

The integrals to be evaluated are (equations 3.46 and 3.47)

$$N_{tot} = \int_{E}^{\infty} f(E)g(E)dE \tag{B.18}$$

and $$N_{sub} = \tfrac{1}{2}\int_{B}^{\infty} f(E)g(E)dE \tag{B.19}$$

where f(E) is the Fermi-Dirac distribution function (equation 3.37):

$$f(E) = [1+\exp((E-E_F)/kTe)]^{-1} \tag{B.20}$$

and g(E) is the density of available states function (3.42):

$$g(E) = 4\pi \sqrt{\left(\frac{2m_c}{h}\right)^3}\ \sqrt{E-E_c} \tag{B.21}$$

Evaluating (B.18)

Combining (B.20) and (B.21) in (B.18) yields:

$$N_{tot} = \frac{4\pi}{h^3}\sqrt{(2m_c)^3}\int_{E_c}^{\infty}\sqrt{(E-E_c)}\,[1+\exp((E-E_F)/kTe]^{-1}\,dE$$

Making the substitution $x = (E-E_c)/kTe$ and simplifying gives

$$N_{tot} = \frac{4\pi}{h^3}\sqrt{(2m_c kTe)^3}\int_{o}^{\infty} x^{\frac{1}{2}}\,[1+\exp(x-\frac{E_F-E_c}{kT})]^{-1}\,dx$$

The integral is the Fermi-Dirac integral of order 0.5, $F_{1/2}(m)$, [B.2] which can be approximated, with less than 0.53 percent error, by the expression [B.3]:

$$F_{1/2}(m) \quad \{3\sqrt{2}[b+m+(|m-b|^c+a^c)^{1/c}]^{1.5} + \frac{\sqrt{\pi}}{2}\exp(-m)\}^{-1}$$

where $a = \sqrt[c]{9.6}$

$b = 2.13$

$c = 2.4$

and $m = (E_F-E_c)/kTe$ (B.22)

Evaluating (B.19):

Equation (B.19) cannot be evaluated directly and hence the equivalent expression

$$N_{subs} = \tfrac{1}{2} \int_0^\infty g(E)f(E)dE - \tfrac{1}{2} \int_0^\infty g(E)f(E)dE \qquad (B.23)$$

is used. The first term is simply one half of equation B.18 and can be calculated in the same way using (B.22). The second term cannot be integrated directly and so is numerically integrated using Simpson's rule.

REFERENCES

B-1 Hayt W.H., "Engineering Electromagnetics", 3rd. Edition, McGraw-Hill, 1974.

B-2 Blakemore J.S., "Intrinsic density $n_i(T)$ in GaAs: Deduced from band gap and effective mass parameters and derived independently from Cr capture and emission coefficients", J. Appl. Phys., Vol. 53(1), Jan. 1982, pp. 520-532

B-3 Aymerich-Humet X., Serra-Mestres F. and Millan J., "An Analytical Approximation for the Fermi-Dirac Integral $F_{3/2}(\eta)$", Sol. St. Elec., Vol. 24, No. 10, 1981, pp. 981-982

APPENDIX C

The paper in this appendix was presented at the 1985 IEEE/ Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits, Cornell University, Ithaca, New York, 19-31 July 1985.

LARGE SIGNAL MODELLING AND PRACTICAL PERFORMANCE OF INVERTED HEMT

A.J. Hill*, P.H. Ladbrooke, S. Ransome and D. Westwood
GEC Hirst Research Centre, East Lane, Middlesex, England HA9 7PP
(* on leave from University of Natal, Durban, South Africa)

ABSTRACT

Inverted High Electron Mobility Transistors (IHEMTs) with an AlAs/n-GaAs superlattice have been fabricated and exhibit high extrinsic transconductance (180 mS/mm at 300K) and large drain-source breakdown voltage (15V for 0.5 micron device). These values are the highest reported for IHEMTs. An analytic large signal model has been developed for the IHEMT, and details of this model are presented.

EXPERIMENTAL RESULTS

We have fabricated Inverted High Electron Mobility Transistors (IHEMTs) incorporating an n-GaAs/AlAs superlattice in place of a bulk n-AlGaAs layer. The structure is similar to that reported by Baba et al [1]. Each superlattice period consists of 15 Å AlAs and 20 Å GaAs, of which the centre 10 Å was Si-doped to 10^{18} cm^{-3}. The spacer layer thickness was 85Å. The undoped GaAs layer is 500 Å thick and is capped with 300 Å of n^+GaAs. The layers were grown by MBE at 620 °C and at a growth rate of 0.1 micron/hr. One half micron gate length devices were fabricated using conventional optical lithography.

Results obtained with these devices as well as results from Hall measurements are summarised in Table 1. The extrinsic transconductance (180 mS/mm at 300K) and the drain-source breakdown voltage (15V) are the highest reported for IHEMT. The average electron transit velocity was calculated from the intrinsic transconductance [2] as 1.7×10^7 cm/s. This value is also the largest reported for IHEMT or comparable structures.

THEORETICAL MODEL FOR IHEMT

HEMT and IHEMT models reported in the literature to date, have concentrated on the accurate calculation of the 2DEG sheet carrier concentration, and in general, have used standard JFET equations to describe the I-V characteristics [3,4]. We have

CH 2173-3/85/0000-0128 $1.00 © 1985 IEEE

developed a specific analytic model for the IHEMT, based on fundamental principles. A simplified cross-section through an IHEMT is shown in figure 1. In this model, the channel of the intrinsic device is divided into three regions.

In region I, the electrons are assumed to have constant mobility, μ, and the drain-source current per unit gate width is calculated using the Schockley model:

$$I_{ds} = q.\mu.n_s(x).E(x) \qquad (1)$$

where q is the electronic charge and E(x) is the longitudinal field at point x. $n_s(x)$ is the 2DEG sheet concentration at x and is calculated using a simple capacitor approximation, which has been shown to be reasonably accurate for the IHEMT [3]:

$$n_s(x) = \frac{\varepsilon}{qY_G}(V_{GS}-V_T-V(x)) \qquad (2)$$

where Y_G is the thickness of the undoped GaAs layer, V_{GS} is the applied gate-source voltage, V_T is the device threshold voltage and V(x) is the channel potential at x. A two-piece velocity-field characteristic is assumed and the boundary between regions I and II is the point at which the longitudinal field is equal to the critical field for velocity saturation, E_s (2.4 kV/cm).

In region II, the 2DEG concentration is assumed constant, n_{sv}, determined using equation 2 with the channel potential at the boundary between the two regions. The electric field in this region has both an x- and y-directed component, and a two-dimensional solution to Poisson's equation is obtained. The 2DEG and underlying thin n-AlGaAs layer is modelled as a sheet of charge with density $q(N_DY_A-n_{sv})$ where N_D and Y_A are the donor density and thickness of the n-AlGaAs layer respectively. The longitudinal component of electric field at the 2DEG, as a function of x-coordinate is given by:

$$E_x(x) = E_s.\cosh((x-L_1)/Y) \qquad (3)$$

where

$$Y = \sqrt{Y_A.Y_G} \qquad (4)$$

and L_1 is the length of region I. Equation 3 is similar to that obtained by Grebene and Ghandhi [5] for the MESFET. The potential at the drain edge of the gate, V_2, is obtained by integrating (3) and adding the potential, V_1, at the boundary between regions I and II:

$$V_2 = V_1 + E_s.Y.\sinh((L-L_1)/Y) \qquad (5)$$

where L is the gate length. The drain current per unit gate width in region II is given by $q.n_{sv}.v_s$, where v_s is the saturation velocity. By equating the currents in regions I and II, and using equation 5, V_1 and L_1 can be determined.

In region III, the electric field also has longitudinal and transverse components. In order to partition the total field arising from the sheet charge, it is assumed that the y-directed component is determined by the occupied surface state density, n_{ss}, and is independent of x-coordinate:

$$E_y = \frac{q}{\varepsilon}.n_{ss} \qquad (6)$$

The 2DEG sheet concentration is assumed to remain constant at n_{sv}. The remainder of the field arising from the net positive sheet charge density is x-directed and decreases linearly with distance towards the drain.

Figure 2 shows the variation of longitudinal electric field along the channel of the device for three values of n_{ss}. Solutions are obtained by successive approximation for V_2, until the sum of V_2 and the voltage drop across region III is equal to the applied drain-source voltage. The field distribution may be used to calculate the ionisation integral and hence obtain the drain-source breakdown voltage.

Substrate Current

Electrons in the 2DEG may gain sufficient energy from the longitudinal field to be scattered over the energy barrier at the interface, into the AlGaAs. This real space transfer results in a component of drain current in parallel with the 2DEG current, which increases the saturated device drain conductance and alters the transconductance. Since the effect of the injected current in the AlGaAs is the same as that of substrate current in MESFETs, the term is retained for the AlGaAs component in IHEMTs.

The barrier height which limits electron transfer from the 2DEG to the AlGaAs layer is determined by the conduction band edge discontinuity, ΔEc, at the heterojunction. ΔEc is dependent on the aluminium mole fraction, x, in the $Al_xGa_{1-x}As$, but the exact dependence is not well established.

The sum of the conduction- and valence-band discontinuities is equal to the energy bandgap difference between GaAs and the larger-bandgap AlGaAs. The energies of the Γ and X conduction band minima in AlGaAs, measured from the top of the valence band (**K**=0,0,0) [6], as a function of aluminium mole fraction, are shown in figure 3. The L band is not shown because it has an energy between the Γ and X bands for all mole fractions and does not affect this discussion in any major way.

For mole fractions less than approximately 0.43, $Al_xGa_{1-x}As$ is a direct bandgap semiconductor and the conduction band discontinuity is simply a fraction of the energy gap difference between GaAs and AlGaAs. Assuming a value of approximately 0.4 [7] for the ratio of the valence band discontinuity (ΔEv) to direct energy gap difference (ΔEg), means that $\Delta Ec/\Delta Eg$ is approximately 0.6 for mole fractions less than 0.43.

At mole fractions greater than 0.43, AlGaAs has an indirect bandgap. The valence band discontinuity, however, has been reported to remain a fraction (approximately 0.4) of the direct (Γ) energy gap difference between GaAs and AlGaAs [8] ((1)in figure 3). If hot electrons in the GaAs 2DEG can only transfer to the Γ-valley in the AlGaAs, then the effective conduction band discontinuity is 0.6 times the direct energy gap difference and increases monotonically as x increases ((2) in figure 3). If however, the electrons in the GaAs can transfer to the X-valley in the AlGaAs, the effective conduction band discontinuity will decrease as x increases above 0.43 ((3) in figure 3), because the sum of conduction and valence band discontinuities must equal the total bandgap difference. There is no a priori reason to assume that the latter transition should not occur, which means that the maximum barrier height to hot electrons will occur at a mole fraction of approximately 0.43.

A simple energy conservation equation is used to calculate the electron temperature of the 2DEG electrons and hence, assuming a Maxwellian energy distribution, the number of electrons with energy greater than the barrier height. Fifty per cent of those electrons with sufficient energy are assumed to be scattered into the AlGaAs where they form the "substrate" current component of drain current. Figure 4 shows simulated I-V characteristics for a 0.5 micron IHEMT with an aluminium mole fraction of 0.3 and source resistance of 2 ohm.mm. Figure 5 shows the increase in drain conductance arising from the reduction of the mole fraction from 0.45 to 0.2, while maintaining the same gate bias.

SUMMARY

Electrical characteristics of fabricated IHEMTs have been presented. Details of an analytic model for the IHEMT, which includes the effects of hot electron injection from the 2DEG into the AlGaAs and channel length shortening, have also been presented.

ACKNOWLEDGEMENT

A.J. Hill wishes to acknowledge the financial support of the University of Natal, Durban, South Africa.

REFERENCES

[1] T. Baba, T. Mizutani, M. Ogawa and K. Ohata, "High Performance (AlAs/n-GaAs Superlattice)/GaAs 2DEGFETs with stabilized Threshold Voltage", Jpn. J. Appl. Phys., Vol. 23, No. 8, Aug. 1984, pp. L654-L656

[2] L. F. Eastman, private communication

[3] K. Lee, M.S. Shur, T.J. Drummond and H. Morkoc, "Charge control model of inverted GaAs-AlGaAs modulation doped FET's (IMODFET's)", J. Vac. Sci. Technol. B 2 (2), Apr. 1984, pp. 113-116

[4] Vinter B., "Subbands and charge control in a two-dimensional electron gas field-effect transistor", Appl. Phys. Lett., Vol. 44(3), Feb. 1984, pp. 307-309

[5] A.B Grebene and S.K. Ghandhi, "General Theory for Pinched Operation of the Junction-Gate FET", Solid-State Elec., Vol. 12, 1969, pp. 573-589

[6] P.K. Bhattacharya, U. Das and Ludowise M.J., "Transport properties of n-type metalorganic chemical-vapor-deposited $Al_xGa_{1-x}As$ $(0<x<0.6)$", Phys. Rev. B, Vol. 29, No. 12, pp. 6623-6631, 15 June 1984

[7] J. Batey, S.L. Wright and D.J. DiMaria, "Energy band-gap discontinuities in GaAs:(Al,Ga)As heterojunctions", J. Appl. Phys., Vol. 57, No. 2, pp. 484-487, Jan. 1985

[8] D. Arnold, A. Ketterson, T. Henderson, J. Klem and H. Morkoc, "Determination of the valence-band discontinuity between GaAs and (Al,Ga)As by the use of p^+GaAs-(Al,Ga)As-p^-GaAs capacitors", Appl. Phys. Lett., Vol. 45(11), pp. 1237-1239, Dec. 1984

Mobility	20 000	cm^2/Vs
2DEG Sheet Concentration	$7.5x10^{11}$	cm^{-2}
Extrinsic g_m	180	mS/mm
Intrinsic g_m	270	mS/mm
Source resistance	1.8	ohm.mm
Drain-source breakdown	15	V
Ave. electron velocity	$1.7x10^7$	cm/s
I_{dss}	230	mA/mm

Table 1: Results obtained on 0.5 micron IHEMTs incorporating an n-GaAs/AlAs superlattice

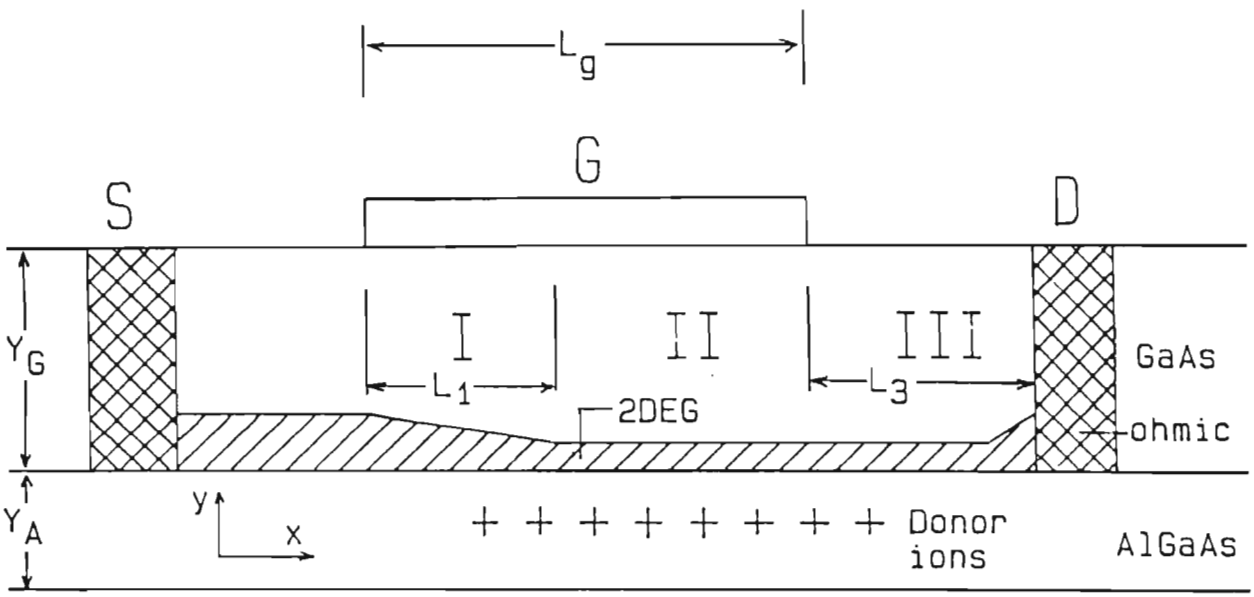


Figure 1: Simplified cross-section through an IHEMT showing the regions used in the model.

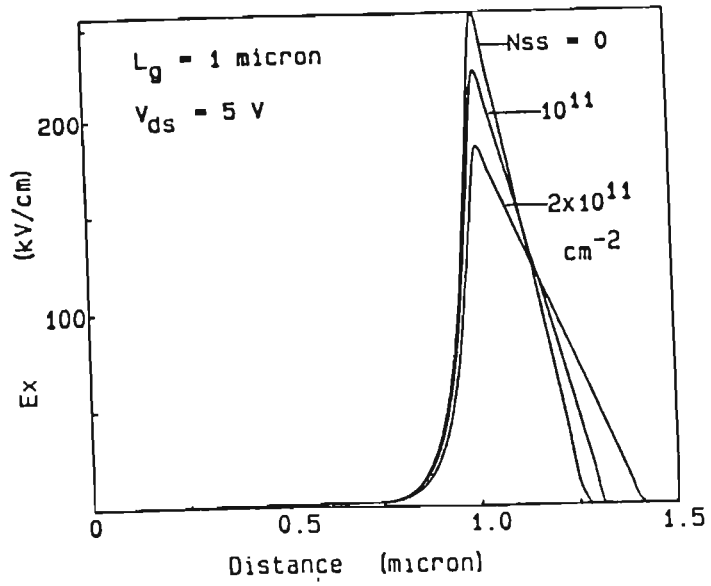


Figure 2: Typical variation of longitudinal electric field along the channel of a 1 micron IHEMT, for 3 values of occupied surface state density. Distance on the x-axis is from the source end of the gate, towards the drain.

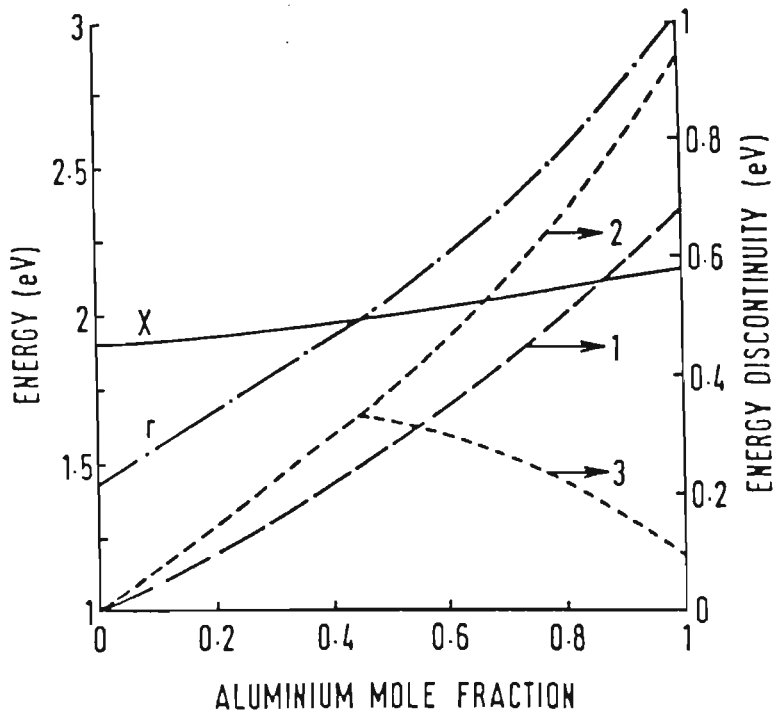


Figure 3: Conduction band edge energies (from [6]), and discontinuities as a function of aluminium mole fraction. (1) Valence band discontinuity. (2) and (3) are the conduction band discontinuities discussed in the text.

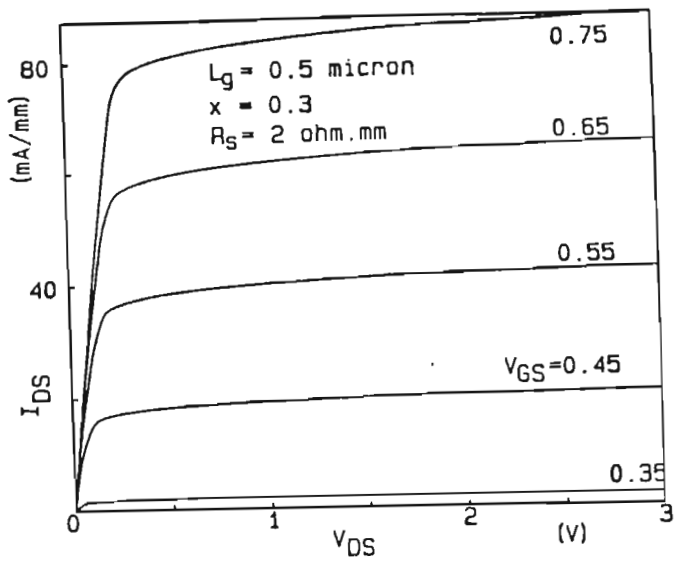


Figure 4: I-V characteristics for 0.5 micron enhancement mode IHEMT, V_T = +0.32 V.

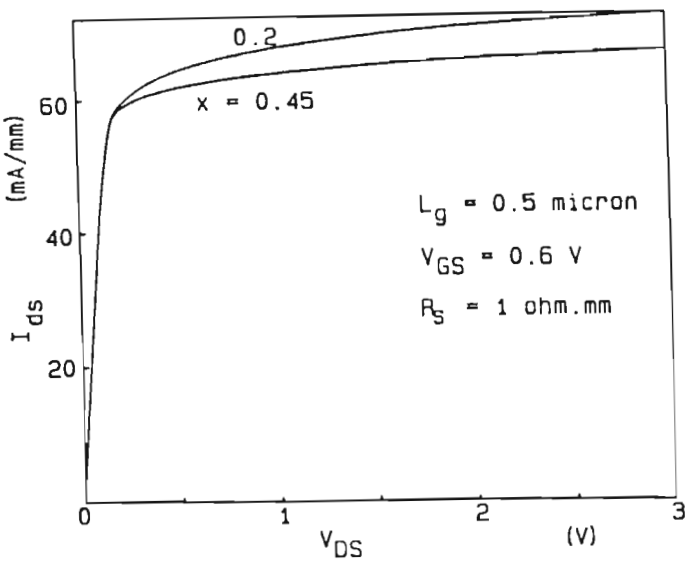


Figure 5: I-V characteristics at fixed gate bias showing the effect on drain conductance of changing mole fraction.

APPENDIX D

The paper in this appendix was presented at the Twelfth International Symposium on Gallium Arsenide and Related Compounds, Karuizawa, Japan, 23-26 Sept. 1985.

Inst. Phys. Conf. Ser. No. 79: Chapter 10
Paper presented at Int. Symp. GaAs and Related Compounds, Karuizawa, Japan, 1985

Large signal characteristics of inverted high electron mobility transistors

A.J. Hill*, P.H. Ladbrooke, S. Ransome and D. Westwood

GEC Hirst Research Centre, East Lane, Wembley, Middlesex, England, HA9 7PP
(* on leave from University of Natal, Durban, South Africa)

Abstract. Inverted High Electron Mobility Transistors (IHEMTs) with improved electrical performance have been fabricated. Extrinsic transconductances of up to 180 mS/mm at 300K, and drain-gate breakdown voltages of greater than 15V were measured on 0.5 micron devices. An analytic large signal model is presented which indicates that the IHEMT has promise as a power device.

1. Introduction

The IHEMT has received less attention than the conventional HEMT because of the larger mobilities obtained in the latter structure. However, the IHEMT may have advantages as a large signal device. Injection of carriers into the substrate results in poor drain conductance in short gate length HEMTs (Berenz et al. 1984 and Camnitz et al. 1984). In the IHEMT, enhanced electron confinement is inherent in the structure because there is a barrier (the conduction band discontinuity) to hot electron injection into the substrate. The optimised IHEMT should also withstand larger drain-gate voltages than HEMTs with comparable gate length.

In MESFETs, the maximum sustainable drain potential is limited by avalanche breakdown at the drain edge of the gate. Fixed ionized donors in the depleted channel induce a non-uniform electric field under the gate, which is largest at the drain edge of the gate. When this field exceeds a critical value, impact ionization occurs and the device breaks down. In the IHEMT, the fixed donor ions are displaced from the gate and induce a more uniform field on the gate, resulting in an increased drain breakdown voltage. In the regular HEMT, the donor ions are close to the gate and increased breakdown voltage is not anticipated. In both the IHEMT and HEMT, the zero gate bias saturation current is larger than in the MESFET (because of the greater average electron velocity in the two-dimensional electron gas) and hence offer possibilities as higher power devices.

2. Experimental Results

We have fabricated Inverted High Electron Mobility Transistors (IHEMTs) incorporating an n-GaAs/AlAs superlattice in place of the bulk n-AlGaAs layer. The superlattice is similar to that used by Baba et al (1984) in conventional HEMTs. Each superlattice period consists of 15 AlAs and 20 GaAs, of which the centre 10 was Si-doped to 10^{18} cm^{-3}. The spacer layer thickness was 85 . The undoped GaAs layer is 500 thick and is capped with 300 of n^+GaAs. The layers were grown by MBE at a temperature of 620 °C and at a growth rate of 0.1 micron/hr. Hall mobilities of 20 000 cm^2/Vs at 77K and 4000 cm^2/Vs at 300K were measured. The low 300K value is due to conduction by carriers in other layers, in parallel with the two dimensional electron gas (2DEG) carriers. The 2DEG sheet carrier concentration was 7.5×10^{11} cm^{-2}.

One half micron gate length devices were fabricated using conventional optical lithography. Typical I-V characteristics are shown in figure 1. The devices could not be completely pinched off because of a parallel conduction path in the partially depleted thick superlattice. Extrinsic transconductances of up to 180 mS/mm at 300K, and gate-drain breakdown voltages of greater than 15V were measured. The breakdown voltage is the highest reported for either HEMT or IHEMT structures. The source resistance was 1.8 ohm.mm from which the intrinsic transconductance was calculated to be 270 mS/mm. From this value, the average electron transit velocity was calculated (Eastman 1985) as 1.7×10^7 cm/s. This value is the largest reported for either IHEMT or Quantum Well HEMT structures.

3. IHEMT Large Signal Model

An analytic large signal model has been developed for the IHEMT. The electric field components acting on the 2DEG electrons are obtained from the solution of Poisson's equation along the channel. In this model, the intrinsic device is divided into three regions (figure 2).

In region I, the electrons are assumed to have constant mobility and the electric field is assumed one-dimensional in the channel (y-directed). The drain-source current is calculated using the standard Schockley formulation (equation 1, Table 1). The 2DEG sheet concentration decreases towards the drain in this region because of the voltage drop along the channel. The 2DEG concentration is calculated using a simple capacitor approximation which has been shown by Lee et al. (1984) to be reasonably accurate for the IHEMT (equation 2). A two-piece velocity-field characteristic is assumed and the boundary between regions I and II is the point at which the longitudinal field is equal to the critical field for velocity saturation, E_s (2.4 kV/cm).

In region II, the carriers are assumed to travel at the saturated drift velocity. The drain-source current is calculated using equation 3 and clearly this current must be equal to that calculated using equation 1. The 2DEG concentration is assumed constant in this region and is calculated using equation 2 with the potential at the drain end of region I (equation 4). In order to obtain a two-dimensional solution for the field in region II, the 2DEG and the adjacent thin n-AlGaAs layer are modelled as a sheet of charge with density given by equation 5. Solving Poisson's equation with suitable boundary conditions yields an expression for the longitudinal electric field acting on the 2DEG

(equation 6). This equation is similar to that obtained by Grebene and Ghandhi (1969) in their analysis of the MESFET. The potential at the drain edge of the gate is obtained by integrating equation 6 and adding the potential at the boundary between regions I and II (equation 7).

In region III, a two-dimensional solution for the field components is obtained by assuming that the y-directed component of electric field is determined by the occupied surface state density and is independent of x-coordinate (equation 8). The remainder of the field arising from the net positive sheet charge density is x-directed. The sheet carrier concentration and velocity are assumed to be the same as in region II.

For given bias voltages, the equations are solved simultaneously by successive approximation. The effect of source resistance is modelled by calculating the current in the intrinsic device and then calculating the voltage drop across the resistance. New terminal voltages for the intrinsic device are obtained and the procedure is repeated until the current in the intrinsic device and in the source resistance are equal. Figure 3 shows the variation of longitudinal electric field along the channel of the device for three values of n_{ss}.

Electrons in the 2DEG are accelerated by the longitudinal field and may be scattered over the energy barrier at the interface, into the AlGaAs. This real space transfer results in a component of drain current in parallel with the 2DEG current, which increases the saturated device drain conductance. The barrier height which limits electron transfer from the 2DEG to the AlGaAs layer is determined by the conduction band edge discontinuity at the heterojunction. For aluminium mole fractions less than 0.45, the discontinuity is approximately 0.6 of the direct energy gap difference between the GaAs and AlGaAs (Batey et al. 1985).

A simple energy conservation equation is used to calculate the electron temperature of the 2DEG electrons and hence, assuming a Maxwellian energy distribution, the number of electrons with energy greater than the barrier height. Fifty per cent of those electrons are assumed to be scattered into the AlGaAs where they form a parallel current component and increase the drain conductance. Figure 4 shows simulated I-V characteristics for an IHEMT.

The gate-drain breakdown voltage for the IHEMT can be determined by calculating the ionization integral from the electric field distribution, assuming the breakdown mechanism to be the onset of impact ionization in the channel. Using this model, we estimate that an optimised IHEMT may achieve a maximum Class A output power of 1.6 W per millimeter gate width per micron gate length. Optimising the IHEMT involves a compromise between breakdown voltage and transconductance; as the GaAs layer thickness is increased, the breakdown voltage increases but the transconductance decreases.

4. Conclusion

IHEMTs with improved electrical characteristics have been fabricated. By decreasing the source resistance and decreasing the GaAs layer thickness, extrinsic transconductances of greater than 300 mS/mm at room temperature should be achievable. An analytic model for the IHEMT has been discussed. The model will be improved by the use of a three-piece velocity-field characteristic and more accurate calculation of hot

electron scattering into the AlGaAs layer.

Acknowledgement

A.J. Hill wishes to acknowledge the financial support of the University of Natal, Durban, South Africa.

References

Batey J., Wright S.L. and DiMaria D.J., J. Appl. Phys., 57(2), 1985, pp. 484-487

Berenz J.J., Nakano K., and Weller K.P., IEEE MTT-S Digest, 1984, pp. 98-101

Camnitz L.H., Maki P.A., Tasker P.J. and Eastman L.F., Inst. Phys. Conf. Ser. No. 74, 1984, pp. 333-338

Eastman L.F., private communication

Grebene A.B. and Ghandi S.K., Solid State Elec., Vol. 12, 1969, pp. 573-589

Lee K., Shur M.S., Drummond T.J. and Morkoc H., J. Vac. Sci. Technol., Vol. B 2(2), 1984, pp. 113-116

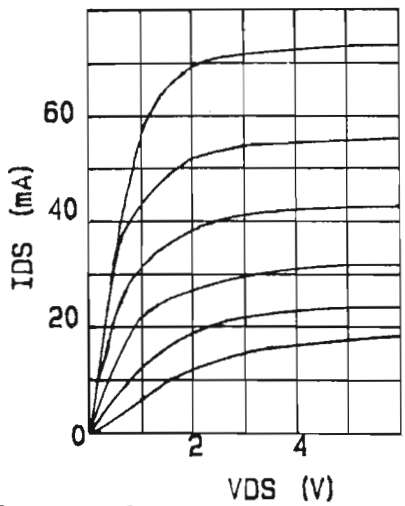


Figure 1: Typical measured I-V characteristics for 250 by 0.5 micron IHEMT. Scales: Vert. 10 mA/div, Horiz. 1V/div, 0.4 V steps.

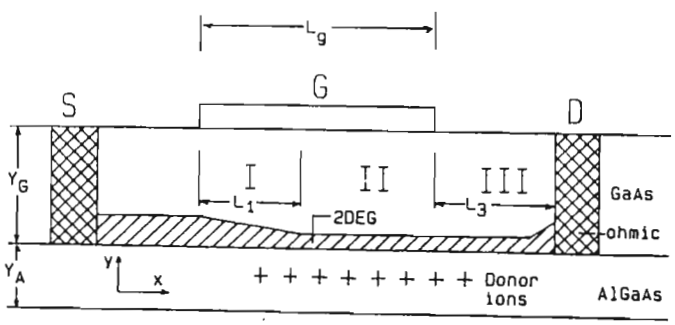


Figure 2: Simplified cross-section through an IHEMT. The three regions of the intrinsic device used in the model are shown. The vertical scale is greatly expanded.

$$I_{DS1} = \frac{\mu\varepsilon}{Y_G L_1}[(V_{GS}-V_T)V_1 - 0.5V_1^2] \quad (1)$$

$$n_s(x) = \frac{\varepsilon}{qY_G}(V_{GS}-V_T-V(x)) \quad (2)$$

$$I_{DS2} = q.n_{sv}.v_s \quad (3)$$

$$n_{sv} = \frac{\varepsilon}{qY_G}(V_{GS}-V_T-V_1) \quad (4)$$

$$\rho = q(N_D.Y_A - n_{sv}) \quad (5)$$

$$E_x(x) = E_s.\cosh((x-L_1)/Y) \quad (6)$$

$$Y = \sqrt{Y_A.Y_G} \quad (6a)$$

$$V_2 = V_1 + E_s.Y.\sinh((L-L_1)/Y) \quad (7)$$

$$E_y = \frac{q}{\varepsilon}n_{ss} \quad (8)$$

$E_x(x)$	Longitudinal (x-directed) electric field component, at x
E_s	Critical field for velocity saturation
I_{DS1}	Drain-source current (per unit gate width) in region I
I_{DS2}	Drain-source current (per unit gate width) in region II
L	Gate length
L_1	Length of region I
N_D	Ionized donor density in the n-AlGaAs layer
$n_s(x)$	2DEG sheet concentration, at x
n_{sv}	2DEG sheet concentration in regions II and III
n_{ss}	Occupied surface state density (region III)
q	Electronic charge
V_{GS}	Gate-source voltage
V_T	IHEMT threshold voltage
V_1	Potential, relative to source, at drain end of region I
V_2	Potential, relative to source, at drain end of gate
$V(x)$	Potential, relative to source, at x
v_s	Saturated drift velocity of 2DEG electrons
Y_A	n-AlGaAs layer thickness
Y_G	GaAs layer thickness
ρ	Sheet charge used in Poisson's equation
ε	Dielectric constant
μ	Electron mobility in region I

Table 1: Equations used in the IHEMT model.

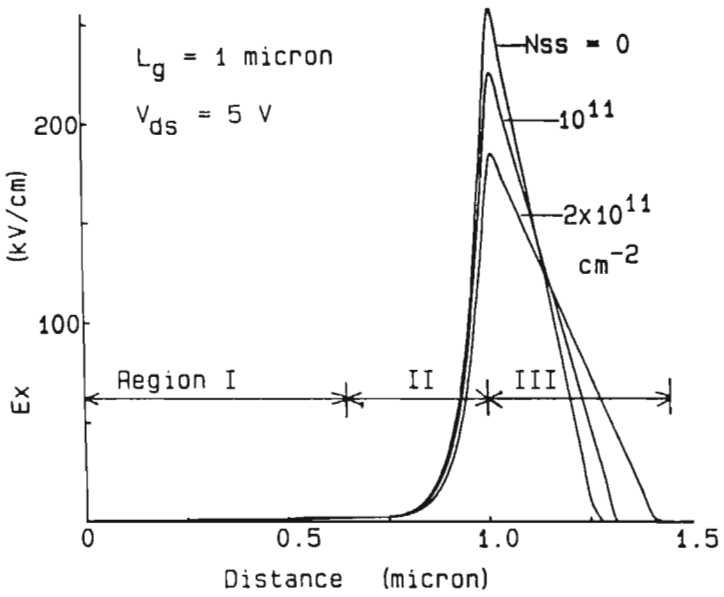


Figure 3: Simulated variation of longitudinal electric field along the channel of a 1 micron IHEMT showing the effect of occupied surface state density.

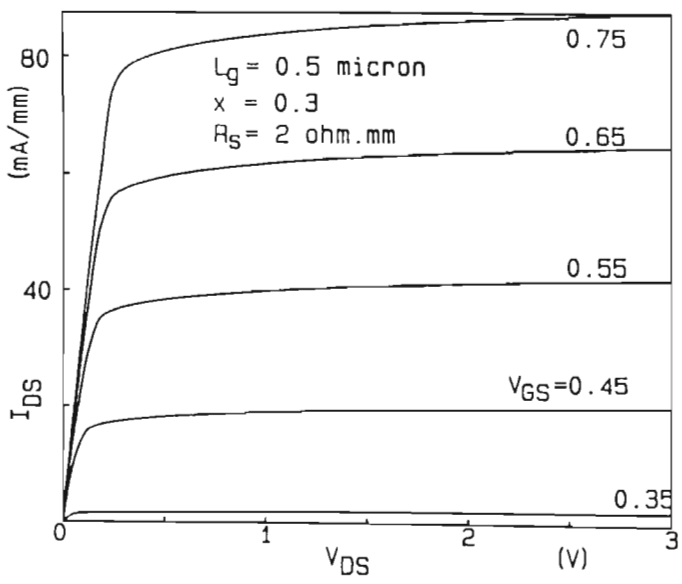


Figure 4: Simulated I-V characteristics for enhancement mode IHEMT (V_T = +0.32V).

APPENDIX E

This paper was presented at the IEE Colloquium on 'ICs above 1 GHz', London, 14 March 1985.

INVERTED HEMT, PROSPECTS AND UNCERTAINTY

A.J. Hill and P.H. Ladbrooke

Abstract

This paper discusses the large signal prospects of GaAs Inverted HEMT (IHEMT) devices and presents first calculations of substrate current and maximum output power in these devices.

IHEMT Output Power

The maximum available output power from a GaAs FET is commonly estimated as

$$P = 0.125\ I_{dss}\ V_b \qquad (1)$$

where I_{dss} is the zero gate bias saturation current and V_b is the gate-drain breakdown voltage. In the IHEMT, the physical limits on these quantities are different from those of the MESFET and indicate that the IHEMT structure is the logical consequence of correcting the known deficiencies of the MESFET structure.

I_{dss} is determined by the average electron drift velocity in the device. The higher average velocity obtained in the IHEMT (because the carriers propogate through undoped material) yields a greater than 50 per cent improvement in I_{dss} at room temperature over that obtained in MESFETs. The gate-drain breakdown voltage in the MESFET is limited by the non-uniform electric field under the gate induced by fixed ionic charge from the donor atoms in the channel. The induced field is largest at the drain edge of the gate. When this field exceeds a critical value impact ionization occurs and the device breaks down. In the IHEMT (figure 1), these fixed charges are further from the gate, which results in a more uniform field distribution along the gate and hence larger breakdown voltage. In the standard HEMT, the donor ions are closer to the gate and the device will exhibit similar breakdown characteristics to the MESFET.

A first estimate of I_{dss} and V_b, assuming breakdown to be caused by impact ionization in the channel (in the absence of better knowledge) and neglecting substrate current, yields

$$I_{dss} = qN_dY_av_s \qquad (2)$$

where N_d is the doping density in the AlGaAs, Y_a is the thickness of the AlGaAs layer and v_s is the average electron drift velocity, and

$$V_b = \frac{\varepsilon F_c^2 v_s Y_a L}{2(Y_a+Y_g)I_{dss}} \qquad (3)$$

where ε is the permittivity of GaAs, F_c is the critical field for the onset of impact ionization, Y_g is the thickness of the GaAs layer and L is the gate length.

GEC Hirst Research Centre, Wembley

Combining (1), (2) and (3) yields the output power

$$P_o = \frac{F_c^2 v_s Y_a L}{16(Y_a+Y_g)} \qquad (4)$$

On this basis the available output power from a 0.5 micron IHEMT is 725 mW/mm which compares favourably with the reported behaviour of 0.5 micron MESFETs (figure 2).

Substrate Current

The substrate current component of the drain current has been calculated as a fraction of the current in the two dimensional electron gas (2DEG) at the interface between the GaAs and AlGaAs. Substrate current determines the drain conductance of the device and should be minimized for maximum gain. Electrons in the 2DEG may gain sufficient energy from the field to scatter over the energy barrier at the interface and transfer into the AlGaAs. It is assumed that the scattering is isotropic so that 50 per cent of electrons with energy greater than the barrier energy will transfer into the AlGaAs and constitute the substrate current.

The energy barrier at the interface is determined by the conduction band edge discontinuity at the heterojunction interface. The sum of the conduction- and valence-band discontinuities is equal to the energy bandgap difference between GaAs and the larger bandgap AlGaAs. The energies of the Γ and X conduction band minima measured from the top of the valence band (K=0,0,0) [1], as a function of the aluminium mole fraction, x, are shown in figure 3. The L band is not shown because it has an energy intermediate between the Γ and X bands for all mole fractions and does not affect this discussion.

For mole fractions less than 0.43, $Al_xGa_{1-x}As$ is a direct bandgap semiconductor and the conduction band discontinuity is simply a fraction of the energy gap difference between GaAs and AlGaAs. This fraction, however, is not well established and values between 0.5 and 0.88 have been reported in the literature [2,3]. A value of 0.6 is assumed in this work.

At mole fractions greater than 0.43, the energy barrier for elecrons in the GaAs is even more uncertain since the AlGaAs is indirect for these mole fractions. It has been reported that the valence band discontinuity is a fraction (approximately 0.4, (1) in figure 3) of the direct (Γ) energy gap difference between GaAs and AlGaAs. If hot electrons in the GaAs Γ-valley can only transfer to the Γ-valley in the AlGaAs, then the conduction band discontinuity is 0.6 times the direct energy difference and is a monotonically increasing function of x ((2) in figure 3). If however, the electrons in the GaAs can transfer to the X-valley in the AlGaAs, the conduction band discontinuity will decrease for increasing x, because the sum of conduction and valence band discontinuities must equal the total bandgap difference ((3) in figure 3).

The ratio of substrate to 2DEG current as a function of drain- and gate-source voltage for a typical 1 micron gate length IHEMT with an aluminium mole fraction of 0.3 is shown in figure 4. Figure 5 shows the same ratio as a function of aluminium mole fraction for drain- and gate-source voltages of 5V and 1V respectively. The ratio is shown for both conduction band discontinuity possibilities outlined above. This figure clearly illustrates the importance of resolving the uncertainties surrounding the conduction band discontinuity since both accurate device modelling and optimization are at stake.

IHEMT in GaAs ICs

IHEMTs will play an increasingly important role in GaAs ICs in the future. IHEMT development is still in its infancy and seems certain to supplant the MESFET in many applications as its potential is realized in practical devices. The IHEMT offers advantages in both small and large signal performance but improved models are necessary to determine the optimum device parameters for particular requirements.

Acknowledgement

One of us (A.J. Hill) wishes to acknowledge the financial support of the University of Natal, Durban, South Africa.

References

[1] Bhattacharya P.K., Das U. and Ludowise M.J., "Transport properties of n-type metalorganic chemical-vapor-deposited $Al_xGa_{1-x}As$ ($0<x<0.6$)", Phys. Rev. B, Vol. 29, No. 12, pp. 6623-6631, 15 June 1984

[2] Dingle R., Wiegmann W. and Henry C.H., "Quantum States of Confined Carriers in Very Thin $Al_xGa_{1-x}As$-GaAs-$Al_xGa_{1-x}As$ Heterostructures", Phys. Rev. Lett., Vol. 33, No. 14, Sept. 1974, pp. 827-830

[3] Miller R.C., Kleinman D.A. and Gossard A.C., "Energy-gap discontinuities and effective masses for GaAs-$Al_xGa_{1-x}As$ quantum wells", Phys. Rev. B, Vol. 29, No. 12, pp. 7085-7087, 15 June 1984

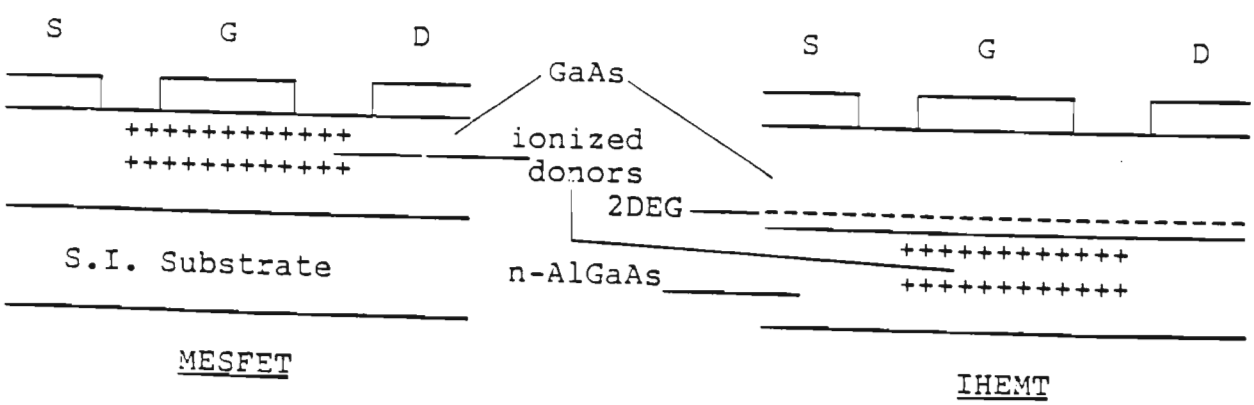


Figure 1. Simplified structure of MESFET and IHEMT

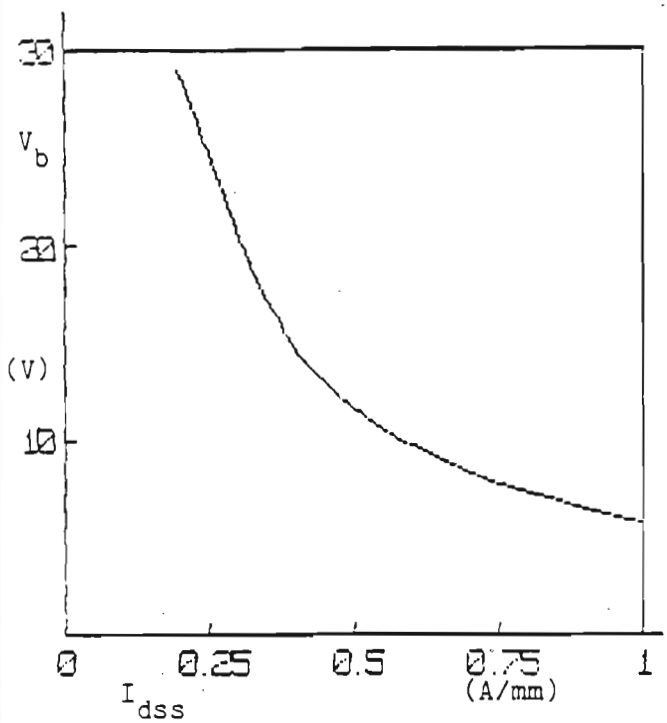


Figure 2: Graph of breakdown voltage is saturation current for a 0.5 micron IHEMT.

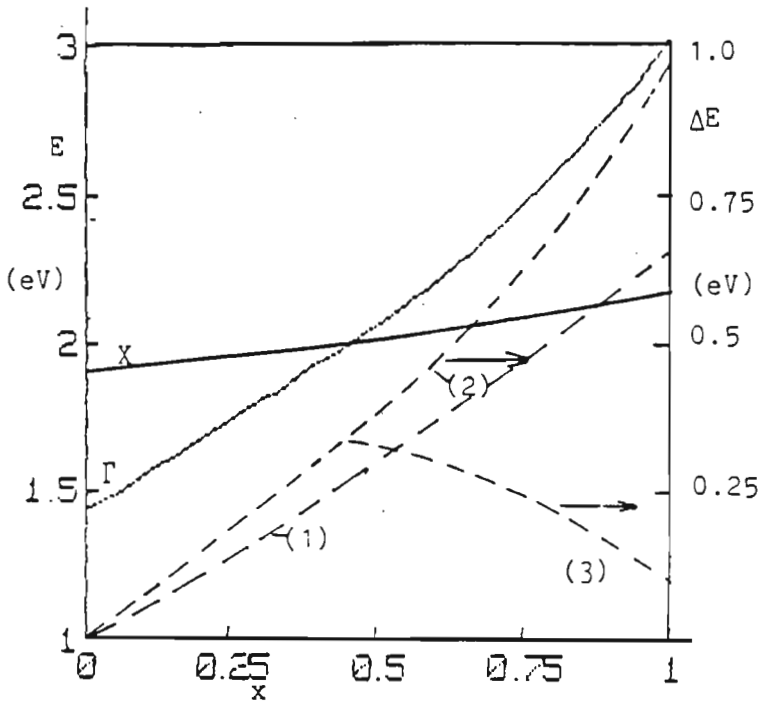


Figure 3: Graph showing Γ and X conduction band minima as a function of Al mole fraction. Discontinuities, as discussed in the text, are also shown.

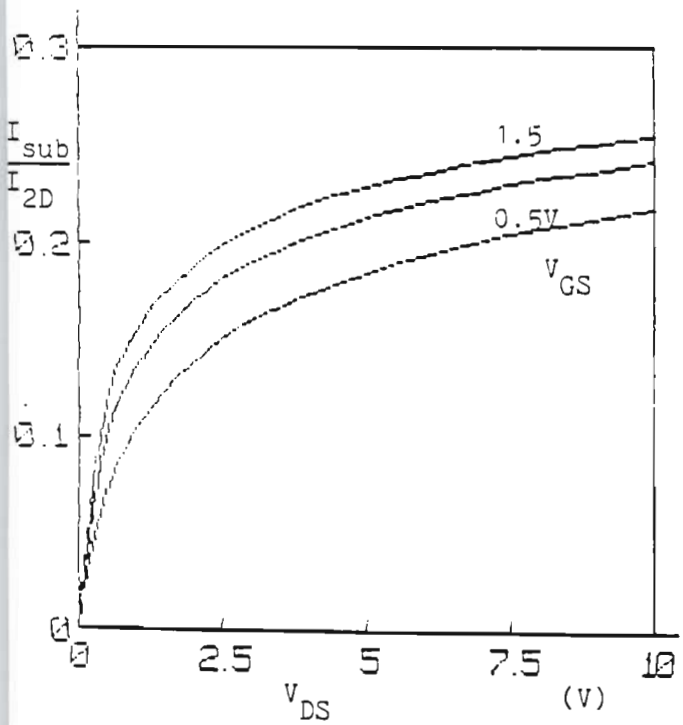


Figure 4: I_{subs} as a fraction of I_{2D} for a 1 micron gate length IHEMT.

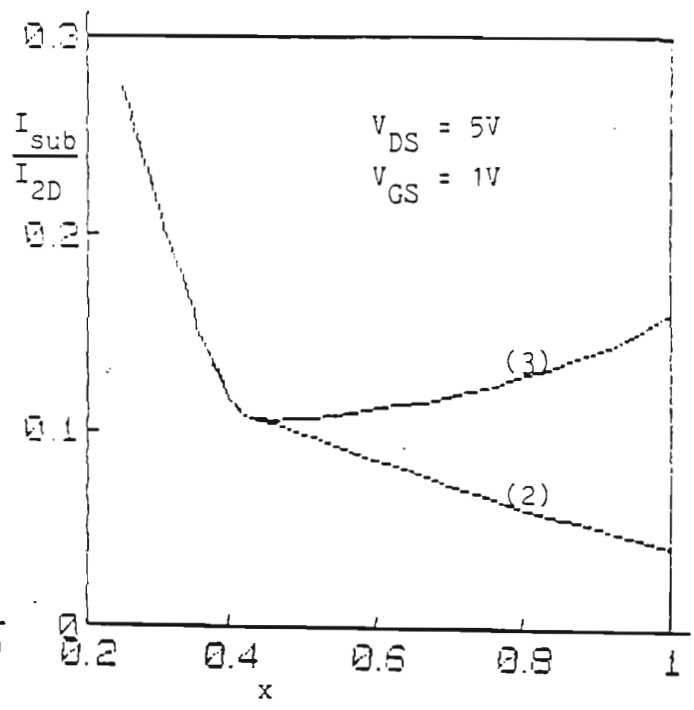


Figure 5: I_{sub} as a fraction of I_{2D} as a function of Al mole fraction for the same 1 micron IHEMT. The numbers correspond to the discontinuity possibilities illustrated in figure 3.

APPENDIX F

This appendix contains a reprint of a paper by the author from Electronics Letters, Vol. 22, No. 4, 14 February 1986, pp. 218-220.

DEPENDENCE OF CONDUCTION-BAND DISCONTINUITY ON ALUMINIUM MOLE FRACTION IN GaAs/AlGaAs HETEROJUNCTIONS

Indexing terms: Semiconductor devices and materials, Semiconductor junctions

The conduction-band discontinuity dependence on aluminium mole fraction for GaAs/AlGaAs heterojunctions is discussed and is shown to have a maximum value at a mole fraction of approximately 0·45.

The ratio Q_c of the conduction-band discontinuity ΔE_c to the total energy bandgap difference ΔE_g is one of the most important parameters determining the electrical properties of a heterojunction. The GaAs/$Al_xGa_{1-x}As$ heterojunction has been extensively studied because of its increasing application to advanced GaAs-based devices such as HEMTs and HBJTs. This letter discusses the dependence of the conduction-band discontinuity on the aluminium mole fraction x in GaAs/$Al_xGa_{1-x}As$ heterojunctions, particularly for Al mole fractions greater than 0·45.

The energies of the conduction-band minima in AlGaAs, relative to the top of the valence band at the Γ point (K = 0, 0, 0), as a function of Al mole fraction, are given[1] (in electron volts at 300 K) by

$$E_g^{\Gamma} = 1{\cdot}424 + 1{\cdot}247x \qquad x \leq 0{\cdot}45 \quad (1a)$$

$$= 1{\cdot}424 + 1{\cdot}247x + 1{\cdot}147(x - 0{\cdot}45)^2 \qquad x > 0{\cdot}45 \quad (1b)$$

$$E_g^{L} = 1{\cdot}708 + 0{\cdot}642x \quad (2)$$

$$E_g^{X} = 1{\cdot}900 + 0{\cdot}125x + 0{\cdot}143x^2 \quad (3)$$

These relationships are plotted in Fig. 1. For mole fractions less than approximately 0·45, the AlGaAs has a direct bandgap. For larger mole fractions, the alloy is indirect, with the X-valley having the lowest energy. The L-valley has an energy intermediate between the Γ- and X-valleys and plays no part in the discussion which follows.

The ratio of the valence-band discontinuity ΔE_v to direct Γ bandgap difference has been the subject of several recent studies (see, for example, References 2, 3 and 4, and further references therein). The majority of the measured values for this ratio lie in the range 0·33 to 0·41, and further show that the ratio is independent of aluminium mole fraction. We

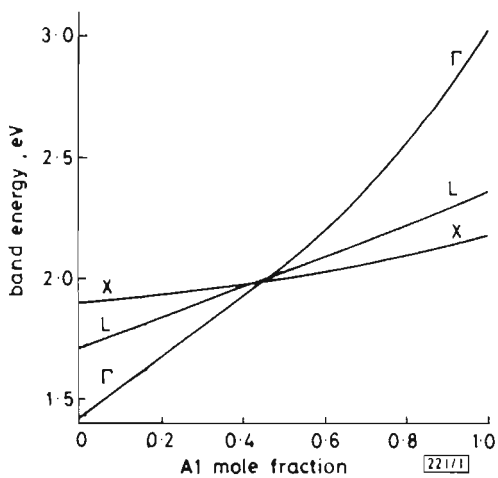


Fig. 1 *Energies of conduction-band minima as a function of aluminium mole fraction relative to top of valence band at Γ point*

assume a value of 0·35 for this ratio[3] which, together with eqn. 1, results in a valence-band discontinuity, as a function of mole fraction, given by

$$\Delta E_v = 0{\cdot}436x \qquad \text{for } x \leq 0{\cdot}45 \quad (4a)$$

$$= 0{\cdot}436x + 0{\cdot}401(x - 0{\cdot}45)^2 \qquad x > 0{\cdot}45 \quad (4b)$$

This result is shown as curve A, Fig. 2. Several measured[3-7] values for ΔE_v are also shown for comparison (crosses).

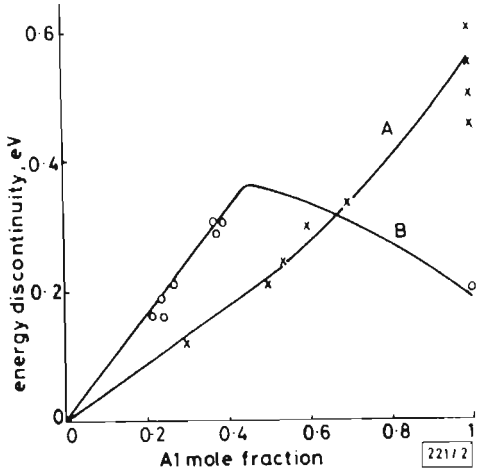


Fig. 2 *Dependence of valence- and conduction-band discontinuities on aluminium mole fraction*

Measured values (○ and × symbols) are discussed in text

Since the sum of the valence- and conduction-band discontinuities must equal the energy bandgap difference between the GaAs and AlGaAs, for mole fractions less than 0·45, Q_c must equal 0·65, and from eqn. 1*a* the conduction-band discontinuity is

$$\Delta E_c = 0{\cdot}81x \qquad x \leq 0{\cdot}45 \quad (5)$$

This relationship is shown as curve B, Fig. 2. The conduction-band discontinuity has been experimentally determined by several groups on heterojunctions with Al mole fractions less than 0·45. Some of these measured values[4,8] are also shown in Fig. 2 (circles).

For Al mole fractions greater than 0·45, AlGaAs has an indirect bandgap with the X-valley lowest in energy. The bandgap in this region increases slowly with increasing mole fraction in comparison to the rapid increase in the Γ-valley energy (Fig. 1). This means that the energy bandgap difference increases slowly while the valence-band discontinuity increases rapidly (eqn. 4*b* and Fig. 2). Since the sum of the valence- and conduction-band discontinuities must still equal the bandgap difference, the conduction-band discontinuity must decrease with increasing Al mole fraction beyond 0·45.[9,10] From eqns. 1 and 3, the energy bandgap dependence on Al mole fraction is given by

$$\Delta E_g = 0{\cdot}476 + 0{\cdot}125x + 0{\cdot}143x^2 \qquad x > 0{\cdot}45 \quad (6)$$

From eqns. 6 and 4*b* the conduction-band discontinuity is given by

$$\Delta E_c = 0{\cdot}395 + 0{\cdot}05x - 0{\cdot}258x^2 \qquad x > 0{\cdot}45 \quad (7)$$

This relationship is shown in curve B, Fig. 2. This result is very different from that predicted by extrapolating the results from heterojunctions with direct bandgap AlGaAs,[3] or predicted by theoretical guides such as Anderson's electron affinity rule, which predicts $\Delta E_c = 0{\cdot}43 + 0{\cdot}14x$, for x greater than 0·45 (Reference 11, eqn. 27). The validity of eqn. 7 is demonstrated by the recently reported value[6] for ΔE_c of approximately 0·2 eV (Fig. 2) for a GaAs/AlAs heterojunction, which was calculated from the measured sheet concentration of accumulated electrons adjacent to the heterointerface.

The maximum conduction-band discontinuity is approximately 0·365 eV and occurs for an Al mole fraction of 0·45.

corresponding to the transition from direct to indirect AlGaAs. This is important for the optimisation of GaAs devices incorporating a heterojunction, such as HBJTs and HEMTs.

In HEMTs, for example, electrons in the two-dimensional electron gas (2DEG) may gain sufficient energy from the applied field (drain-source potential) to be scattered over the energy barrier at the heterointerface, into the AlGaAs. This real space transfer results in a component of drain current in parallel with the 2DEG current which has the deleterious effect of increasing the saturated device drain conductance. The barrier height is simply ΔE_c, and hence to minimise this effect a mole fraction of approximately 0·45 should be used for the spacer layer to yield the largest possible barrier height.[9,10]

The exact Al mole fraction at which the direct indirect transition occurs is not well established,[8] but it is clear that the maximum value for ΔE_c will occur for heterojunctions with an AlGaAs layer of this mole fraction. More work is required using Al mole fractions between 0·45 and 1 to verify the exact dependence of eqn. 7, as has been done for lower mole fractions.

A. J. Hill wishes to acknowledge the financial support of the University of Natal, Durban, South Africa.

A. J. HILL

19th December 1985

P. H. LADBROOKE

GEC Hirst Research Centre
East Lane
Wembley, Middx. HA9 7PP, United Kingdom

References

1 CASEY, H. C., and PANISH, M. B.: 'Heterostructure lasers, Part A: fundamental principles' (Academic Press, New York, 1978)

2 WATANABE, M. O., YOSHIDA, J., MASHITA, M., NAKANISI, T., and HOJO, A.: 'Band discontinuity for GaAs/AlGaAs heterojunction determined by C-V profiling technique', *J. Appl. Phys.*, 1985, **57**, pp. 5340–5344

3 ARNOLD, D., KETTERSON, A., HENDERSON, T., KLEM, J., and MORKOC, H.: 'Electrical characterisation of GaAs/AlGaAs semiconductor-insulator-semiconductor capacitors and application to the measurement of the GaAs/AlGaAs band-gap discontinuity', *ibid.*, 1985, **57**, pp. 2880–2885

4 BATEY, J., WRIGHT, S. L., and DIMARIA, D. J.: 'Energy band-gap discontinuities in GaAs: (Al, Ga)As heterojunctions', *ibid.*, 1985, **57**, pp. 484–487

5 WANG, W.I., MENDEZ, E. E., and STERN, F.: 'High mobility hole gas and valence-band offset in modulation-doped p-AlGaAs GaAs heterojunctions', *Appl. Phys. Lett.*, 1984, **45**, pp. 639–641

6 WANG, W. I., and STERN, F.: 'Valence band offset in AlAs/GaAs heterojunctions and the empirical relation for band alignment', *J. Vac. Sci. Technol. B*, 1985, **3**, pp. 1280–1284

7 KELLY, M. K., NILES, D. W., COLAVITA, E., MARGARITONDO, G., and HENZLER, M.: 'Valence-band discontinuities at AlAs-based heterojunction interfaces', *Appl. Phys. Lett.*, 1985, **46**, pp. 768–770

8 HEIBLUM, M., NATHAN, M. I., and EIZENBERG, M.: 'Energy band discontinuities in heterojunctions measured by internal photoemission', *ibid.*, 1985, **47**, pp. 503–505

9 HILL, A. J., and LADBROOKE, P. H.: 'IHEMT, prospects and uncertainty'. IEE colloquium on GaAs ICs, London, 14 March 1985

10 HILL, A. J., LADBROOKE, P. H., RANSOME, S. R., and WESTWOOD, D.: 'Large signal modelling and practical performance of inverted HEMT'. Proc. 10th IEEE/Cornell Conf., July 1985, to be published

11 ADACHI, S.: 'GaAs, AlAs and $Al_xGa_{1-x}As$: Material parameters for use in research and device applications', *J. Appl. Phys.*, 1985, **58**, 1 Aug. pp. R1–R29

APPENDIX G

This paper was published in the GEC Journal of Research, Vol. 4, No. 2, 1986. This author's contribution to the paper was the calculations used for the comparison of GaAs and Si FETs.

APPLIED DEVICE PHYSICS IN THE GaAs INDUSTRY

P. H. LADBROOKE, J. P. BRIDGE and A. J. HILL

Applied device physics is an activity based upon developing a physical understanding and mathematical description of semiconductor devices, involving close interaction with basic research, materials scientists, fabrication technologists and circuit engineers. Broadly, the initial intention is twofold. First, to assimilate the electrical performance required of an active device by a circuit, and to translate that requirement into a material and device structure specification for implementation by materials growers and process technologists. Secondly, to identify new and alternative materials and device structures which offer lower noise, higher gain, higher power, higher speed or frequency response, or operation over an extended range of temperature, depending upon the application.

Once a capability of this kind has been established, it may be applied to the commercial question of forecasting the yield of a given semiconductor component from a process line with fixed tolerances, and to the inverse problem of predicting the material and processing tolerances required for any desired yield, and hence chip selling price. Applied device physics, therefore, when developed into a full capability, has the potential to direct investment and effort into the most rewarding fields without the need to commit resources to the time-honoured but expensive approach of repeated practical trial. An outline of the principles involved, and a survey of some typical applications to the GaAs FET and integrated circuit industry, are given in this paper.

1. INTRODUCTION

In the continuing quest to accommodate ever more complex signal processing functions into semiconductor integrated circuits of ever smaller dimensions, new materials, new process technologies and new amplifying and switching device structures are coming to the forefront all the time. Much of what is done depends upon the efforts of the solid-state physicist in the first instance, yet it may take some years to realise the commercial potential of his work. It is not unknown for ideas to remain unexploited, for example, for the lack of a suitable technology, and experience has shown that hitherto unthought-of phenomena can be revealed by the results of experiments made possible by new technology. Thus there exists, in the semiconductor device and integrated circuit business, a strong interdependence between solid-state physics, technology and device engineering.

By broad category, the tasks which have to be addressed by a commercial enterprise may be stated as follows:

(a) Design and development of semiconductor components, which can be manufactured commensurate with (b) below.
(b) Specification of a technology, and its attendant tolerances, required to manufacture semiconductor components at affordable cost.
(c) Diagnosis of faults and failures in finished components.
(d) Forecasting the benefits to be had from the development and, perhaps, the introduction of new materials or technologies.

Fig. 1. The dual role of modelling.

(e) Assessment of the feasibility and worth of new ideas of principle in device and circuit fields.

A common response to these needs is the age-old one of practical trial. Practical trial is widely invoked as the initial response to a new idea or difficulty, not, as it ought to be, the final or proving one which constitutes the ultimate test. In a high technology business, where each individual piece of major equipment costs a million pounds, practical trial can be prohibitively expensive. Thus, an alternative must be sought. What ideally is required is a 'software' capability to predict the electrical characteristics of a component in terms of the semiconductor material properties and detailed dimensions which result from the process technology. Such is the realm of device physics applied to the needs of industry. Once such a capability has been developed, it can be operated in reverse to cater for the issue identified at (b) for example (see fig. 1), and incidentally all the other needs as well.

What might be termed the 'power' of modelling has been adequately expressed by Getreu[1] in his published work on the bipolar transistor. On the grounds that almost all GaAs analogue circuits, and very many digital circuits as well, are based on the Schottky-gated FET (or MESFET), the present

paper deals exclusively with that particular structure. The emphasis is given to the underlying principles of applied devices physics and its uses, rather than to the extensive mathematical and experimental background to the development of such methods.

2. OBJECTIVES AND METHODS

(a) Objectives

The fundamental objective is to determine how the electrically active devices in a circuit will behave as a function of the material and device structural properties which are the result of technological processing*. In round numbers, there are ten features of the MESFET which determine its electrical behaviour (Figure 2a):

(i) R_c: ohmic metal contact resistance to n-GaAs
(ii) L_G: gate length
(iii) L_{SG}: source-gate spacing
(iv) N: n-layer doping
(v) W: n-layer thickness under the gate
(vi) ϕ_B: interfacial barrier of potential between the n-layer and the buffer/substrate (which depends heavily on buffer/substrate properties)
(vii) R: gate recess depth
(viii) N_T: deep level density
(ix) D_S: density of occupied surface states
(x) L_{GD}: gate-drain spacing

All of these factors affect, in a not entirely obvious way, the electrical behaviour as expressed by:

(1) I_{DSS}—open channel current with gate shorted to source
(2) I_f—open channel current with positive gate voltage
(3) V_p—pinch-off voltage
(4) V_{DSB}—initial drain-source avalanche voltage (for $|V_{GS}| \geq |V_p|$)
(5) self-regulating impact ionisation current which follows initial avalanche
(6) all element values in the equivalent circuit, and their bias dependence (Figure 2*b*).

The objective, therefore, is to find part or all of the set (1–6) in terms of the set (i–x). It is this procedure which is referred to by the term 'modelling'.

(b) Methods

Methods of modelling range from compact analytic theories at one extreme to full Monte Carlo simulators at the other. The GEC Hirst Research Centre has a Monte Carlo simulation capability of its own, but generally finds simulators no more useful than

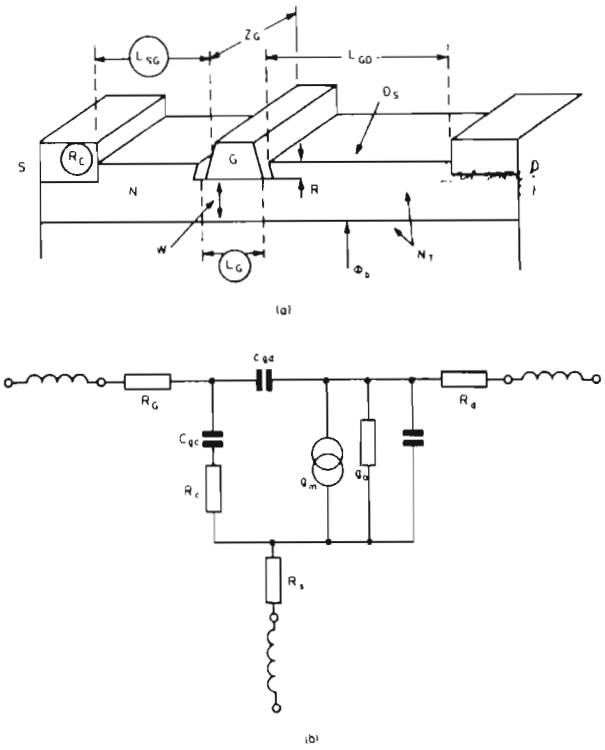


Fig. 2. GaAs MESFETs—Equivalent circuit element dependences.

models which treat electron transport in a less sophisticated manner. The suspected reason is that uncontrolled and, to a degree, poorly understood processes occurring in GaAs FETs render most (if not all) simulators unrepresentative of the real device or, at least, no more representative than well-considered analytic theories. Uncertainty rests with transfer of charge to and from the device surface during operation, the technological factors which govern it, active layer/buffer-substrate interface charging and transport, and poor screening of the high-field region in the gate-drain space in practical circuits. All of these factors affect the electric field distribution within the operating FET, and hence the terminal current I_D which flows in response to the impressed potentials V_{GS} and V_{DS}. With reference to fig. 3,

$$I_D = \int_{A_D} \boldsymbol{J} \cdot d\boldsymbol{A}$$

But $\boldsymbol{J}(x, y)$ is a function of the two-dimensional vector electric field, $\boldsymbol{E}(x, y)$. The connecting relationship is the way in which electrons are transported in the field, a relationship which for present purposes will be symbolically denoted by $\tau(\boldsymbol{E})$. Thus I_D depends upon $\tau(\boldsymbol{E})$.

In order of commonly accepted accuracy, the descriptions of electron transport are:

(i) The Drift-Diffusion Approximation

In this (least accurate) description, the electrons are assumed all to have a directed component of velocity (drift) and random components of velocity (diffusion) which are instantaneous functions of the local electric field. Relationships for drift velocity

* *The same requirement also exists, of course, for the passive components in an integrated network. In monolithic microwave integrated circuits (MMICs), this can be no easy task: electromagnetic coupling between closely spaced components on a chip is a particularly complicating factor and will become even more so as functional densities increase in the drive for smaller chips and lower unit costs.*

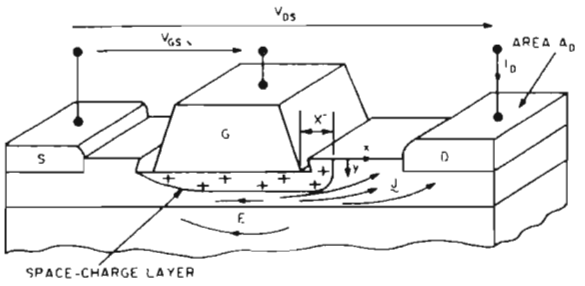


Fig. 3. Illustrating transport of electrons in the field $E(x, y)$ to give rise to current density J.

$v(E)$ and diffusion coefficient $D(E)$ derived from transport in long, homogeneous samples are assumed to apply even in devices (such as the sub-micron gate length GaAs MESFET) where there are large potential gradients over distances comparable with a mean free path for scattering events.

(ii) Transport (or moment) equations

In order to allow for the physical fact that electrons possess inertia, and so do not in reality respond instantaneously to changes in electric fields, a set of coupled equations expressing conservation of momentum and energy among the electrons can be derived[2] based upon some assumed form for the electron distribution function $f(\boldsymbol{k}, \boldsymbol{r}, t)$. For this purpose the explicit dependence upon wave vector (or momentum) $\boldsymbol{k}$, position in real space, $\boldsymbol{r}$, and time t is reduced to the single explicit functional dependence $f(\boldsymbol{k})$: any dependence on $\boldsymbol{r}$ and t is allowed for implicitly, i.e. by regarding $\boldsymbol{k}$ as a function of $\boldsymbol{r}$ and t. Commonly used approximations for $f(\boldsymbol{k}[\boldsymbol{r}, t])$ are Maxwellian and drifted Maxwellian distributions[3].

(iii) Monte Carlo simulation

Calculation of the distribution function $f(\boldsymbol{k}, \boldsymbol{r}, t)$ without the ab-initio assumption of a particular form yields transport results which, from the point of view of fundamental principle, should be the most accurate. The Monte Carlo method is used to simulate the timing and direction of scattering events which determine the trajectory of an electron moving in the internal electric fields in a device. By averaging over the trajectories of many electrons, the electron velocity, current density and terminal current can be found.

Regardless of what should constitute the most accurate method in principle, what matters is which of these choices is the most workable in practice, from all points of view. The problem would seem to be that Monte Carlo simulators, transport equation simulators and, up until recently*, even simulators based upon drift-diffusion transport have largely ignored the fact that surface and interface charges and potentials have a major bearing on the electric field $\boldsymbol{E}(x, y)$ in which the electrons are transported via the symbolic relationship $\tau(\boldsymbol{E})$. It is not surprising, therefore, that simulators often predict device characteristics quite different from those measured in practice. There would appear to be little point in going to great lengths to describe the transport relationship accurately if the field $\boldsymbol{E}(x, y)$ in which transport takes place is in sizeable error. It is worthwhile, moreover, to examine the merits of accepting the lowest-level transport description while making some effort to include the effects of surface state occupancy, D_s, and interface potential, ϕ_b, (see fig. 2a) on the field, the space-charge layer (fig. 3), the drain current, and the equivalent circuit elements (fig. 2b). It turns out that much of value can be achieved with this approach while retaining a completely analytic formulation (the only equivalent circuit element which cannot be calculated completely analytically is the drain conductance, g_0 in fig. 2b).

To illustrate the usefulness of the approach, fig. 4 compares the calculated and independently measured[4] dependences of gate-drain capacitance, which is one of the most convoluted electrical quantities to interpret in the GaAs MESFET. For a planar FET with no surface charge, the expression for the space-charge contribution to C_{gd} is particularly simple in the current saturation region:

$$C_{gd} = 2\varepsilon/(1 + 2X/L) \quad \text{per unit gatewidth}$$

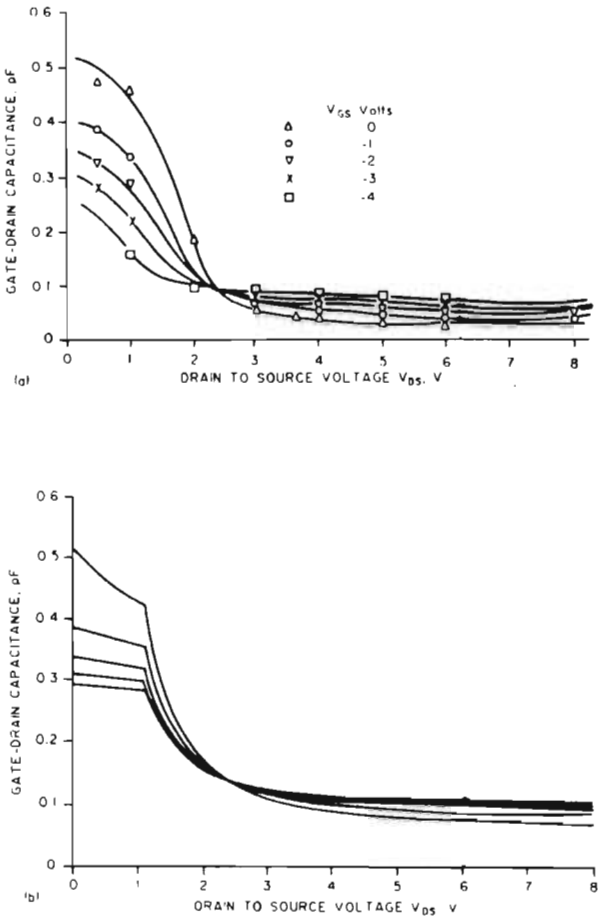


Fig. 4. Gate-drain capacitance variation in a GaAs MESFET. (a) Experimental data from Willing et al[4]. (b) Model calculations using Willing's published data on material and device structure.

* *T. M. Barton of Leeds University is working under GEC sponsorship to develop a simulator which includes surface effects.*

where ε is the permittivity of GaAs, X is the space-charge layer extension (see fig. 3) and L is the gate length.

The real trick, however, lies in determining how this expression needs to be modified to account for surface charge, and how X depends upon the applied potentials V_{GS} and V_{DS}. These factors have been accounted for in fig. 4.

It is as well to spell out the differences between the simulation and physical or analytical approaches to help place them in context with one another. Generally speaking, simulators aim to include the best possible description of fundamental phenomena, particularly scattering modes and rates, and to solve self-consistently for the interdependence of all these effects, subject to the boundary conditions imposed by the device structure, where known. With such a procedure it is necessary only to write equations describing the basic phenomena and to solve them as accurately as numerical schemes will allow. It is largely a question of carrying out systematic, logical and numerous operations with a high order of repetition.

The development of analytical models is quite different. To be analytically tractable, equations have to be acknowledged approximations to begin with: there is no hope whatsoever of solving a full equation set analytically in two dimensions subject to FET boundary conditions. There is then the difficulty, not faced by the simulator, as to what it is one can discard and what must be retained when making these approximations.

Many clues can be found in purpose-designed, carefully executed, experiments. Continued iteration between theory and experiment leads eventually to a model for the device in the form of a set of analytic equations in which the uncertainties are no worse than the uncertainties present in very carefully fabricated and measured practical devices. This procedure may take very much longer than the simulator to develop but, once developed, the speed with which devices may be designed and analysed within practical limits is unparalleled.

GEC has found analytic models to be as representative as Monte Carlo models of the kind of FET structures commercial organisations deal with. From the point of view that the time to analyse, diagnose or design practical FETs with such models is very much less than that required by a simulator, analytic models have found greatly more application within GEC Research in GaAs FET device and circuit design and development. Accordingly, GEC has applied such models to the following specific problems in recent times:

- Diagnosis of discrete 0.5 μm FETs and X-band integrated circuits for improved specification limits.
- Design of 0.5 μm gate length FETs specific to particular circuit applications (for example, for a 0.1–26.5 GHz wideband amplifier for instrument applications). Both the active device and the passive network are designed concurrently, i.e. the entire network is designed in this approach, not just the passive network around previously ascertained FET s-parameters.
- Examination of the merits of using different materials in FETs and MMICs.
- Prediction of the kind of device structure that may be necessary to realise a 100 GHz transistor for integrated use.
- Analysis of the sensitivity of a travelling wave 0.5–18 GHz amplifier to material and process variations (such an activity is sometimes referred to as 'Process Modelling').
- Yield prediction for HRC's current technology.
- Forecasting of the material and process tolerances required to achieve a desired chip selling price

Examples of these uses are given in Section 3.

3. APPLICATIONS

(a) Engineering sub-micron FETs for specific applications

Many high-speed circuits, be they analogue or digital, require active device elements which are specific to that particular application. The requirements of a MESFET structure to be integrated into a low-noise amplifier are, for example, quite different from those of a MESFET for an integrated power amplifier. Furthermore, it is increasingly the case that circuits demand of the device the ultimate performance which the material and structure will allow—a principle which applies to both analogue and digital integrated circuits, and to devices of both field-effect and bipolar type.

Two key applications of GaAs MESFETs are in low noise amplification (for receivers) and in broadband systems. Circuit parameters and operating conditions of the FET elements differ in the two cases.

Consider first the low noise requirement. From Fukui[5], the noise figure is given in terms of the equivalent circuit elements as

$$NF = 1 + K_f \omega C_{gc} ((R_s + R_G)/g_{m0})^{1/2}$$

with $K_f = (2 \rightarrow 2.5)$.

The use of a highly doped n-layer ($\geq 2 \times 10^{17}\,\mathrm{cm}^{-3}$) gives rise to low R_S and high g_{m0}, both of which are favourable. The bulk contribution to R_S may be further reduced by recessing the gate strip below the surface of the semiconductor to a depth R (see fig. 2a): a 50 nm recess may reduce R_S by up to 3 ohms relative to a planar device. The introduction of the recess incidentally reduces the gate-channel capacitance by a few percent or a few tens of percent, but it also degrades the drain-source breakdown (fig. 5a). To minimize the risk of excess noise due to impact ionization, the drain bias voltage V_{DS} is required to be, say, 2V below V_{DSB}.

It follows that for normal FET operation with a drain bias between 3–4V, a minimum V_{DSB} of 6V is

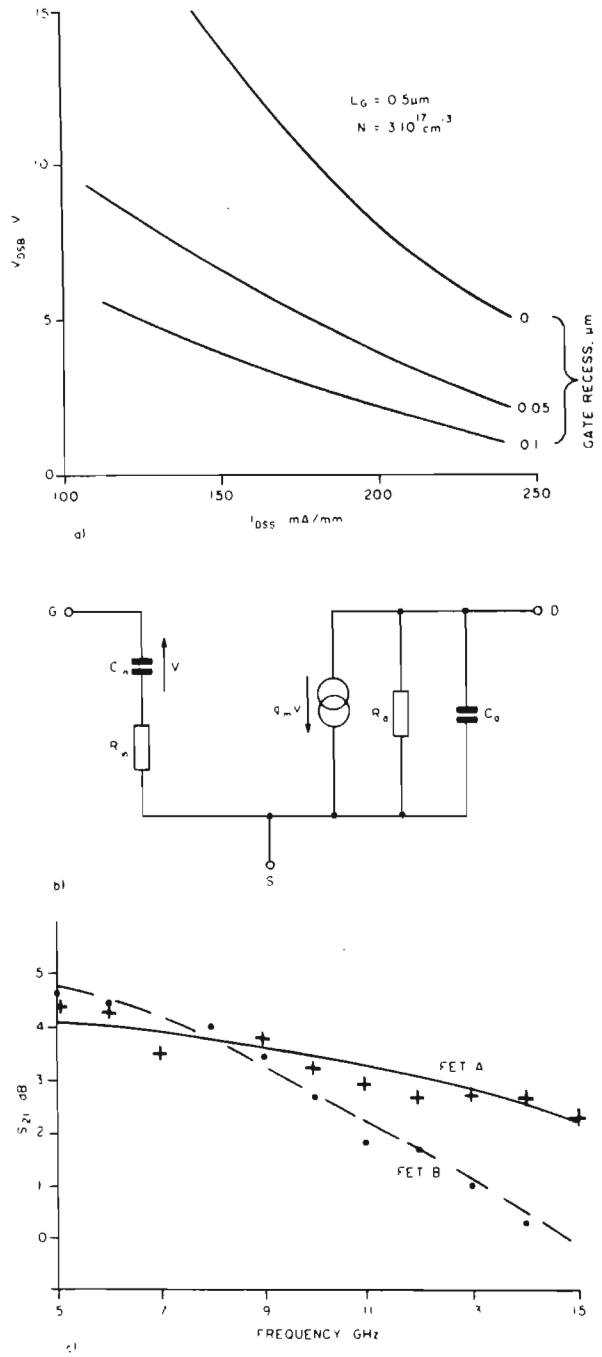


Fig. 5. 0.5 µm small-signal GaAs FET considerations. (a) Drain-source breakdown voltage (V_{DSB}) as a function of open channel current (I_{DSS}) and gate recess depth. (b) Simplified equivalent circuit (common source). (c) Measured scattering parameters, 5–15 GHz.

required. Figure 5a illustrates the dependence of V_{DSB} on the open channel current (I_{DSS}) for a doping level of 3×10^{17} cm^{-3}; from these curves it is evident that to achieve a usable breakdown voltage of ≥6V from a 50 nm recessed gate device, a maximum I_{DSS} of only 160 mA/mm gate width can be tolerated. For a typical FET this limits I_{DSS} to around 40 mA. Commercially available 0.5 µm low noise FETs adopt a highly doped (≥3.10^{17} cm^{-3}) gate recessed structure with relatively low I_{DSS} (≤40 mA) and pinch off (≤1.5 V). Unfortunately such a device is unlikely to be suitable for high frequency or broadband applications. The high doping gives rise to increased input capacitance (C_{in}) which enhances gain-frequency roll off and reduces the upper frequency limit of the device.

For broadband applications, it is generally true that noise performance is a secondary consideration and that maximum gain-bandwidth is the overriding objective. With reference to the highly simplified equivalent circuit of fig. 5b, the forward gain (S_{21}) is dominated by the intrinsic transconductance (g_{m0}) at low frequencies and the input time constant ($C_{in}R_{in}$) at higher frequencies.

Considering the two major components of C_{in} and R_{in}, namely the channel resistance R_i and the gate-channel capacitance C_{gc}, the following very much simplified approximations can be written

$$C_{gc} \simeq \varepsilon_0 \varepsilon_r L Z_G / W(1 - I_D/I_F)$$

$$R_i \simeq L v_{sat}/\mu I_D$$

where, for a given channel doping and bias condition, I_D increases with layer thickness, W.

Note that both C_{gc} and R_i display inverse proportionality with respect to W. It follows that for broadband operation the FET should comprise a thick active layer. It is necessary to reduce N_D to maintain the open channel current at a level which will yield an acceptable breakdown voltage (fig. 5a). It is worth noting that the removal of the gate recess provides the benefit of enhanced V_{DSB} at a given I_{DSS}. It follows that a greater I_{DSS} may be tolerated and a proportionally thicker channel may be adopted. Ultimately, the reduction in doping concentration N_D is limited by the effect on g_{m0}, which varies as the square root of N_D; a small reduction in g_{m0} is tolerable in view of the reduced roll off, but a definite compromise is required between maximum gain (necessitating high N_D) and bandwidth (low N_D). In practice useful concentrations lie in the range $1.5 - 2 \times 10^{17}$ cm^{-3}.

The foregoing suggestions regarding FET engineering by control of process parameters have been verified in the laboratory by thorough evaluation of several batches of devices. Two specific examples are presented in Table 1, together with appropriate material process, d.c. and microwave data.

From fig. 5c, it can be seen that there exists excellent correlation between the doping concentration/channel thickness (as indicated by I_{DSS}) and the microwave performance of the two FETs. In particular, the extended high frequency characteristic of FET 'A' suggests that such a low doped, thick channel device is useful for amplifiers at frequencies up to and beyond 18 GHz.

(b) Merits of different semiconductor materials

In this section, a commonly debated issue will be examined, namely, the age-old topic of Si versus GaAs. Much to the continued irritation of some Si adherents, there are still commentators who persist with the throw-away remark that GaAs devices and

TABLE 1

Comparison of two FETs

Parameter	FET 'A'	FET 'B'
Doping conc. (N_D), cm^{-3}	1.7×10^{17}	3×10^{17}
Gate recess, μm	0	0.05
Open channel current (I_{DSS}), mA	68	37
Transconductance (g_m), mS	26	35
Pinch off (V_p), V	3	1.5
$F_{min}(1)$, dB	2.5	2.0
$G_{ass}(1)$, dB	8	10

(1): $f = 10$ GHz, $I_{DSS} = 10$ mA, $V_{DS} = 3$V.

circuits are five times faster than their Si counterparts because GaAs has an electron mobility five times greater than that of Si. This particular issue will be examined from the point of view of discrete 0.5 μm gate length MESFETs used as small-signal amplifiers.

In fig. 6 are shown three comparisons of the calculated maximum unilateral transducer power gain (MUG) for a typical 0.5 μm FET structure built in Si and in GaAs. In all cases the calculations have been based upon the simplest of transport models, characterised at low fields by an electron mobility (3000 cm^2/volt sec for GaAs and 600 cm^2/volt sec for Si) and at high fields by a saturated drift velocity which is the same in both materials and equal to 10^7 cm/sec. The cases are

(i) fig. 6a—The 'intrinsic' part of the transistor is considered alone with no substrate effects, and no series parasitic resistances.

(ii) fig. 6b—Parasitic resistances in series with the source, drain and gate electrodes, R_S, R_d and R_g respectively, have been added to the intrinsic transistor. Over what one normally takes to be the useful operating frequency range of a 0.5 μm FET, i.e. up to 20 GHz, the Si FET actually exhibits a greater MUG than the GaAs FET by virtue of the different input and output matches.

(iii) fig. 6c—Finally, substrate effects are now included. At typical resistivities and substrate thicknesses, the depletion layer under the gate pad may not reach the reverse side metallization (ground) in the Si device, whereas for GaAs it does. Consequently, parasitic capacitance is larger in the Si case. This is a major cause of gain degradation. Further, by taking an n-layer/buffer-substrate barrier potential of a few tenths of a volt, injection of hot electrons from the channel into the buffer-substrate is allowed for. There is also bulk resistance between the source and drain via a path through the substrate. (A very crude estimate of its value has been made using the one-dimensional formula for bulk resistance. Note that this procedure implicitly assumes a very simple one-dimensional field distribution within the substrate, whereas in an operating FET the field distribution is quite different. This resistance turns out to be quite large

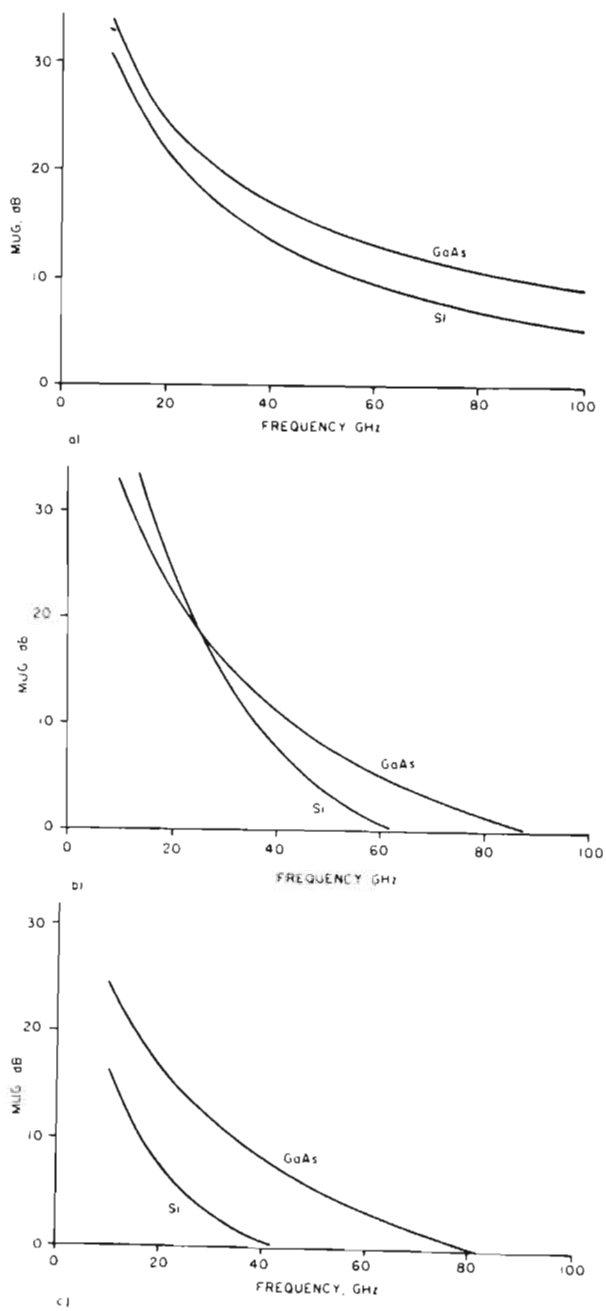


Fig. 6 a) Calculated maximum unilateral gain for the intrinsic part of 0.5 μm Si and GaAs FETs; b) as for a), but with series parasitic resistances R_S, R_d and R_g added; c) calculated MUG of complete Si and GaAs MESFETs, with substrate effects included.

(~kΩ) even for the Si FET). Both mechanisms of currents flow degrade the output conductance of the FET to some degree.

It is worth noting in passing that calculated comparisons of the maximum available gain (MAG)/maximum stable gain (MSG) show similar trends to figs. 6a–6c. MUG is a more appropriate guide to realizable circuit gain, at least in the present context, and so affords a better measure of what can be expected in practical circuits.

Recalling that fig. 6b includes all the regions of the FET structure which mobility affects, it is by no means clear from the figure that the five times higher mobility in GaAs leads to a five times higher speed. In fact, over the useful frequency range of practical 0.5 μm FETs (≤20 GHz), fig. 6b suggests that as much gain in a practical circuit might be available from Si as from GaAs. Fig. 6c, however, indicates a quite different conclusion: it suggests that the substrate plays a major role in degrading the gain of Si MESFETs. At 10 GHz, the GaAs FET offers approximately 8 dB more gain: at a MUG of 15 dB, it offers roughly twice the frequency that the Si FET offers.

The impression is gained, therefore, that any demonstrated superiority of the GaAs FET over the Si FET as a small-signal microwave amplifier (noise considerations apart) is due more to the semi-insulating substance in GaAs than to mobility differences. Further, it must be emphasised that the quantitative improvement GaAs offers is dependent upon device structural features such as gate length, pad sizes, layer doping levels and so forth, and so the 'improvement' factor may be expected to vary from organisation to organisation.

A comparison with independent experiment is possible (fig. 7, after Baechtold et al[6]) with the differences shown in Table 2.
(Note that nowadays a channel doping of $\sim 10^{17}$ cm^{-3}, as used in ref. (6), might be considered somewhat low for a 0.5 μm general purpose MESFET.) From fig. 7 one concludes that the frequency at which the GaAs FET achieves a given gain is about 3–3.5 times that for the Si FET; equivalently, at a given frequency, the GaAs FET offers 7–8 dB more gain than the Si FET. A further vitally important consideration, not dealt with here, is that the GaAs FET generates typically 2 dB less noise[6].

Bearing in mind the effects that unknown detailed differences in doping and structure could have, it is evident from figs. 6c and 7 that, for FETs used as amplifiers, GaAs offers perhaps a factor of two to three improvement in frequency response, not five as the arguments based on mobility ratios suggest.

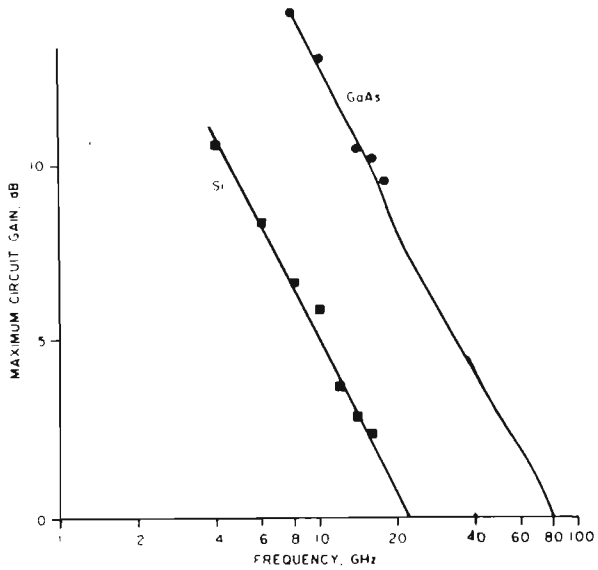


Fig. 7. Measured maximum circuit gain for 0.5 μm Si and GaAs FETs (after ref. 6).

This factor depends upon device geometry (including source-gate spacing and bonding pad sizes) and upon the differences between high-resistivity silicon substrates and semi-insulating GaAs substrates. At shorter gatelengths GaAs may further benefit from electron velocity overshoot[7] which increases the average transit velocity under the gate to values in excess of v_{sat}, but such arguments serve only to highlight the inadequency of the simple view based on mobility ratios.

(c) Sensitivity of circuits to process variables

One of the challenges in MMIC engineering is to interpret, and to correct, any differences between designed and measured scattering (S) parameters of fabricated chips. Leaving aside design errors and inadequacies in design rules, such differences originate from variations in material and process technology—in fact, all those things which go into the practical realization of the integrated circuit.

A logical first step in the problem diagnosis is to substitute an equivalent circuit model for the active devices in the network, and to simulate the circuit behaviour using a network analysis program such as SUPERCOMPACT (indeed, it is highly likely that the design phase involved such a procedure in the first place). The circuit engineer's approach then consists, as a second step, of varying the equivalent circuit elements one-by-one to see what effect each

TABLE 2

	Gate width	Layer doping × thickness
Calculated MUG (fig. 6c)	150 μm	8×10^{11} cm^{-2}
Measured Maximum Circuit Gain (MCG) (fig. 7)	200 μm	$<1.5 \times 10^{12}$ cm^{-2}

has on the scattering parameters of the circuit overall. Although this approach can be successful, it suffers from the disadvantage that just about nothing that one can do to the material or device structure alters just one element in the equivalent circuit: a single material or process change generally alters several equivalent circuit elements, and some alter them all. Applied device physics offers an answer.

Giving recognition to the uncertainties present at all stages in device fabrication, assessment and modelling, the application and proving of the method may be represented as in fig. 8.

It is commonly found that the calculated and measured equivalent circuit element values differ by perhaps 10–20%, depending upon the element. Because there are uncertainties at stages 1 and 2, and at 3 and 5, it is legitimate to adjust the material and device structural data input to the modelling scheme (task 4), within the perceived uncertainty limits attendant to boxes 1 and 2, until the two equivalent circuits agree exactly. Once this is done, the revised material and device structural parameters become the 'standard' values around which variation is imposed.

Having derived an initial equivalent circuit in this way, the effect of varying technological parameters (FET gate length, for instance) on the entire circuit can be assessed. Not only can one determine the effect of variation in any of the material and structural features identified on fig. 2a, together with the effect of changing bias conditions, but much physical insight is generated into the reasons for the changes to circuit behaviour because the manner in which all the equivalent circuit elements are affected is known. Thus, a quick-response diagnostic tool is available to help guide the efforts of materials scientists and process technologists.

As an example of such 'process modelling' in use, fig. 10 forecasts the effect of (a) layer doping variations and, (b) surface potential variations, upon the power gain of the distributed amplifier in fig. 9. These two are of interest because the first is relatively easy to control (at least on a prior measurement basis), whereas surface condition, by and large, still is something of an unknown, and is in any case determined by surface treatment, aging and the presence of any passivating dielectric overlayers. The value of fig. 10b, therefore, is that it gives some idea of how damaging process parameters which are partly controlled might be.

(d) Dependence of chip yield upon technology

Although the potential advantages of monolithic microwave integrated circuits (MMICs) over hybrid realizations are well known, their adoption in systems will ultimately depend upon the chip cost to the end-user. In turn, cost depends upon yield. Reliable estimates of the yield required to give a chip price which is attractive to the user are difficult to make, but a typical figure might be 20–30% for a circuit of average complexity by today's standards. It is a common experience that yields in practice fall short of this goal.

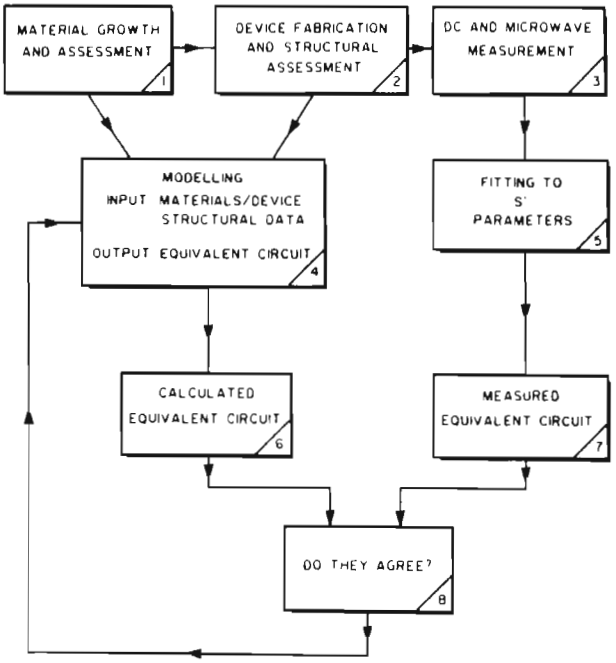


Fig. 8. Procedure followed to establish initial theoretical FET equivalent circuit in process modelling.

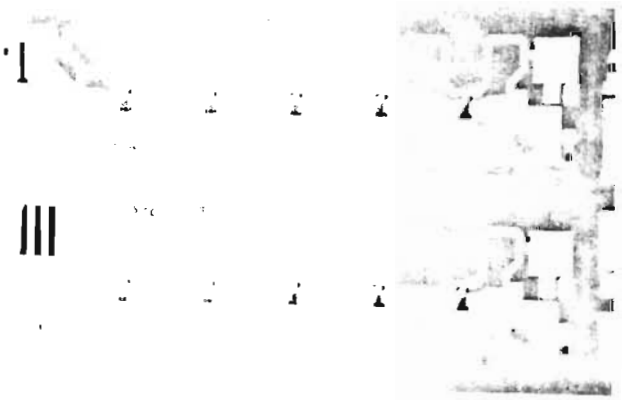


Fig. 9. 5-stage 18 GHz distributed amplifier.

In endeavouring to identify the principal yield-limiting factors, the following distinction can be useful:

$$\text{Overall Yield} = (\text{Mechanical Yield}) \times (\text{Intrinsic Yield})$$

Mechanical yield embraces such factors as wafer breakages, photolithographic lift-off failure and the like. Intrinsic yield is taken to mean the yield of chips within r.f. specification where, given 100% mechanical yield, some chips fail to meet the specification because of tolerances or variations in materials and processing parameters (for example, active layer doping and thickness, FET gate length, ohmic contact resistance, surface charge and so on).

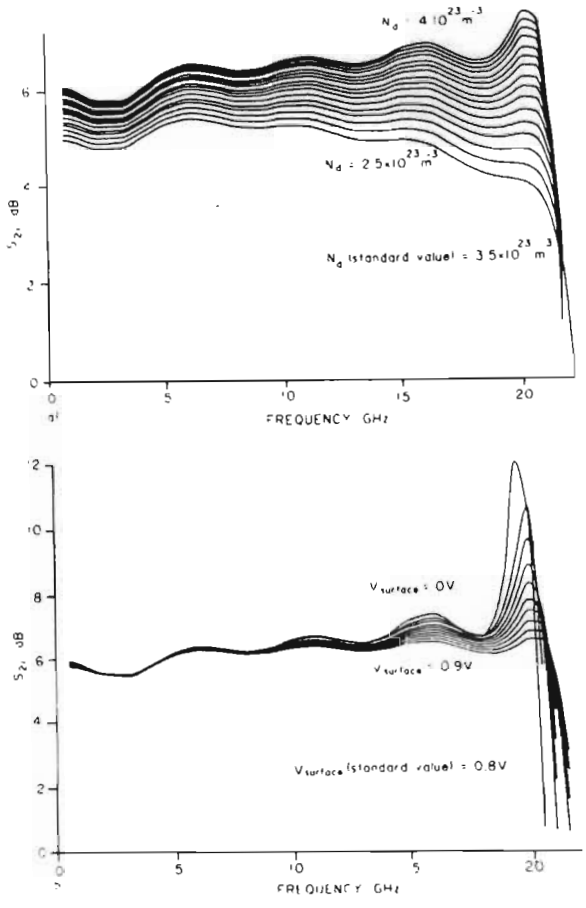


Fig. 10. Effect of variation in (a) layer doping, (b) surface potential upon the magnitude of the forward gain $|S_{21}|$ of the amplifier in Figure 10.

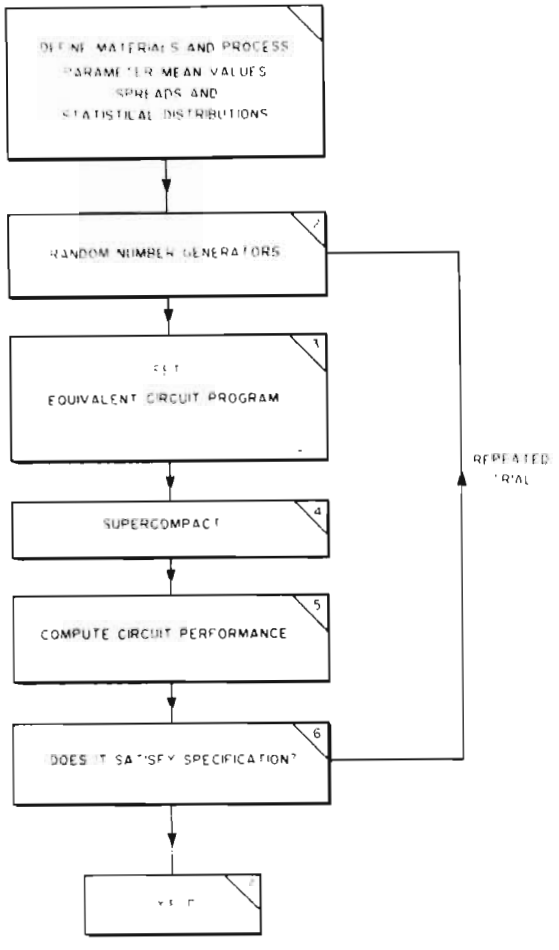


Fig. 11. Flow chart for predicting intrinsic yield.

To determine, in turn, the factors governing intrinsic yield, an extension of the method described in the previous section is employed (fig. 11). Each material and FET structural parameter is varied randomly but in such a manner as to satisfy, over a large enough number of trials, the mean value, deviation in that value and shape or form of the statistical distribution governing that particular parameter in practice. (Thus, a necessary adjunct is to determine these distributions in real life, which is a major undertaking in its own right.) A set of random but statistically constrained values is generated (at box 2 in fig. 11), thereby specifying a set of material and device structural values with which a chip could conceivably have been manufactured. Its performance can then be calculated (steps 3–5) and tested against the target specification (step 6). The intrinsic yield is determined as the ratio of the number of chips which satisfy the specification to the total number of trials. Using such a scheme it should be possible not only to understand present day yields, but also to identify those material and structural features which need to be more tightly controlled to achieve any desired yield.

The following results relate to the 18 GHz travelling wave amplifier of fig. 9 and are based on contrived statistics (so as to protect commercially confidential data):

Distribution assumed	Gaussian
Standard deviation (all relevant parameters)	±10%
Circuit	5 section, 0.5–18 GHz distributed amplifier
Intrinsic yield	≈30%

Dominant contributions to yield figures such as this have been identified, with important ramifications for technological development. Furthermore, two different designs utilizing different numbers of stages have been investigated and found to be different in their sensitivity to the stated spreads, a finding which provides a basis for the concept of 'robustness' of design. Thus, the foregoing software tool has the capability to:

(a) Determine the intrinsic yield of a given circuit design fabricated to a fixed process with known in-process variations.
(b) Ascertain the 'robustness' of alternative circuit designs to the uncontrolled variations in a given process.
(c) Determine the degree of materials and process control needed to achieve any desired yield target.

(d) From (a), and knowing the wafer processing cost and mechanical yield, determine the chip selling price required for economic viability.
(e) From (b), formulate a set of design rules for circuit 'robustness'.
(f) From (c), forecast the investment in capital equipment, time, and human resources necessary to achieve economic viability in GaAs IC ventures.

There may well be further applications for this tool which will become evident as experience is gained using it. For the time being, however, it appears to be the only alternative to the vastly expensive and time-honoured approach of practical trial and iteration.

4. CONCLUSIONS

While recognizing the importance of computer simulation as a method for investigating fundamental physical phenomena in active devices, it is equally recognized that the results of such studies are often not of immediate value to device technologists and circuit engineers. What the latter groups require are active device models in compact analytical form which provide a rapid diagnostic and/or design capability in support of their efforts to fabricate working circuits and devices. It is this need that analytic modelling of active devices is intended to fulfil. The activity relies heavily on good experimental data to provide clues for the formulation of analytical expressions, and it relies also on computer simulation to verify or improve its findings in certain cases. The applications for such a capability are manifold.

A tool has been described which allows forecasts of GaAs MMIC intrinsic yield to be made, depending upon circuit design, complexity and variations in materials and processing parameters. It provides an alternative to the costly and customary approach of practical trial. A first application to two designs for a wideband amplifier has indicated intrinsic yields for 'not atypical' parameter spreads which are different for the two designs. Thus the method offers the opportunity of developing a set of design guidelines, or even perhaps rules, for 'robustness' of circuit design to process variations. A readily foreseen application exists in the general field of economic and investment planning attendant to GaAs initiatives, in both tactical and strategic senses.

ACKNOWLEDGEMENT

The work of colleagues too numerous to mention individually is represented in various ways in this paper. Their support and collaboration in the developments described herein has been especially rewarding. A J Hill gratefully acknowledges the financial support of the University of Natal, Durban, South Africa.

REFERENCES

1 GETREU, I. E., 'Modelling the bipolar transistor', Tektronix Inc., Beaverton, Oregon, 1976.
2 STRATTON, R., 'Diffusion of hot and cold electrons in semiconductor barriers', Phys. Rev., *126*, 6, p. 2002–2014, June 1962.
3 CONWELL, E. M., 'High field transport in semiconductors', Academic Press, New York, 1967.
4 WILLING, H. A., RAUSCHER, C. and DE SANTIS, P., 'A technique for predicting large signal performance of a GaAs MESFET', IEEE MTT-S Internat. Symp. Digest, p. 132–134, Ottawa, Canada, 1978.
5 FUKUI, H., 'Optimal noise figure of microwave GaAs MESFETs', IEEE Trans. Electron. Devices, *ED-26*, 7, p. 1032–1037, July 1979.
6 BAECHTOLD, W., DAETWYLER, K., FORSTER, T., MOHR, T. O., WALTER, W. and WOLF, P., 'Si and GaAs 0.5 μm-gate Schottky-Barrier Field-Effect Transistors', Electronics Letters, *9*, 10, p. 232–234, 17 May, 1973.
7 HUANG, R. S. and LADBROOKE, P. H., 'The physics of excess electron velocity in submicron-channel FETs', J. Appl. Phys., *48*, 11, p. 4791–4798, November, 1977.