ProQuest Number: 10131289

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.

ProQuest 10131289

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346
THE DEVELOPMENT OF AN INDIUM GALLIUM ARSENIDE JUNCTION
FIELD EFFECT TRANSISTOR FOR USE IN OPTICAL RECEIVERS

by

D. WAKE

A thesis submitted to the University of Surrey in candidature for the degree of Doctor of Philosophy.

British Telecom Research Laboratories
Martlesham Heath
Ipswich
Suffolk

Department of Physics
University of Surrey
Guildford
Surrey

© May 1987
Summary

The objective of this work was to design and develop a high performance field effect transistor to be suitable for monolithic integration with a photodetector for use in long wavelength optical communication systems. It was decided that the most promising type of device for this application was a junction field effect transistor (JFET), fabricated using the alloy In\textsubscript{53}Ga\textsubscript{47}As grown epitaxially onto an InP substrate.

The requirements for such a device were that it should have high transconductance, low input capacitance, and low gate leakage current (for high receiver sensitivity), and that it should have a structure which would be easily integrated monolithically with the desired type of photodetector - an In\textsubscript{53}Ga\textsubscript{47}As PIN-photodiode.

Although this alloy semiconductor has favourable electron transport properties, at the start of this work, high performance field effect transistors had not been realised in this material. In particular, the In\textsubscript{53}Ga\textsubscript{47}As FETs that had been made at that time were characterised by low transconductance. Using a device design that incorporated many novel and efficacious features, the JFET described in this work gave results which greatly surpassed all previous (and current) published results of similar devices. This device not only showed high performance, but the novel design features also enabled a simple fabrication scheme.

Having developed this very high performance discrete device, the feasibility of monolithic integration with a In\textsubscript{53}Ga\textsubscript{47}As PIN-photodiode was demonstrated. Although the physical size and material requirements of these two devices were very different, novel design features enabled the construction of a monolithic PIN-FET combination, in which the performance of the JFET was not compromised.
## CONTENTS

1. Introduction

2. Literature survey
   2.1 Discrete In$_{53}$Ga$_{47}$As field effect transistors
      2.1.1 Prior to this work (up to early 1983)
      2.1.2 More recent results (from early 1983 onwards)
   2.2 Monolithic InGaAs PIN-FET literature
   2.3 FET device modelling

3. Discrete InGaAs FET development
   3.1 Design
      3.1.1 Design philosophy
      3.1.2 Modelling
         3.1.2.1 Device modelling
         3.1.2.2 Circuit noise model
      3.1.3 Structural design
   3.2 Fabrication
      3.2.1 Techniques
         3.2.1.1 Molecular beam epitaxy
         3.2.1.2 Photolithography
         3.2.1.3 Metallisation
         3.2.1.4 Lift-off
         3.2.1.5 Etching
      3.2.2 Layer assessment
      3.2.3 Gate formation
      3.2.4 Source and drain formation
      3.2.5 Device isolation
   3.3 Optimisation
      3.3.1 Channel mobility
      3.3.2 Contacts
         3.3.2.1 Source and drain contact
         3.3.2.2 Gate contact
      3.3.3 Rib profile
4. Discrete InGaAs FET performance - basic device
   4.1 Measurement techniques
      4.1.1 dc current-voltage (IV)
      4.1.2 Capacitance-voltage (CV)
      4.1.3 Microwave (S-parameters)
   4.2 Standard device characteristics

5. Variation of JFET characteristics with channel properties
   5.1 Transconductance
      5.1.1 Variation with gate bias
      5.1.2 Variation with gate length and electrode spacing
   5.2 Gate-source capacitance
   5.3 Output conductance
   5.4 Gate leakage current
   5.5 Microwave measurements

6. Monolithic integration of InGaAs JFET and PIN-photodiode
   6.1 Design
   6.2 Performance
      6.2.1 FET performance
      6.2.2 PIN performance
      6.2.3 PIN-FET characterisation

7. Conclusions

References

Acknowledgements

Appendix 1 Device model

Appendix 2 Derivation of figure of merit for PIN-FET design

Appendix 3 Papers published as part of this work
Chapter 1. Introduction

The use of optical transmission systems in telecommunications networks is becoming increasingly widespread throughout the world. In Britain, for example, the whole of the trunk (inter-city) network will use optical systems by 1988, and the first international submarine optical route (Britain-Belgium) is expected to open imminently. A transatlantic optical system is due to open in 1988 (Britain & France - USA), and it is planned that all future international submarine routes will be optical.

The evolution of these systems has tended towards longer unrepeatered spans, and higher transmission rates, in order to increase efficiency and therefore reduce costs. Longer unrepeatered spans are achieved by reducing the fibre attenuation, by increasing the launched optical power from the source, or by increasing the sensitivity of the receiver. First generation optical communications systems operated at a wavelength of 0.85μm, because of the availability of optical sources and detectors at this wavelength. Laser sources used in these systems were fabricated from gallium arsenide (GaAs) based materials, and the detectors were fabricated from silicon (Si). Although these devices could be produced with high power or high sensitivity, the absorption of radiation by silica-based optical fibre is relatively high at this wavelength, which resulted in short unrepeatered transmission distances.

Second generation systems take advantage of the low-loss windows in the spectral response of silica-based fibre, at wavelengths of 1.3, and 1.55μm. For these systems, semiconductor laser sources are constructed using the quaternary alloy system indium gallium arsenide phosphide (In$_{1-x}$Ga$_x$As$_y$P$_{1-y}$), lattice matched to InP substrates. The composition of this InP-lattice matched alloy system can be varied to produce direct bandgap material with an absorption edge which can range between 0.92μm (InP), and 1.67μm (In$_{.53}$Ga$_{.47}$As - henceforth referred to as InGaAs). These lattice matched compositions have the relation $y = 2.2x$, which conveniently fixes the range from $y=0$ (InP) to $y=1$ (InGaAs). For example, the active region of a 1.3μm semiconductor laser has the composition In$_{.73}$Ga$_{.27}$As$_{.60}$P$_{.40}$. 


Detectors in these systems are either constructed from germanium (Ge), or the ternary alloy InGaAs. The latter alternative is generally preferred, due to the high quantum efficiency of this material, especially at 1.55 μm, where the electronic transition in Ge is indirect. InGaAs, on the other hand, has a direct bandgap, and is responsive to radiation up to a wavelength of 1.67 μm, and can therefore detect at both wavelengths of interest with high quantum efficiency.

The type of detector generally constructed from this alloy is the PIN-photodiode. This device consists of a thick undoped 'intrinsic' layer, bounded by a highly doped p-type layer and a highly doped n-type layer (hence the name). The 'intrinsic' layer forms the depletion region of the pn junction, and carriers created in this depletion region from incident radiation are swept out by the high electric field, to give a photocurrent in an external circuit. Since the maximum quantum efficiency (number of electron-hole pairs created per incident photon), is unity, this device must be followed by a low-noise pre-amplifier to provide sufficient gain to give adequate receiver sensitivity. For high sensitivity receivers, this is achieved presently by connecting the InGaAs PIN-photodiode to a GaAs metal-semiconductor field effect transistor (MESFET) in a thick-film hybrid package using a bondwire. Although this approach can give high sensitivity, the inherent parasitic capacitance and inductance of the hybrid technique means that performance is degraded.

Monolithic integration of the photodiode and field effect transistor can minimise the parasitics and thereby increase sensitivity, if the performance of the equivalent discrete devices can be matched or bettered. Other potential advantages of monolithic integration include lower cost and increased reliability, if the analogy with silicon integrated circuits is reasonably valid.

Going on from this stage, it was envisaged that more complicated opto-electronic integrated circuits (OEICs) could be built, which could combine optical components with small-scale or medium-scale electronic circuits, which could ultimately lead to a complete monolithic optical repeater. Initially, greater gains were perceived in the monolithic integration of the receiver, rather than transmitter, because external parasitics play a less dominant role in the latter case.
For these reasons, the aim of this work was to develop a low noise field effect transistor, and then monolithically integrate it with an InGaAs PIN-photodiode.

Obviously, to match the substrate requirements of the photodiode, a field effect transistor had to be developed using InP-based materials. Attempts to fabricate MESFETs using InP were less successful than the equivalent GaAs devices\(^1\). Not only does InP produce a lower barrier height than GaAs, but the tendency of the InP surface to become phosphorus deficient makes the Schottky contact irreproducible. Ohmic contacts to this material are also more difficult to achieve because of this surface problem.

In a comparative study of various semiconductors for high performance FETs, Cappy et al\(^2\) showed that InGaAs had great potential. The band structure of this material is such that electrons in the central \(\Gamma\) valley have low effective mass (0.041\(\bar{m}_0\)\(^3\)), where \(\bar{m}_0\) is the free electron mass). In fact, this material has the lowest electron effective mass in this InP-lattice matched quaternary system. Low-field electron mobility depends upon effective mass and the scattering processes that the electrons undergo, and InGaAs has the highest low-field mobility of this system. At an electron concentration of 10\(^{17}\) cm\(^{-3}\), for example, the mobility of InGaAs is typically 9000 cm\(^2\) V s\(^{-1}\), which compares favourably with GaAs (4500 cm\(^2\) V s\(^{-1}\)), and InP (3500 cm\(^2\) V s\(^{-1}\)).

The energy gap between the central \(\Gamma\) valley and satellite \(L\) valley is relatively large (0.55 eV\(^4\)), and so electrons can be accelerated to high velocities in an applied electric field before being scattered into the high effective mass \(L\) valley. This gives electrons a high average velocity in this material, and values close to 3.10\(^7\) cm/s have been observed\(^5\).

This combination of electronic properties makes InGaAs an attractive material for FETs. InGaAs, therefore, was the obvious material to use for fabricating the FET, partly because of its excellent electronic transport properties, but also because of material compatibility with the PIN-photodiode.

Presently, the vast majority of OEICs are based on GaAs. In these circuits, the transistor technology is already well established, and therefore the major effort in their construction is
in the design of the optical device, and the interconnection scheme. However, for InP-based OEICs, FET technology is still immature, and so their development and emergence has been retarded. To illustrate this difference, small scale and medium scale GaAs microwave monolithic integrated circuits (MMICs) have been available commercially from several sources for the past couple of years, whereas discrete FETs constructed from the quaternary alloy InGaAsP are still in the research phase of their development.

The first step to achieve the objectives of this work was the development of the FET. Due to the immaturity of this type of FET, very little prior knowledge existed. The development of a discrete InGaAs FET therefore presented a considerable challenge, and was considered suitable as the basis of this thesis.

The principal requirements of a field effect transistor for high sensitivity receiver applications are that it should have high transconductance (rate of change of drain current with gate bias at a fixed drain bias), low gate capacitance, and low gate leakage current. An additional requirement for the device in this study was that it should be readily integrable with a PIN-photodiode. Having established designs for discrete devices that satisfied these requirements, an interconnection scheme had to be devised to integrate the devices without degradation of performance.

The layout of this thesis follows this sequence of work in a roughly chronological order. Before any practical work on this topic, a thorough literature search was performed, in order to glean any prior knowledge that existed, and also to put device performance into context. The results of this survey are presented in chapter 2.

Design, fabrication, and optimisation of the discrete InGaAs JFET is discussed in chapter 3, and its performance is described in chapters 4 and 5. Demonstration of the monolithic integration of this device with a PIN-photodiode forms the basis of chapter 6.

Finally, in chapter 7, an assessment is made regarding the status of this work, and how closely the initial objectives were met. Further work towards the development of InP-based OEICs is also discussed in this chapter.
Due to the nature of the work described in this thesis, collaboration with many colleagues at BTRL was necessary. It is appropriate, therefore, to outline my personal contribution to this work, or rather, to indicate the areas of work that were not directly due to me.

Virtually all of the epitaxial growth of material used in this work was carried out in the molecular beam epitaxy group at BTRL. Close liaison between myself and this group was needed to achieve device quality material. For instance, my interpretation of device results enabled feedback of useful information to the material growers, which was used to modify the growth technique to achieve better device material.

Technical assistance was used for some of the device fabrication in this study - i.e. the photolithography and metallisation stages. All other processing stages, and all process development was my undertaking.

Measurements that were performed by others were SIMS profiling and Hall mobility.

Apart from the exceptions listed above, all work contained in this thesis was my endeavour.
Chapter 2. Literature survey

In this chapter, a survey of relevant published work is presented, and is composed of three sections. The first section deals with work on discrete InGaAs field effect transistors, and is further subdivided into literature published prior to the start of this work, and literature published during the course of this work. The second section reviews the literature for monolithically integrated InGaAs PIN-FET results. Finally, in section 3, a review of GaAs FET modelling is presented in order to justify the model used in this study.

2.1 Discrete In$_{53}$Ga$_{47}$As field effect transistors

2.1.1 Prior to this work (up to early 1983)

Historically, the emergence of the quaternary alloy system In$_x$Ga$_{1-x}$As$_y$P$_{1-y}$ lattice matched to InP in the mid-1970s was due to favourable optical properties of this material for optoelectronic devices used in the wavelength range between 1.0 and 1.7µm. In particular, semiconductor laser sources at 1.3 and 1.55µm have been developed for optical communication systems, and photodetectors using the alloy In$_{53}$Ga$_{47}$As have found widespread use for receivers in these systems.

Calculations of the electronic properties of this alloy system were published in 1977 by Littlejohn, Hauser, and Glisson, based on computer simulation of transport processes$^6$. This work highlighted the excellent electron transport properties of this system, especially those of the ternary composition, In$_{53}$Ga$_{47}$As, which showed a twofold increase in mobility over GaAs at a carrier concentration of $10^{17}$cm$^{-3}$, and therefore indicated that this material would be desirable for high speed microwave devices such as the field effect transistor.

The first report of such a device was published in 1980 by Leheny et al$^7$. They described a JFET formed by zinc diffusion through a silicon nitride mask into unintentionally doped LPE-grown InGaAs (a carrier concentration of mid-$10^{16}$cm$^{-3}$). This device was obviously
intended as a demonstration of technology since a long gate length of 20μm precluded high performance. A transconductance of 1mS/mm was achieved.

The JFET approach was chosen because the Schottky barrier height of metals on n-type InGaAs is only 0.2eV[^8], which makes the transfer of the GaAs MESFET technology difficult, without severe problems with gate leakage current. Leakage current is important for optical receiver applications because of the shot noise it produces. High transmission rate systems can tolerate more leakage current due to the differing bit rate dependences of the noise sources. As a rough guide to levels of leakage current that can be tolerated, 100nA produces significant noise at 140Mbit/s, whereas at 1.2Gbit/s, an equivalent noise would be produced by a leakage current of 1μA[^9]. However, no indication of the level of leakage current experienced in this design was given.

However, Ohno et al, later the same year, produced a MESFET on this material by incorporating a thin layer of InAlAs to form a Schottky contact with the gate metal[^9]. The lattice matched composition, In_{52}Al_{48}As, has a direct bandgap of 1.46eV and gave a Schottky barrier height of 0.80eV with a test structure. MBE was used to grow the material for this device, and was also used to grow aluminium in-situ for the gates. Although the device dimensions were not optimised, a maximum transconductance of 57mS/mm was achieved for a gate length of 2.75μm. However, the leakage current was still unacceptably high at 80μA for a gate-drain voltage of -8V.

The same device with a submicron gate was reported soon afterwards by the same team[^10]. The maximum transconductance was 135mS/mm for a gate length of 0.6μm, but again, gate leakage current was high (12μA at a gate-drain voltage of 4V). This result, however, showed the potential of this material, since an electron velocity of 1.8 \times 10^7 cm/s was inferred.

A similar technique was used by Bandy et al to fabricate MESFETs using InP as the Schottky-assist layer[^5]. Because they used liquid phase epitaxy (LPE) to grow the structure, a quaternary anti-meltback layer was included. Electron beam lithography was used to define 0.4μm gates, which gave the device a transconductance of 132mS/mm. Their analysis of the
variation of drain current with depletion width gave a saturated electron velocity of
2.95 \times 10^7 \text{cm/s} for this material. Although they cited the use of the Schottky-assist layer as a
means of reducing gate leakage current, they gave no quantitative results.

As an alternative to the use of a high bandgap semiconductor Schottky assist layer,
O'Connor et al tried a thin layer of silicon nitride, deposited by plasma-enhanced chemical
vapour deposition (PECVD), to achieve the same effect\cite{11}. For a gate length of 1.2\mu m, a
transconductance of 130mS/mm was achieved using this technique. Again, no gate leakage
current results were given for this device.

Chang et al published results of a JFET design which used a grown-in junction, formed by
molecular beam epitaxy (MBE)\cite{12}. The motivation for the JFET approach was low gate
leakage current, although no quantitative results were given. The device was formed by wet
etching of the top p-type layer to define the gate area using PECVD silicon nitride as an etch
mask. Source and drain contacts were spaced 18\mu m apart. For a gate length of 24\mu m, a
transconductance of 50mS/mm was achieved.

The first metal insulator semiconductor field effect transistor (MISFET) in this material
was reported by Wieder et al in 1981\cite{13}. This was an n-channel inversion mode device on a
Zn-doped In_{55}Ga_{47}As epilayer grown by LPE, using PECVD SiO_2 as the gate dielectric.
The performance of this device was poor, showing a transconductance of 0.1mS/mm for a
4\mu m gate length, although improvement was envisaged by development of the dielectric
technology.

Soon afterwards, Liao et al reported a similar device using PECVD Si_3N_4 as the gate
dielectric, and achieved a transconductance of 2mS/mm for a 2\mu m gate length\cite{14}. They
attributed the improvement mainly to better dielectric quality. At the same time, Gardner et al
published results of a MISFET on n-type material grown by vapour phase epitaxy (VPE),
using PECVD SiO_2 for the gate dielectric, and achieved a transconductance of 25mS/mm for
They reported that the SiO$_2$/InGaAs interface was well behaved with an interface state density of $10^{12}$ cm$^{-2}$ eV$^{-1}$.

Liao et al consolidated their work on MISFETs by publishing results on devices using native oxide for the gate dielectric$^{[16]}$. Firstly, a p-channel device on LPE material was fabricated, which had a transconductance of 4 mS/mm for a 3 $\mu$m gate length. The native oxide was grown in an oxygen plasma, and was 45 nm thick. Results of an n-channel device using the same gate dielectric on MBE material were also published$^{[17]}$. Measurement of transconductance using a curve tracer indicated a value of 20 mS/mm, but drift in drain current was observed. When pulsed measurements were made, the transconductance increased to 33 mS/mm. An interface state density of $5 \times 10^{12}$ cm$^{-2}$ eV$^{-1}$ was deduced from capacitance voltage measurements, which was blamed for causing a fivefold reduction in channel mobility.

Ishii et al reported results of a MISFET which used a double-layer gate dielectric of anodic Al$_2$O$_3$/anodic native oxide on LPE grown material$^{[18]}$. This approach was used because of its success on InP, and was formed by the anodisation of a deposited Al layer in an aqueous electrolyte. A transconductance of 17 mS/mm for a 10 $\mu$m gate was achieved, but drift in drain current was observed. Interface state density was calculated to be $2 \times 10^{13}$ cm$^{-2}$ eV$^{-1}$ at the Fermi level.

The first high electron mobility transistor (HEMT) produced using InGaAs was reported by Chen et al in 1982, using a InAlAs/InGaAs heterostructure grown by MBE$^{[19]}$. Hall mobility measurements of the wafer prior to fabrication showed a room temperature mobility of nearly 9000 cm$^2$/Vs, and a 77K mobility of over 60000 cm$^2$/Vs. Device transconductance was 31 mS/mm at 300K and 69 mS/mm at 77K, for a gate length of 5.2 $\mu$m. This relatively poor performance was attributed to having non-optimised InAlAs layer thickness, and to ohmic contacts. Gate leakage current was 30 $\mu$A at -2 volts, which is an unacceptably high level for receiver applications.
In a subsequent publication, this performance was improved by shortening the gate to 1.2 \mu m, and by including a top n\textsuperscript{+} InGaAs contact layer\textsuperscript{[20]}. Also, a InAlAs buffer layer was grown to improve output conductance. Although the Hall mobility of this structure was lower than the previous case, the device transconductance was increased to 70mS/mm at 300K and 125mS/mm at 77K. No specific mention was made to the gate leakage current, and it was assumed to be unchanged from the previous result.

The same group of workers followed this up by publishing results on an improved device which had a transconductance of 90mS/mm at 295K for the same gate length\textsuperscript{[21]}. Analysis of these results, however, suggested that a large proportion of the drain current was carried via bulk conduction. Gate leakage current was given as 45\mu A at -3 volts, which again, is unacceptably high for receiver applications.

A summary of these results in chronological order is shown in table 2.1 below. The work presented in this table shows all of the published material, with one or two minor exceptions, concerning the development of InGaAs as a semiconductor for field effect transistors, covering a time span of two and a half years from mid-1980 to the beginning of 1983, which was approximately the start date of the work covered in this thesis.

Clearly, the MISFET was the most popular device type, although results from this approach were the poorest. The attractions of this device are relative ease of fabrication, simplified circuit design, and potential packing density in a manner analogous to silicon MOSFET technology. This type of device is so successful on silicon due to the well behaved, and easily formed native oxide of this material. However, it appears from the work surveyed above that this situation is not the case for InGaAs. Whether a native oxide or PECVD dielectrics are used, high interface state densities are produced, which lead to low device transconductance and drain current drift. These problems also occur in GaAs MISFETs for the same reasons, and considerable effort to improve these devices over many years has not been successful. The inference from this observation to the case of the InGaAs MISFET was that stable, high performance devices were unlikely to be achieved by this means.
<table>
<thead>
<tr>
<th>Ref.</th>
<th>Device type</th>
<th>L (μm)</th>
<th>$g_m$ (mS/mm)</th>
<th>$I_L$ (μA@V)</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>JFET LPE Zn-diffused</td>
<td>20</td>
<td>1</td>
<td>-</td>
<td>Jun 80</td>
</tr>
<tr>
<td>9</td>
<td>MESFET MBE InAlAs Sch.</td>
<td>2.7</td>
<td>57</td>
<td>80@-8</td>
<td>Aug 80</td>
</tr>
<tr>
<td>10</td>
<td>MESFET MBE AlInAs Sch.</td>
<td>0.6</td>
<td>135</td>
<td>12@-4</td>
<td>Sep 80</td>
</tr>
<tr>
<td>13</td>
<td>MISFET LPE SiO$_2$</td>
<td>4</td>
<td>0.1</td>
<td>-</td>
<td>Mar 81</td>
</tr>
<tr>
<td>5</td>
<td>MESFET LPE InP Sch.</td>
<td>0.4</td>
<td>132</td>
<td>-</td>
<td>May 81</td>
</tr>
<tr>
<td>14</td>
<td>MISFET LPE Si$_3$N$_4$</td>
<td>3</td>
<td>2</td>
<td>-</td>
<td>Nov 81</td>
</tr>
<tr>
<td>15</td>
<td>MISFET VPE SiO$_2$</td>
<td>3</td>
<td>25</td>
<td>-</td>
<td>Dec 81</td>
</tr>
<tr>
<td>11</td>
<td>MESFET MBE Si$_3$N$_4$ Sch.</td>
<td>1.2</td>
<td>130</td>
<td>-</td>
<td>Mar 82</td>
</tr>
<tr>
<td>12</td>
<td>JFET MBE grown jn.</td>
<td>2</td>
<td>50</td>
<td>-</td>
<td>Mar 82</td>
</tr>
<tr>
<td>16</td>
<td>MISFET LPE native ox.</td>
<td>3</td>
<td>4</td>
<td>-</td>
<td>Jun 82</td>
</tr>
<tr>
<td>19</td>
<td>HEMT MBE InAlAs</td>
<td>5.2</td>
<td>31</td>
<td>30@-2</td>
<td>Jun 82</td>
</tr>
<tr>
<td>17</td>
<td>MISFET MBE native ox.</td>
<td>4.5</td>
<td>33</td>
<td>-</td>
<td>Aug 82</td>
</tr>
<tr>
<td>20</td>
<td>HEMT MBE InAlAs</td>
<td>1.2</td>
<td>70</td>
<td>-</td>
<td>Aug 82</td>
</tr>
<tr>
<td>18</td>
<td>MISFET LPE Al$_2$O$_3$/nat.</td>
<td>10</td>
<td>17</td>
<td>-</td>
<td>Nov 82</td>
</tr>
<tr>
<td>21</td>
<td>HEMT MBE InAlAs</td>
<td>1.2</td>
<td>90</td>
<td>45@-3</td>
<td>Jan 83</td>
</tr>
</tbody>
</table>

Table 2.1 Discrete InGaAs FET publications up to early 1983.

The MESFET, using either a high bandgap semiconductor or thin insulator to give a Schottky gate, was the next most popular approach. Although this type of device gave the highest performance in terms of transconductance, the level of gate leakage current which these techniques were meant to minimise were too high for low-noise operation. Half of the papers describing this approach did not even refer to this aspect explicitly.
Only two of the papers that were published in this time period dealt with discrete JFETs. Neither gave details of gate leakage current, and showed middling performance. Although this type of device should give the lowest levels of gate leakage current (after the MISFET) due to its higher built-in potential, it suffers from the disadvantage that it is usually a more difficult device to fabricate. This is reflected in the initial lack of interest in this device.

The HEMT is potentially a very attractive device, although the first attempts using the InAlAs/InGaAs heterointerface did not fulfil this potential. This type of device, however, is by far the most difficult to produce. Very tight control of layer thickness and doping, coupled with the need for an extremely abrupt heterointerface, poses considerable difficulty for the crystal grower. Added to that, ohmic contact to the 2-deg region is not easy, and this can reduce the device performance drastically. These limitations account for the late start of this type of device.

Of the epitaxial techniques available, LPE was the most commonly used in this survey. This reflects the fact that not many groups had access to more sophisticated epitaxial techniques such as MBE, during this time period. LPE is not ideally suited to producing thin, uniform layers with the precision required, and would not have been used as a first choice.

Having assessed the literature, it was decided to adopt the JFET approach for the present work. The reasons for this choice are given below;

* the poor performance and instability shown by the MISFET was considered to be a very difficult problem to overcome,
* the MESFET approach gave consistently high levels of gate leakage current,
* the HEMT was considered to be too great a challenge to the epitaxy available for this work.
The literature discussed in the next section was published during the course of this work, and serves to put the present work into context, and acts as an indicator of relative performance.

2.1.2 More recent results (from early 1983 onwards)

Work on MISFETs continued to develop, with high values of transconductance being reported. Wieder et al reported preliminary results of an accumulation mode transistor made using Fe-compensated LPE-grown InGaAs, with free electron concentration of $2.5 \times 10^{12} \text{cm}^{-3}$ \cite{22}. Gate dielectric was PECVD SiO$_2$. For a gate length of 8\,\mu m, a transconductance of approximately 10\,mS/mm was achieved, which was two orders of magnitude higher than their previously published work.

Cheng et al produced a self-aligned, recessed gate MISFET on MBE material with a submicron gate length (0.5\,\mu m), which gave transconductance in excess of 250\,mS/mm\cite{23}. This depletion mode device used PECVD Si$_3$N$_4$ for the gate dielectric, and an angled evaporation technique to form the submicron gate. However, observation of the source-drain characteristic showed that it was far from ideal, with high output conductance and incomplete pinch off.

Another promising result was published by Itoh and Ohata\cite{24}. Their enhancement mode MISFET showed a transconductance of 120\,mS/mm for a 1\,\mu m gate length. The average electron velocity was estimated to be greater than $2.5 \times 10^7 \text{cm/s}$ in this device, which used PECVD SiO$_2$ as the gate dielectric on VPE material. Results of microwave performance were given, and showed a 5.6dB gain at 12\,GHz and an extrapolated cut-off frequency of 24\,GHz.

Work on the MESFET also continued to develop, based on the high bandgap Schottky-assist layer principle. Chen et al described a MESFET which had a very thin, highly doped channel (sheet charge field effect transistor), in order to give high transconductance\cite{25}. 


However, poor ohmic contact limited this to 80mS/mm for a 2μm gate. InAlAs was used as the Schottky-assist layer, but no leakage current details were given.

Cheng et al reported a MESFET using an electron beam evaporated silicon oxide Schottky-assist layer in a self aligned, recessed gate structure\[26\]. Reduction of series resistance by these techniques gave devices with a transconductance of 150mS/mm for a gate length of 1.5μm. Gate leakage current was reported to be less than 10μA at -2 V.

A combination of the insulator/high bandgap semiconductor Schottky-assist layer techniques was used by Chang et al\[27\]. In this device, an aluminium oxide barrier, formed by oxidation of an in-situ deposited aluminium layer, was used in conjunction with a layer of InAlAs to reduce leakage current. Transconductance of this MESFET was 130mS/mm for a 2μm gate length. Although leakage current was low on test structures, the leakage current of this device was not given directly.

A new approach to this problem was published by Chen et al, where a mismatched layer of GaAs was used to increase the apparent Schottky barrier height\[28\]. Although good morphology was obtained despite the large mismatch (3.7%), leakage current was very high at 200μA for a bias of -1V, rising to 700μA at -3V. Transconductance was 108mS/mm for a gate length of 1.4μm. In a subsequent paper from the same team, the cause of the high leakage current was identified to be due to misfit dislocations at the GaAs/InGaAs interface\[29\]. These dislocations were shown to be confined within a short distance of the interface, therefore by growing a thicker GaAs layer, leakage current was reduced to 0.48μA at -1V, and 3.4μA at -2V. Transconductance of this device was approximately the same as the previous device.

A MESFET grown by MOVPE was described by Scott et al, which used InAlAs as the Schottky-assist layer\[30\]. This device had a gate leakage current of 10μA at -10V, and a transconductance of 50mS/mm for a gate length of 4μm.
The first result of a HEMT using the InP/InGaAs heterointerface was published by Takikawa et al\cite{31}, and they achieved comparable device transconductance to previous InAlAs/InGaAs HEMT results (90mS/mm for a 2μm gate). Gate leakage current of 5μA at -3V was reported.

The true potential of the HEMT was realised by Hirose et al\cite{32} and Itoh et al\cite{33}, in separate papers presented at the International Symposium on GaAs and related compounds in 1985. Both devices had a 1μm gate length, and used the InAlAs/InGaAs system, grown by MBE. Itoh et al achieved a transconductance of 250mS/mm, but the best result was given by Hirose, who presented a device having a transconductance of 440mS/mm at room temperature, rising to 700mS/mm at 77K. These results, especially the latter, demonstrate the advantages gained from the HEMT approach.

A variation of the HEMT was reported by Seo et al, who described results of a single quantum well field effect transistor, grown by MBE\cite{34}. The 100 Å well was formed in a InAlAs/InGaAs/InAlAs heterostructure, and showed a transconductance of 130mS/mm for a 1.8μm gate length. Gate leakage current was given as 3μA at -3V for this device.

Interest in the JFET was greatly increased during this time period. Chai et al published a design that enabled submicron gates to be produced\cite{35,36}. In this design, Be ion implantation was used to form a pn junction in gate regions. After metallisation, the gate metal was then used as an etch mask to form a rib of p-type semiconductor, which formed the gate. Etching proceeded in steps until the transconductance was observed to decrease. In this way, a gate length of 0.5μm was achieved, which gave a transconductance of 85mS/mm. Gate leakage current in this design was 10μA at a drain voltage of 3V.

An improved design was published by Wake et al, as part of this work, where the T-shaped undercut gate rib was used to self align source and drain contacts\cite{37}. This publication can be seen in appendix 3. This device used a grown-in junction formed by MBE. Although the contacts and channel mobility had not been optimised, a
transconductance of 60mS/mm was reported for a 3µm gate length. Another novel feature of this design was the use of air bridges to isolate the gate bondpads. Gate leakage current was given as 100nA at -3V, which was the lowest value so far reported for any InGaAs FET device. For the first time, an InGaAs FET had been produced that had a gate leakage current low enough for optical receiver applications.

Cheng et al used the same design with a p-type InGaAs buffer layer, and achieved a transconductance of 90mS/mm for a 1µm gate length\[38\]. A contact was made to the buffer layer, which gave some control of device characteristics. Gate leakage current was 200-400nA at -4V. The same team later published results of the same JFET design on n+ InP substrates using either an MBE-grown InAlAs or MOVPE-grown Fe-compensated InP buffer layer\[39\]. Transconductance of 70-80mS/mm was achieved for both types of device. Although the InP buffer was shown to be more effective because of its superior semi-insulating properties, device uniformity was poor due to the two-stage epitaxy used in this design.

A similar gate rib approach was taken by Schmitt and Heime, although self alignment was not used, where the pn junction was formed by zinc diffusion from a spin-on source\[40\]. Using a gate length of 1.2µm, a transconductance of 100mS/mm was achieved. Microwave measurements on this device gave an extrapolated frequency limit of 30GHz.

A JFET using a p-InP buffer layer was reported by Wake et al\[41\], as part of this work, and can be seen in appendix 3. This device was grown by MOVPE, and used the self alignment technique described previously\[37\]. The p-InP buffer layer was shown to be effective in achieving high electron velocity in the channel (a saturated electron velocity of 2.7 \(10^7\)cm/s was inferred), which lead to a device transconductance of 210mS/mm for a 1.5µm gate length. Gate leakage current, however, was excessive (15µA at -1V), due to a large density of interface states in the high field region of the gate.
Selders et al used an undoped InP buffer layer to achieve the same effect. Using VPE, they obtained a carrier concentration of $10^{13}\text{cm}^{-3}$ in the buffer layer. The JFETs were formed using Be ion-implantation in a structure with a gate length of 1.6\textmu m, and a source-drain spacing of 4\textmu m. Although they noted improved performance over similar devices without buffer layers, the maximum transconductance obtained was only 130mS/mm.

Selders et al tried the GaAs surface layer approach to produce JFETs and MESFETs. GaAs layers were grown on LPE-grown n-InGaAs FET channels by MOVPE (n-type for MESFETs and p-type for JFETs). Device results were poor, however, with a transconductance of 105mS/mm for the JFET (gate length of 4\textmu m) and 100mS/mm for the MESFET (gate length of 1.6\textmu m). The JFET showed incomplete pinch-off, and the MESFET showed high gate leakage current (not given directly).

The results discussed above are summarised in table 2.2. Papers are listed chronologically, and those not specifically discussed above have also been included. It can be seen from this table that less interest was shown in the MISFET approach than previously. The problem of drain current drift due to high interface state densities had not been effectively overcome, although high device transconductance had been achieved by Cheng et al. However, this MISFET suffered from incomplete pinch-off and high output conductance, and was therefore not a good device.

The MESFET continued to attract attention, using new approaches to the Schottky-assist principle. The lattice mismatched GaAs gate layer tried by Chen, Selders, and Garbinski did not fulfil the promise of low leakage current operation, however.

The leakage current reported for the latter was probably the lowest, but at 3-4\textmu A is still far too high for low noise operation, except, perhaps, for very high bit-rate applications.

The HEMT was shown to give excellent performance with the publication of a device with a 440mS/mm transconductance at room temperature.
<table>
<thead>
<tr>
<th>Ref.</th>
<th>Device type</th>
<th>L</th>
<th>$g_m$</th>
<th>$I_L$</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>μm</td>
<td>mS/mm</td>
<td>μA@V</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>JFET MBE Be-implanted</td>
<td>0.5</td>
<td>85</td>
<td>10@-3</td>
<td>Jul 83</td>
</tr>
<tr>
<td>22</td>
<td>MISFET LPE SiO$_2$</td>
<td>8</td>
<td>10</td>
<td>-</td>
<td>Aug 83</td>
</tr>
<tr>
<td>25</td>
<td>MESFET MBE InAlAs</td>
<td>2</td>
<td>80</td>
<td>-</td>
<td>Sep 83</td>
</tr>
<tr>
<td>31</td>
<td>HEMT VPE InP/InGaAs</td>
<td>2</td>
<td>90</td>
<td>5@-3</td>
<td>Mar 84</td>
</tr>
<tr>
<td>23</td>
<td>MISFET MBE Si$_3$N$_4$</td>
<td>0.5</td>
<td>250</td>
<td>-</td>
<td>May 84</td>
</tr>
<tr>
<td>37</td>
<td>JFET MBE grown-jn</td>
<td>3</td>
<td>60</td>
<td>0.1@-3</td>
<td>Jul 84</td>
</tr>
<tr>
<td>26</td>
<td>MESFET MBE SiO$_2$</td>
<td>1.5</td>
<td>150</td>
<td>10@-2</td>
<td>Dec 84</td>
</tr>
<tr>
<td>24</td>
<td>MISFET VPE SiO$_2$</td>
<td>1</td>
<td>120</td>
<td>-</td>
<td>KOBE 84</td>
</tr>
<tr>
<td>27</td>
<td>MESFET MBE Al$_2$O$_3$/InAlAs</td>
<td>2</td>
<td>130</td>
<td>-</td>
<td>IEDM 84</td>
</tr>
<tr>
<td>28</td>
<td>MESFET MBE GaAs</td>
<td>1.4</td>
<td>108</td>
<td>700@-3</td>
<td>Jan 85</td>
</tr>
<tr>
<td>38</td>
<td>JFET MBE p-buffer</td>
<td>1</td>
<td>90</td>
<td>0.2@-4</td>
<td>May 85</td>
</tr>
<tr>
<td>40</td>
<td>JFET LPE Zn-diffused</td>
<td>1.2</td>
<td>100</td>
<td>-</td>
<td>May 85</td>
</tr>
<tr>
<td>39</td>
<td>JFET MBE/MOVPE</td>
<td>1</td>
<td>70</td>
<td>1@-2</td>
<td>Jul 85</td>
</tr>
<tr>
<td>32</td>
<td>HEMT MBE InAlAs/InGaAs</td>
<td>1</td>
<td>440</td>
<td>-</td>
<td>GaAs 85</td>
</tr>
<tr>
<td>33</td>
<td>HEMT MBE InAlAs/InGaAs</td>
<td>1</td>
<td>250</td>
<td>2@-3</td>
<td>GaAs 85</td>
</tr>
<tr>
<td>30</td>
<td>MESFET MOVPE InAlAs</td>
<td>4</td>
<td>50</td>
<td>10@-10</td>
<td>GaAs 85</td>
</tr>
<tr>
<td>41</td>
<td>JFET MOVPE p-buffer</td>
<td>1.5</td>
<td>210</td>
<td>15@-1</td>
<td>Dec 85</td>
</tr>
<tr>
<td>34</td>
<td>HEMT MBE SQW InAlAs</td>
<td>1.8</td>
<td>130</td>
<td>3@-3</td>
<td>Dec 85</td>
</tr>
<tr>
<td>42</td>
<td>JFET LPE/MOVPE GaAs</td>
<td>4</td>
<td>105</td>
<td>-</td>
<td>Jan 86</td>
</tr>
<tr>
<td>42</td>
<td>MESFET LPE/MOVPE GaAs</td>
<td>1.6</td>
<td>100</td>
<td>-</td>
<td>Jan 86</td>
</tr>
<tr>
<td></td>
<td>JFET LPE Zn-diffused</td>
<td></td>
<td>105</td>
<td>.01@-1</td>
<td>Feb 86</td>
</tr>
<tr>
<td>29</td>
<td>MESFET MBE GaAs</td>
<td>1.4</td>
<td>104</td>
<td>3.4@-2</td>
<td>Feb 86</td>
</tr>
<tr>
<td>43</td>
<td>JFET VPE InP buffer</td>
<td>1.6</td>
<td>130</td>
<td>-</td>
<td>Mar 86</td>
</tr>
</tbody>
</table>

Table 2.2 Discrete InGaAs FET publications from early 1983 onwards.
Results published as part of this work are in bold type.
However, this performance is critically dependent on having good quality epitaxy with tight control of layer thickness and doping, and the ability to produce very abrupt interfaces. If this epitaxy is available routinely, then the HEMT becomes an extremely attractive device.

Most of the publications in this time period were concerned with JFETs. Innovations in the design of this type of device lead to high performance levels. The rib gate design, first published by Chai et al\[^{35}\], enabled submicron gate lengths to be achieved in JFETs for the first time. As part of this study, Wake et al took this idea a stage further by using the gate rib for self-alignment of the source/drain contacts\[^{37}\]. This approach, coupled with improvements in channel properties due to more effective buffer layers, lead to high levels of transconductance. The highest value reported was 210 mS/mm, by Wake et al\[^{41}\], again, as part of this study.

The conclusions from this survey are as follows:

* the highest transconductance is offered by the HEMT, but very high quality epitaxy is needed for this performance,
* the JFET gives adequate transconductance and the lowest levels of gate leakage current, and requires less demanding epitaxy than the HEMT,
* the MESFET should be capable of achieving adequate transconductance, but gate leakage current has been unacceptably high to date,
* the MISFET is unattractive due to problems with instability as a result of high interface state densities found so far.

2.2 Monolithic InGaAs PIN-FET literature

The primary interest in InGaAs FETs was for the potential advantages to be gained by the monolithic integration of these devices in PIN-FET optical receivers used for long wavelength systems. Not surprisingly, therefore, the first report of such a structure was published at around the same time as the first discrete FET result, by the same team of workers, i.e., Leheny et al\[^{44}\], using LPE grown material.
This structure was obviously intended as a preliminary demonstration, since the device dimensions that were used precluded high performance. The FET part of the structure was the same design as their discrete device\cite{7}, i.e., a JFET which used zinc diffusion through a silicon nitride mask to form the gate region. This device had a transconductance of 1mS/mm for a 20\textmu m long gate. The PIN photodiode was formed as an integral part of the JFET gate electrode, and so no independent measurements could be made. Photoresponse was observed, but was not fully characterised. An improved result using the same design was subsequently reported by the same team, in which the JFET transconductance was increased to 60mS/mm by reduction of the gate length to 6\textmu m\cite{45}. The PIN/gate combination exhibited a leakage current of 40nA at -3V, and a total input capacitance of 1pF at this bias. No photoresponse measurements were given, however.

A similar structure incorporating a planar PIN photodiode was reported by Inoue et al\cite{46}, again using LPE material. Sequential zinc diffusions through a SiO$_2$ mask formed the p$^+$ device isolation layers, and the JFET gate/PIN diode regions. The JFET showed a transconductance of 30mS/mm for a 5\textmu m gate length, and the combined leakage current was 200nA at -5V. Total input capacitance for this structure was 2pF at -3V, and 1.5pF at -5V. Simple photoresponse measurements at 1.3\textmu m were performed for demonstration purposes only.

An improved structure was reported subsequently by the same team\cite{47}. Maximum transconductance was increased to 40mS/mm, and leakage current was reduced to 20nA at -5V by passivating the junction with polyimide. More comprehensive photoresponse measurements were made on this structure. The quantum efficiency was given as 60\% over the wavelength range 1.0 - 1.6\textmu m, and a sensitivity of -25dBm was measured for a bit error rate of 10$^{-9}$ at 400Mbit/s, although the operating wavelength was not given. For comparison, a typical hybrid receiver would give a sensitivity of around -37dBm at this bit rate.

Hata et al published results of a planar PIN/JFET formed by Be ion implantation into a single layer of LPE material\cite{48}. The JFET employed a p-column gate which controlled the drain
current by the radial and axial spread of the depletion region around the columns. This gate consisted of 130 columns, each 4μm square, separated by 2μm. Again, the PIN diode was incorporated as part of the JFET gate as in previous structures discussed above. Transconductance was 13mS/mm, and combined leakage was 50nA at -5V. No capacitance data was given, and photoresponse measurements were arbitrary.

The basic limitation in the designs outlined above is that a single epitaxial layer is used for both the PIN photodiode and FET channel. The intrinsic layer of the PIN must be thick to absorb most of the incident light, and have a low carrier concentration to give low capacitance and low leakage current. Conversely, the FET channel layer must be thin to give optimum pinch-off voltage, and moderately doped to provide high transconductance for sufficient gain to give low noise operation. Since these designs use the same layer for both devices, the performance of both must be compromised. To overcome this limitation, Wake et al published a design, as part of this work, that used separate layers for each device in a vertically integrated structure[49]. This is reproduced in appendix 3. In this design, the FET layers were grown on top of the PIN layers by MBE. The JFET design was essentially the same as reported for the discrete device[41], except all layers were InGaAs, including the p-buffer layer. This layer also acted as the p-contact layer for the PIN photodiode. Another novel feature in this design was the use of air bridges to isolate the devices, so that each could be operated and biased independently. The other designs had incorporated the PIN diode into the gate of the FET, and therefore bias conditions could not be set separately in these cases. Maximum transconductance was 170mS/mm, which is the highest reported figure for a monolithically integrated InGaAs FET. However leakage current was high (10μA at -2V for the FET and 400nA at -5V for the PIN). Total input capacitance was 0.8pF at operating bias, comprising 0.3pF for the FET gate and 0.5pF for the PIN.

Photoresponse was demonstrated at 1.55μm wavelength, but was intended for demonstration purposes only, since results were arbitrary.

A different approach to the problem was to integrate an InGaAs PIN-photodiode with an InP MISFET. Although this type of structure is not directly relevant to the work considered in
In this context, inclusion of these monolithic structures is useful for comparison and completeness.

Kasahara et al. were the first to publish results on this type of structure, which used a hybrid growth technique; LPE for the FET contact layer, and VPE for the PIN absorbing layer. Zinc diffusion was used to form the PIN junction, and device definition was performed by mesa etching. PIN capacitance was 0.5pF at -6V, and the leakage current was 200nA at -5V. The MIFSFT used PECVD SiO₂ as the gate dielectric, and achieved a transconductance of 12mS/mm for a 3.5μm gate length. Gate capacitance was 0.6pF for this device. Interconnection was achieved by a metal track across the dielectric layer, which contributed stray capacitance. Photoresponse measurements were performed at 1.3μm, and a sensitivity of -30.6dBm was achieved at a bit error rate of 10⁻⁹ for 100Mbit/s operation. For comparison, at this bit rate, the sensitivity of a hybrid receiver would be typically -45dBm.

The same team improved on this result in a subsequent publication. By reducing the gate length of the MISFET to 1.4μm, the transconductance was increased to 45mS/mm, and leakage in the PIN was reduced to 3nA at -5V. These improvements lead to a receiver sensitivity of -34.5dBm under the same measurement conditions as previously.

Tell et al. reported a InGaAs PIN - InP MISFET receiver which included a PIN bias resistor. In this structure, the PIN absorbing layer was grown by VPE in channels formed in the InP substrate, and selective ion implantation of Be was used to form the pn junction. Leakage currents less than 25nA at -5V were achieved by this technique. The MISFET channel was formed by ion implantation of Si into the semi-insulating InP substrate. Silicon nitride was used as the gate dielectric, which was self aligned to the source and drain contacts. Gate length was 3μm, which gave a transconductance of 75mS/mm. The PIN bias resistor was also formed during the Si ion implantation stage, and had a value of 1.5kΩ. Total input capacitance of this structure was 1.5 pF, of which over 1pF was contributed by the FET. Photoresponse measurements were performed using a 1.54μm laser, at bit rates of 90Mbit/s and 295Mbit/s, and extrapolated sensitivities of -34dBm and -29.5dBm were given for a bit error rate of 10⁻⁹. Typically, a hybrid receiver would have
given sensitivities of -46dBm and -39dBm respectively. In their analysis, this performance was dominated by the noise from the PIN bias resistor.

Recently, Ohtsuka et al. reported a low leakage current PIN integrated with an InP MISFET\(^5^3\). The PIN layers were grown by LPE on a stepless substrate, and cadmium diffusion formed the pn junction. However, in order to achieve low leakage current operation, the mask dielectric (silicon nitride) had to be removed, and the surface passivated by polyimide. A leakage current of 2nA at -5V was given for this device. The MISFET was formed on the semi-insulating substrate using SiO\(_2\) as the gate dielectric, and had a transconductance of 19mS/mm for a gate length of 3\(\mu\)m. Interconnection of devices was achieved by running a metal track down the side of the PIN mesa. Although this mesa was 4\(\mu\)m high, interconnection of devices could still be made, but yield and repeatability must be questioned. No capacitance measurements were given, and the only photoresponse quoted was a PIN responsivity of 0.5A/W over most of the wavelength range.

This work is summarised in table 2.3, which shows the relevant results from the publications in chronological order. This work basically falls into three categories. Firstly, the majority of publications have been concerned with using a single layer of InGaAs to form the PIN and FET (invariably JFET). Inevitably, the performance of either device (or possibly both) was compromised by this arrangement. For example, the highest JFET transconductance using this approach has been 60mS/mm\(^4^5\). The most important figure of merit for a PIN-FET combination is its sensitivity. However, only one publication in this category gave such a result, but this was not too encouraging\(^4^7\).

The second approach was the integration of an InGaAs PIN with an InP MISFET. This type of structure gave more promising results because each device could be optimised separately. For instance, high FET transconductance and low PIN capacitance can be achieved more easily with this approach (e.g. 75mS/mm and 0.15pF respectively), which lead to higher sensitivities being achieved\(^5^2\). However, the InP MISFET has not shown as high a performance as the InGaAs JFET.
<table>
<thead>
<tr>
<th>Ref.</th>
<th>Device type</th>
<th>$g_m$ [mS/mm]</th>
<th>$C_T$ [pF]</th>
<th>$I_L$ [$\mu A$]</th>
<th>$S$ [dBm]</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>PIN-InGaAs JFET</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>80</td>
</tr>
<tr>
<td>45</td>
<td>PIN-InGaAs JFET</td>
<td>60</td>
<td>1.0</td>
<td>0.04</td>
<td>-</td>
<td>81</td>
</tr>
<tr>
<td>46</td>
<td>PIN-InGaAs JFET</td>
<td>30</td>
<td>2.0</td>
<td>0.2</td>
<td>-</td>
<td>83</td>
</tr>
<tr>
<td>50</td>
<td>PIN-InP MISFET</td>
<td>12</td>
<td>1.4</td>
<td>0.2</td>
<td>-30.6 (100)</td>
<td>83</td>
</tr>
<tr>
<td>51</td>
<td>PIN-InP MISFET</td>
<td>45</td>
<td>0.9</td>
<td>0.003</td>
<td>-34.5 (100)</td>
<td>84</td>
</tr>
<tr>
<td>48</td>
<td>PIN-InGaAs JFET</td>
<td>13</td>
<td>-</td>
<td>0.05</td>
<td>-</td>
<td>84</td>
</tr>
<tr>
<td>47</td>
<td>PIN-InGaAs JFET</td>
<td>15</td>
<td>1.7</td>
<td>0.02</td>
<td>-25 (400)</td>
<td>85</td>
</tr>
<tr>
<td>52</td>
<td>PIN-InP MISFET</td>
<td>75</td>
<td>0.6</td>
<td>0.025</td>
<td>-34 (90)</td>
<td>85</td>
</tr>
<tr>
<td>53</td>
<td>PIN-InP MISFET</td>
<td>19</td>
<td>-</td>
<td>0.002</td>
<td>-</td>
<td>86</td>
</tr>
<tr>
<td>49</td>
<td>PIN-InGaAs JFET</td>
<td>170</td>
<td>0.8</td>
<td>10</td>
<td>-</td>
<td>86</td>
</tr>
</tbody>
</table>

Table 2.3 Monolithic PIN-FET publications. Bit-rate for sensitivity, $S$, is given in brackets. The result in bold type was published as part of this work.

The best of both approaches has been combined for the third category of device, in which only one result has been published\cite{49}. This is the result of Wake et al which is reported as part of this thesis, and can be seen in appendix 3. This type of device incorporates an InGaAs PIN and an InGaAs JFET in a vertical integration scheme, so that each device can be optimised independently. Although high transconductance was achieved (170 mS/mm), no sensitivity measurements were taken. However, this approach shows the greatest promise, if the performance of the hybrid receiver is to be bettered.
2.3 FET device modelling

Modelling of field effect transistors has attracted a great deal of attention since their inception by Shockley in 1952\[^{54}\]. Several excellent reviews of this subject have been published, e.g., by Liechti\[^{55}\], and by Pucel, Haus, and Statz\[^{56}\], therefore only a brief description of this work is given here. Also, since device modelling was of secondary importance to this work, this brevity is doubly justified.

The JFET is a unipolar device, i.e., majority carrier flow is predominant. Basically, the JFET consists of a channel of n-type semiconductor (electron mobility is significantly greater than hole mobility in III-V semiconductors), on which an ohmic contact is formed at each end, i.e., the source and drain electrodes. The gate electrode is formed between the source and drain, by inclusion of a region of p-type semiconductor. A schematic of a planar style JFET is shown in figure 2.1.

![Figure 2.1 Basic structure of planar JFET.](image)

At low drain bias, this structure acts as a simple voltage-controlled resistor, and the current is proportional to the drain bias. As the gate bias is made more negative, the depletion region under the gate extends into the channel, thereby constricting the carrier flow. The drain current therefore decreases until pinch-off is reached, when the channel is fully constricted, and the drain current drops to zero. However, due to the field dependent mobility of semiconductors, as the drain bias is increased, the mobility in the channel decreases, and the
current saturates. In III-V semiconductors, this effect occurs at low fields, and is due to carriers being scattered from the low effective mass central conduction band valley into the high effective mass satellite valleys. As the gate is made more negative, the level of current saturation is decreased, until the device becomes pinched-off.

In Shockley's analysis of the JFET, the longitudinal and transverse fields in the device are treated independently. Furthermore, it is assumed that the extent of the gate depletion region changes gradually from source to drain, i.e., that the longitudinal field is negligible compared with the transverse field. This assumption is generally referred to as the 'gradual channel approximation', and limits the applicability of this model to operation below current saturation. Since only the transverse field is considered in this model, the depth of the depletion region at any point can be determined by a one dimensional solution of Poisson's equation, if the channel potential at that point is known. An analytical solution describing the variation of the channel current with gate and drain bias may be obtained by using this solution of Poisson's equation in conjunction with Ohm's law in the channel.

Another limitation of this model is that the mobility is assumed to be constant, which is only true at low electric fields. Consequently, this model only provides a good description of long gate FETs prior to current saturation.

This basic model was later extended by Dacey and Ross to include the effects of field dependent mobility, which they suggested was responsible for current saturation[57]. They used experimental mobility-field data for germanium, in which the mobility was observed to decrease as the half power of the field, above a critical field, i.e.,

\[ \mu = \mu_0 \left( \frac{E_c}{E} \right)^{1/2} \]

\[ \text{...... 2.1} \]

where \( \mu_0 \) is the low field mobility and \( E_c \) is the critical field. Turner and Wilson were the first to analyse the implications of drift velocity saturation on the performance of GaAs FETs[58]. They postulated a finite channel width at the drain end of the gate at the onset of current saturation, which remained pinned at this position. Variation of drain current with
gate bias was then accommodated by changes in this width. The gradual channel approximation of Shockley was used in this analysis, with constant mobility. Lehovec and Zuleeg\textsuperscript{[59]} later modified this with an analytical form for the drift velocity, based on the field dependent electron mobility proposed by Trofimenkoff\textsuperscript{[60]}. This velocity is given by,

$$
u = \mu_0 \cdot \frac{E}{1 + \mu_0 \frac{E}{\nu_s}}$$

\hspace{1cm} \text{... 2.2}

where $\nu_s$ is the saturated drift velocity. A different approach to this problem was taken by Grebene and Ghandhi, who produced a two-region model of the FET, based on a two-piece linear approximation of the velocity-field characteristic\textsuperscript{[61]}. Below a critical field, $E_c$, the mobility was assumed constant, and the velocity varied linearly with the field, whereas above $E_c$, the velocity was constant (the actual form of the velocity field characteristic is shown in figure 2.2, towards the end of this section). The FET channel was divided into two regions based on this approximation. In region I, towards the source end of the channel, the field is below $E_c$ and therefore Ohm's law is obeyed. Consequently, Shockley's gradual channel approximation could therefore be used. In region II, towards the drain end of the channel, the field is above $E_c$, and therefore saturated-velocity current flow exists. An added complication is that the gradual channel approximation is not valid in this region due to the two-dimensional field distribution caused by the large longitudinal field. Operation in this region can be described by a two-dimensional solution of Poisson's equation. With suitable boundary conditions and the assumption of a rectangular geometry, solution of this equation was given by a hyperbolic sinusoid functional form. Solution in this region was matched to the solution in region I to give device characteristics. An important aspect of this model was that the position of the interface between regions could vary with bias conditions, so that as the drain bias increased beyond saturation, region II could extend towards the source.

This principal was extended by Pucel, Haus, and Statz, who produced a complete theory which included small signal and noise analysis\textsuperscript{[56]}. This treatise formed the basis of the device model adopted for the present work, and is discussed more fully in appendix 1.
Wang and Hsieh modified this model by including the effects of velocity overshoot\textsuperscript{[62]}. In short channel FETs, i.e., submicron dimensions, the electron transit time can be sufficiently short so that equilibrium transport is not reached. In these circumstances, the electron velocity overshoots the equilibrium velocity, and is dependent on gate length. In GaAs, these effects become important for gate lengths below 1\mu m. In the limit, if the gate is sufficiently short, electrons can transit the device without collision, and no velocity saturation occurs. Such behaviour is known as ballistic transport. Velocity overshoot phenomena were first analysed by Ruch\textsuperscript{[63]}, and later by Maloney and Frey\textsuperscript{[64]}, who used the Monte Carlo method to study the electron dynamics. Wang and Hsieh used the latter results to give values of effective velocity and threshold field for a range of gate lengths, and inserted these values into the Pucel, Haus and Statz model. They also included a third region in this model, between the gate and the drain, to account for the extension of the gate depletion layer in this region, although they concluded that this region gave little effect. This model attempted to overcome the limitations of the previous analytical models, which proved inaccurate at short gate lengths. This was partly due to the non-stationary electron dynamics in this regime, but was also because the simplifying approximations that are necessary with the analytical approach become less valid.

Although Wang and Hsieh went some way to overcoming this limitation, accurate modelling of short gate FETs is the domain of the numerical computer simulation, as first used by Kennedy and O'Brien\textsuperscript{[65]}. Generally, in the more refined examples of this technique, electron dynamics are determined by a Monte Carlo simulation of the material of interest. The FET channel is divided into a two dimensional array of cells, in which the fundamental equations used to describe electron flow and potential distribution are solved in an iterative manner until some convergence criteria are met to achieve consistency. These are Poisson's equation and the current continuity equation, i.e.,

\[
\nabla^2 \psi = - \frac{q}{\varepsilon} (N_d - n) \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{ ...... 2.3}
\]

\[
q \frac{dn}{dt} = \nabla \cdot \mathbf{J} \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{ ...... 2.4}
\]
where $\psi$ is the potential, $n$ is the electron density, and $J$ is the current density. Although this type of model is capable of giving the potential distribution, and electron density distribution in two dimensions over a range of bias conditions, the whole process is very time consuming. Firstly, the development of such a computer program is not a trivial task, and in itself has often been the basis of PhD theses. Secondly, the resources needed in terms of computer hardware and computing time preclude this type of model for most applications. A single current voltage characteristic can take up to a month to produce by this method, for example.

The analytical methods discussed earlier are generally more useful to the device designer. For device design, qualitative information regarding trends is usually sufficient, and fine detail is often superfluous. Also, all models, to some extent, represent idealised structures, with abrupt impurity profiles and perfect contacts. For instance, the presence of a substrate has been shown to drastically alter the device characteristics\cite{ref6}. Therefore the precise nature of the substrate or buffer layer will play an important part in shaping the device characteristics.

![Figure 2.2 Velocity-field characteristic of GaAs and InGaAs (after Littlejohn, Hauser and Glisson - ref. 6)](image)

Figure 2.2 Velocity-field characteristic of GaAs and InGaAs (after Littlejohn, Hauser and Glisson - ref. 6)
All of the models discussed above have been concerned with either Si or GaAs (or in some cases, InP). It has been assumed in this work that the models developed for GaAs MESFETs will apply equally well to InGaAs JFETs, since the form of the velocity field characteristic for both materials is similar. This is illustrated in figure 2.2. Average saturated drift velocity and threshold field differ, and velocity overshoot phenomena occur for different fields and gate lengths in each material. Therefore, for the design aspects of this work, the Pucel, Haus, and Statz model was modified for the material properties of InGaAs. However, the limitations of this approach were such that detailed comparison with device characteristics with this model were considered to be inappropriate.
Chapter 3. Discrete InGaAs FET development

The development of a high performance discrete InGaAs field effect transistor has been split into three main areas, those being design, fabrication, and optimisation. These areas overlap to some extent, and the development process is essentially iterative. For clarity of presentation, however, this chapter is composed of three sections, each covering one aspect of device development. The first section deals with design, and covers design philosophy, modelling, and structure. The second section outlines the fabrication of the device, and finally, the third section discusses the steps taken to optimise the first crude devices, into the high performance transistors produced latterly.

3.1 Design

3.1.1 Design philosophy

Of the many possible types of field effect transistor, the one most suited to the present application had to be evaluated. The literature survey, outlined in section 2.1, showed the JFET approach to be the most promising. The reasons for this are reiterated below.

The GaAs metal semiconductor field effect transistor (MESFET) used extensively for low noise microwave applications and also for the hybrid PIN-FET, relies for its success on the moderately high and reproducible Schottky barrier formed between most metals and n-type GaAs, which forms the gate of such a transistor. A direct transfer of this technology to InGaAs is unsuitable, unfortunately, because the Schottky barrier height formed between metals and n-type InGaAs is only 0.2 eV. Various workers have sought ways to enhance this barrier height; by insertion of a thin layer of insulator or higher bandgap semiconductor between the metal and FET channel, but these approaches suffered from large gate leakage currents. This type of approach was therefore considered unsuitable, because gate leakage current introduces shot noise which degrades sensitivity in optical receivers.

The metal insulator semiconductor field effect transistor (MISFET), which is used for the majority of silicon digital circuits, relies on the reproducibly good quality native oxide grown
on the semiconductor for gate regions of the device. Again, a direct transfer of this
technology to InGaAs is impractical because of the poor quality of the native oxide for this
semiconductor. A more controlled approach, using other insulators such as silicon nitride,
has not been easy to achieve with good results. Problems with high interface state
density and irreproducibility ruled out this type of device for this application.

More recently, the high electron mobility transistor (HEMT) has shown great promise.
This type of device relies on a very tightly controlled modulation-doped heterostructure to
form a sheet of electrons with high charge density, in a material with low doping level (and
hence high mobility). Although this has shown to be feasible in a InGaAs system (e.g.,
InGaAs/InAlAs), the additional complexities involved with material growth precluded
this type of device.

The junction field effect transistor (JFET), uses a pn junction to form the transistor gate.
For a material with a small bandgap, the JFET is advantageous because the potential barrier
formed by the pn junction is nearly equal to the full bandgap, giving relatively low leakage
current operation.

For these reasons, the JFET was considered to be the most practical type of transistor for
the optical receiver.

In order to construct a JFET, a reliable method of producing small area localised pn
junctions had to be found. Diffusion of p-type impurities through a dielectric mask into
epitaxial n-type material is a well tried technique for localised pn junction formation. For III-
V semiconductors, zinc is the usual choice of acceptor. However, earlier experience with
zinc diffusion into InGaAs for PIN photodetector fabrication showed it to be irreproducible.
Precise control of junction depth proved to be extremely difficult to achieve. Another
disadvantage in using this technique for pn junction formation is gate length broadening due
to lateral diffusion along the edges of the dielectric mask.

A pn junction formed by epitaxy (grown-in junction), can be very accurately controlled,
especially with molecular beam epitaxy (MBE). Very precise positional control, with
relatively abrupt doping profiles can be obtained with ease using this epitaxial technique, although some dopant redistribution during growth can occur (see section 3.3).

For the above reasons, a JFET using a grown-in junction was the preferred type of device for this work.

3.1.2 Modelling

From an early stage it was recognised that a device model would be needed to tailor the design of the FET to the specific application of PIN-FET optical receiver. The GaAs MESFETs used in hybrid optical receivers had been designed for use as conventional low-noise microwave amplifiers, and performance advantages could be obtained with specifically designed devices.

The device model was used in conjunction with a circuit model to provide parameters for the initial FET design. These parameters were optimised for maximum receiver sensitivity within constraints imposed by voltage supply limitations and noise feedback from second stage amplification.

3.1.2.1 Device modelling

The requirements of the FET model used in this work were that it should provide a rough guide to the design of a high performance device, and that it should be easy to use. Therefore, a simple approach that could give qualitative results, and could be used on a desktop computer was sought.

A review of FET modelling is given in chapter 2. Obviously, the numerical computer simulation approach was inappropriate for this application, due to the long and complex program development required, and the need for access to large mainframe computers for large amounts of time. The analytic model developed by Pucel, Haus, and Statz was recognised to give good agreement with GaAs MESFETs having gate lengths greater than 1micron, and was considered as a suitable basis for modelling InGaAs JFETs with similar gate dimensions. The use of a model developed for GaAs was considered feasible due to the similarity of the electron transport properties of this material to those of InGaAs, i.e., both materials show negative differential mobility in their velocity field characteristics. It was
assumed that if a piecewise velocity field characteristic gave reasonable agreement for GaAs devices, then the same should also apply to InGaAs.

In this treatment, described more fully in appendix 1, expressions for small signal and noise parameters are derived from device dimensions, channel carrier concentration, and bias levels. Solution of the highly non-linear equations used in this derivation were performed using a HP9836 desktop computer. A BASIC program was written as part of this work, which used the bi-section method for this purpose. The output from this model was used to provide parameters for the circuit model described below.

3.1.2.2 Circuit noise model

The circuit diagram of the front end of a high impedance optical receiver is shown below in figure 3.1a. The bias resistor, $R_b$, has a very large resistance in order to minimise its noise contribution. This, in conjunction with the FET input capacitance, gives an integrating amplifier, the effect of which is equalised later in the receiver circuit. A simplified signal equivalent circuit of this receiver is shown in figure 3.1b, where

- $C_p$ is the PIN capacitance,
- $R_g$ is the FET gate metallisation resistance,
- $C_{gs}$ is the FET gate-source capacitance,
- $R_s$ is the FET source-channel resistance,
- $g_m v$ is the FET current generator, and,
- $r_{ds}$ is the FET output resistance.

The noise equivalent circuit is shown in figure 3.1c, where the noise current generators have the following origins:

- $i_p$ is the shot noise contribution of the PIN and FET gate leakage currents,
- $e_m$ is the thermal noise contribution of the gate metal and source series resistances,
- $i_g$ is the FET induced gate noise;
- $i_b$ is the thermal noise component due to the bias resistor,
Figure 3.1a  Circuit diagram of high impedance front-end receiver

Figure 3.1b  Signal equivalent circuit

Figure 3.1c  Noise equivalent circuit
\( i_d \) is the FET channel noise, and,

\( i_t \) is the thermal noise component due to the load resistor.

In order to derive an expression for the signal to noise ratio of this circuit, the noise sources must be transformed and lumped together as a single noise generator. Therefore, all noise sources were transformed into current generators at the input of the FET, and added to give a total noise current. Appendix 2 shows how this was achieved. N.B. In this noise analysis, only the noise sources directly influenced by FET design have been considered.

From appendix 2, the total mean square noise current of this circuit is given by;

\[
\bar{i}_t^2 = 4kT \Delta f \psi (\omega C_t)^2 \quad \frac{\bar{g}_m}{g_m} \quad \ldots \quad 3.1
\]

when considered at the FET input. \( \psi \) is a dimensionless noise coefficient, and is derived in appendix 2. The signal to noise ratio, and therefore sensitivity, is inversely proportional to this noise current. Therefore, a figure of merit, \( F \), can be defined for the circuit;

\[
F = \frac{g_m}{(C_t^2 \psi)} \quad \ldots \quad 3.2
\]

The figure of merit of a FET used in a conventional microwave amplifier is inversely proportional to the input capacitance, rather than its square, which gives a significant difference to the device design.

The channel carrier concentration and device geometry were optimised for a given set of bias conditions and circuit component values by using the expressions from the Pucel, Haus and Statz model to obtain signal and noise parameters, which were then used to maximise this figure of merit. This optimisation was subject to the following circuit constraints imposed by limitations of supply voltage, or second stage noise:

i) a limited supply voltage means that the voltage dropped across the load resistor must also be limited. A supply voltage of 5 volts, and a FET drain voltage of 3 volts, gives the restriction

\[
I_d R_L < 2
\]
ii) in order to minimise the feedback of second stage noise into the front-end of the receiver circuit, the gain must have a minimum value, of roughly five, i.e.,

\[ g_m R_1 > 5. \]

The device and circuit models, together with the constraints outlined above were combined in one BASIC program, which gave the variation in figure of merit as a function of device design. As an example, figure 3.2 shows the variation of the figure of merit with channel carrier concentration at constant pinch-off voltage. The conclusion from this figure was that a low channel carrier concentration would lead to high receiver sensitivity. The minimum gain restriction gives a limit to how low the transconductance can be set, which in turn dictates the minimum channel carrier concentration. This turns out to be approximately \( 5 \times 10^{16} \text{ cm}^{-3} \).

![Figure 3.2 Variation of figure of merit with channel carrier concentration at constant pinch-off voltage.](image)

Having set the channel carrier concentration, the channel thickness may be set from the practical range of threshold voltage. A thickness of 0.25\( \mu \)m gives a threshold voltage of \(-1.6\) volts, which is a practical choice from supply voltage limitations.
Since a short gate leads to high transconductance and low gate capacitance, then this dimension should be reduced as far as possible. Commercial GaAs MESFETs are now available with gate lengths down to 0.25\(\mu\)m, but this size is difficult to achieve without expensive and sophisticated lithography techniques (e.g., electron beam or X-ray). A more practical target for the smallest feature size using the lithography equipment available (a deep UV projection mask aligner), was set at 2 \(\mu\)m, which can be achieved routinely without much difficulty. By using the design approach described in section 3.1.3, the gate length produced from this feature size is of the order of 1 \(\mu\)m. In principle, smaller gate lengths can be achieved, but yield becomes a problem at this level, due to the difficulties encountered during photolithography.

To a first approximation, transconductance and capacitance scale linearly with gate width. Therefore to maximise the figure of merit, a narrow width is required. However, the width can not be too narrow, because a minimum gain must be achieved. This constraint sets a lower limit for the transconductance, because the load resistor can not have too high a value because of supply voltage limitations. A width of 250\(\mu\)m was adequate to meet this requirement.

The device parameters derived above were used as a basis for wafer specification and photolithographic mask design.

3.1.3 Structural design

A schematic cross section of the InGaAs JFET structure is shown in figure 3.3. Detailed fabrication of this device will be discussed in the next section. Some of the more important design features are outlined here.

The gate of the JFET is formed by etching the top p-type layer using the gate metal as a mask, thereby leaving a rib of p-type material. This rib defines the extent of the gate, and has the important feature of being smaller than the metal mask due to undercutting.
Since short gate length is a prerequisite for high performance FETs, any process that helps to reduce this is desirable. This is especially so since practical photolithography has a lower bound to the feature size that can be reproduced.

The gate metal and rib together form a T-shaped structure. This gives the opportunity to self-align the source and drain electrodes to the gate, by directly evaporating the source/drain metal over this T-shaped structure. The gate metal acts as a shadow mask in this technique. This self-alignment technique ensures that the photolithography need not be a critical process, and therefore, device yield is high. Another consequence of this technique is that the spacing between source and gate is determined by the degree of undercut achieved when the gate rib is etched. This spacing is roughly the same as the top p-type layer thickness (0.3μm). Such a close source-gate spacing ensures that the associated parasitic resistance is kept to a minimum. The resistance between the source and channel degrades the transconductance of the FET and also acts as a source of noise.

Another important aspect of structural design is in the separation of the gate bondpads from the gate. In many devices that are isolated by mesa etching, the gate metal is simply run down the mesa side to connect with a remote bondpad. This introduces a yield problem due to the need for a routinely achievable smooth and shallow angled mesa edge. In this design, the gate and its bondpad are separated by an air bridge, which is formed by undercutting a
narrow link section between the device mesa and bondpad mesa using a selective etchant that does not etch the substrate. A scanning electron micrograph of an air bridge is shown in figure 3.4. This technique gives excellent isolation, which is important because the bondpad would otherwise contribute excess capacitance and leakage current to the gate. Yield is good because the metallisation remains planar, and the air bridge has been shown to be reproducible and stable - in test structures, stable spans of 400\mu m have been produced.

Figure 3.4 Scanning electron micrograph of air bridge
3.2 Fabrication

3.2.1 Techniques

3.2.1.1 Molecular beam epitaxy

Molecular beam epitaxy (MBE), is a technique where epitaxial growth occurs when one or more molecular beams are incident on a suitable substrate. For the growth of InGaAs for example, three cells are aligned with their orifices to the substrate holder, one containing indium, another containing gallium, and the third containing arsenic. The whole assembly is housed in a stainless steel ultra-high vacuum system (10^-11 torr), and the cells are heated until the material they contain is emitted in the form of molecular beams. In this pressure regime, the mean free path of the molecules is greater than the source to substrate separation, and therefore molecular flow predominates. This is significantly different from vapour phase epitaxial techniques, which operate at or close to atmospheric pressure, where the mean free path is very much less than the source to substrate separation, and therefore viscous flow predominates. The flux from each cell is controlled by its temperature. The three beams are arranged to be co-incident on the substrate, and growth of single crystal InGaAs occurs if the conditions are favourable. Cells containing silicon and beryllium are included so that n-type and p-type material can be produced. Because of the molecular flow regime, the molecular beams can be turned on and off using a simple shuttering arrangement, giving the possibility of extremely abrupt heterointerfaces.

Although most of the published literature detail the advantages of this technique, it is worth pointing out some of the difficulties that have prevented MBE from providing on a routine basis, the high quality of material necessary for this work.

To grow the correct alloy composition, the ratio of indium to gallium flux must be maintained constant. Also, to grow at a controlled rate, the total group III flux must be fixed. The arsenic flux must be sufficient to allow stoichiometry, and in practice this flux is much higher to minimise group V depletion at the growing surface. Control of beam flux is accomplished by a spot calibration using an ionisation gauge, while the cell temperature is fixed. However, drift in flux occurs as the material in the cell is depleted, and variations can
occur in some cases when the cell material redistributes itself. The flux measurement is not accurate enough to rely on for control of layer composition. Fine tuning of layer composition is achieved by feedback of information from the slice grown previously. This information is usually an X-ray measurement of the lattice spacing of the epilayer, from which the indium to gallium ratio can be deduced. Once this fine tuning has been done, then growth of the correct composition can continue with only the occasional adjustment of cell temperature. However, it takes over a week to get to this stage once the system has been at atmosphere.

No matter how well the cells are aligned, the cosine beam profiles do not overlap sufficiently at the substrate for uniform growth, so consequently, many MBE systems use a substrate holder that rotates. This substrate rotation produces material that is compositionally uniform everywhere, but thickness variation still occurs (the layer thins out towards the edge of the substrate).

Another critical growth parameter is the substrate temperature. If the growth temperature, \( T_g \), is too low, then the mobility of the adsorbed atoms is reduced, and point defects tend to be frozen in. As a consequence of this, crystal quality is poor. Alternatively, if \( T_g \) is too high, then the group V element is preferentially lost from the surface, again resulting in poor crystal quality. Also dopant redistribution can occur if the temperature is too high. Obviously, the need to maintain the substrate at the optimum temperature is very important for good quality material growth. Problems arise because the \( T_g \) 'window' is narrow for InGaAs, and also because the absolute temperature is difficult to measure. This is especially so with a rotating substrate, where physical connection of a thermocouple is impractical.

Layer morphology is another important aspect of material quality. Surface irregularities can affect photolithography, and if defects are incorporated in sensitive device areas then performance can be degraded, sometimes catastrophically. Careful substrate preparation can avoid defect generation, but often, many defects grow without known cause.

Summarising, MBE is a very useful epitaxial technique, with the potential of being able to produce high quality material with very abrupt interfaces. However, there are many
problems, and the output of good quality material can be irregular in practise. However, it must be emphasised that MBE is a developing technique, and many of the problems discussed above have been minimised with recent work.

3.2.1.2 Photolithography

The following techniques (photolithography, metallisation and lift-off), are standard methods for device and circuit fabrication. However, there are many parameters that have to be optimised for each of these techniques, in order that they work for a particular mask aligner, evaporator, or type of wafer. Therefore, considerable effort has been expended, as part of this work, in the development of these techniques.

Photolithography is a process for defining patterns on a wafer. The wafer is covered with a photosensitive coating (photoresist), which is applied by spinning on from a solution. This is exposed to UV light through a chrome patterned glass mask. For a positive photoresist, the exposed areas are removed using a developer, and the converse is true for a negative photoresist.

For most applications, several levels of masking are necessary to complete a device, and each level has to be very carefully registered with the previous level. The exposure to UV light and the mask level registration is achieved using a mask aligner. In the simplest type of mask aligner, the mask plate is held in contact with the wafer prior to exposure, and is known as a contact aligner. Alternatively, in a projection aligner, the mask plate and wafer are some distance apart, and a series of lenses is used to focus the pattern from the mask plate onto the wafer.

Each type of aligner has advantages and disadvantages. Briefly, the contact aligner can replicate smaller features, but contact can break fragile wafers and mask plates have short lives. Since the slices used in the fabrication of III-V semiconductor devices tend to be very fragile (especially if the epilayers are mismatched), a projection aligner was used for this work.
The ambient conditions in which photolithography takes place are very important. Lighting must be filtered to exclude UV components, so that the resist is not exposed by room lighting. Also, the air must be filtered to remove as many particulates as possible, since dust particles are generally larger than the features to be defined on the wafer, and any contamination will reduce yield. Finally, since the photoresists are sensitive to the water content of the atmosphere, the relative humidity and temperature must be controlled.

The wafers must be baked two or three times during this process, depending on the application. Firstly, a high temperature pre-bake (150°C) drives off any water adsorbed onto the surface, otherwise resist adhesion may degrade. Secondly, a lower temperature bake (90°C) is given after the resist has been spun onto the wafer in order to drive off excess solvent. Finally, another high temperature bake (150°C) must be given after development if the resist is to be used as a wet etch mask, which has a hardening effect on the resist.

In order to remove any residues which may be left on the wafer after photolithography, a process known as 'ashing' is used. This subjects the wafer surface to a low power oxygen plasma, which oxidises the residue (primarily carbon oxides since residues are organic in nature). Since these residues are volatile, the wafer surface is cleaned by this process.

3.2.1.3 Metallisation

All metallisation for the devices fabricated in this work was deposited by thermal evaporation. In this technique the material to be evaporated is inserted into a tungsten spiral or boat, which is heated by passing an electric current through it. This process occurs in a vacuum system, usually at pressures below 10⁻⁵ torr. To obtain an accurate thickness of film, a quartz monitor is placed near to the wafer. As this monitor becomes coated, the frequency of resonance changes, which can be calibrated in terms of film thickness.

3.2.1.4 Lift-off

Lift-off is a process for patterning metallisation, which avoids the necessity to etch the metal film. Firstly, an overhanging resist profile is formed during the photolithography process, as shown in figure 3.5a.
Figure 3.5 Schematic representation of lift-off process.

a. overhanging resist profile
b. metallisation evaporated
c. excess metal 'lifted-off'

Next, the required metallisation is evaporated onto this structure, as shown in figure 3.5b. Lastly, the resist is dissolved, lifting off the excess metal, as shown in figure 3.5c. Care must be taken to ensure that the wafer is perpendicular to the flux of evaporant, and that the source to wafer distance is large enough to give near normal incidence, otherwise metal could be deposited on the side of the resist edge, which would prevent the resist from being dissolved.

There are various ways to achieve the desired overhanging resist profile. One method is to soak the resist in chlorobenzene before development. This hardens the top layers of resist, and makes the top develop slower than the bottom, thus producing a lip on the resist profile. However, a good lift-off profile is critically dependent on the soak time, and the chlorobenzene temperature and water content, and reproducibility is difficult.

A better way is to use a resist that has been specially designed to give lift-off profiles with the type of mask aligner used. For example, Hunts W118 resist has been produced to give good lift-off profiles (concave walls) with projection mask aligners.
An alternative method is to use a multilevel resist technique. In its simplest form, two different types of resist are spun onto the wafer sequentially. The resists used must not mix, and they should be sensitive to different wavelengths of UV light. The top resist layer is exposed and developed in the normal way, and is then used as a mask for the exposure and development of the bottom layer. Lateral development of the bottom layer ensures that the top resist layer overhangs the bottom. This technique gives excellent results, but is very time consuming.

3.2.1.5 Etching

All etching of InGaAs in this work used aqueous solutions (wet etching). The requirements needed for etchants used in this work are that they leave a smooth surface, do not etch InP (for mesa and air bridge formation), and give re-entrant etch profiles without trenching effects at mask edges (for rib formation). Also, a moderate etch rate is required for fine control of rib height. The most widely used etchants for InGaAs are aqueous mixtures of sulphuric acid (H₂SO₄) and hydrogen peroxide (H₂O₂), in which the peroxide acts as an oxidising agent and the acid is used to dissolve the oxide formed. Etchant composition can be varied to give the required properties for a particular application. Compositions with a high sulphuric acid content are more viscous, and the etch rate tends to be limited by mass transport i.e., by the rate of supply of fresh etching species or the rate of removal of reaction products. These etchants are very sensitive to stirring, and are generally used for polishing, due to their smooth etching characteristics. However, this type of etchant gives abnormally high etch rates adjacent to mask edges, and therefore give rise to trenching.

Less viscous etchant compositions tend to be kinetically limited, i.e., by the chemical reactions at the etched surface. This type of etching mechanism is usually highly anisotropic, and can be used for obtaining orientation dependent cross sections which prove to be useful in semiconductor processing.

The standard etchant used for both rib and mesa etching in this work is composed of 5 parts H₂O₂, 1 part H₂SO₄, and 50 parts H₂O, or simply 5:1:50. The notation used gives the peroxide content first, then the acid, and finally the water. When used at 20°C, this etchant has an etch rate of 0.45 μm/min for InGaAs, and leaves the surface smooth. The etching
mechanism is kinetically limited, which gives differing etch profiles with gate direction. A more dilute etchant (1:1:50) is used to clean the InGaAs surface during some of the processing stages. A dip in this solution cleans the surface and removes a negligible amount of material (<100Å).

3.2.2 Layer assessment
Firstly, the morphology of the wafer is assessed by observation under a microscope. Too high a defect density will degrade device yield, or could affect device performance adversely. Morphological features can be divided into small scale roughness and larger scale discrete defects. A grainy texture of the surface indicates poor crystal quality, and the cause is often arsenic deficiency. Large scale discrete defects come in many varieties, the effects of which are difficult to quantify.

A common defect in MBE material is known as the 'oval' defect because of its shape, and has been shown to have a gallium rich nucleus. Such a defect, which can measure 50 by 20 μm, causes catastrophic damage if it occurs in the gate region of a device.

Wafers that have a grainy texture, or a high defect density are generally not used for device processing.

The thickness and doping level of the epilayers are assessed using a test mask to define contact patterns by lift-off. Using the first level, two 100 by 100μm metal pads, separated by 20μm, are defined on the epilayer surface, as shown in figure 3.6a. The current voltage characteristic between these two pads can be used to assess the doping level of the top p-type layer. If this is less than approximately $10^{19}$ cm$^{-3}$, then the metal semiconductor interface becomes a Schottky contact, and only leakage current flows between the metal pads. If the doping level is higher than this, then the interface becomes ohmic (tunneling Schottky), and a current flows which is proportional to the conductivity, and therefore thickness, of the top layer. Therefore, only material with a top layer that forms an ohmic contact can be used for device fabrication.

The top layer is etched in steps, using the metal pads as etch masks, and the current flowing is monitored after each etch step, until it drops to a very low level (usually 1μA or until no further decrease is observed). This current is the leakage current due to two back to
back pn junctions. At this stage, the top p-type layer has been fully removed as shown in figure 3.6b. Use of a \textit{Dektak} surface profiler to give step height before and after etching allows the thickness of the top layer to be found.

The second level of the test mask is then used to place further metal patterns adjacent to those of level 1. These contact pads are then on the surface of the channel layer, one of

\begin{center}
\includegraphics[width=0.8\textwidth]{figure3.6a.png}
\end{center}

Figure 3.6a  Schematic cross section of layer structure with test pattern

\begin{center}
\includegraphics[width=0.8\textwidth]{figure3.6b.png}
\end{center}

Figure 3.6b  Top p$^+$ layer fully etched using contact pads as etch masks

\begin{center}
\includegraphics[width=0.8\textwidth]{figure3.6c.png}
\end{center}

Figure 3.6c  Test structure complete with channel contact pad
which is shown in figure 3.6c. The p⁺n structure formed in this manner can be CV profiled (capacitance voltage) to give the channel doping and thickness, and this technique is described in section 4.1.2. On the basis of this measurement, wafers that fail the requirements of channel carrier concentration and thickness are rejected. By this means, unnecessary device processing can be avoided.

3.2.3 Gate formation

Material that is judged suitable for processing after layer assessment is cleaved so that a piece measuring approximately 12 by 15mm is removed for this purpose. This slice is then cleaned using organic solvents (for de-greasing), and hydrochloric acid to remove the oxide layer on the surface.

Patterning of the gate metal is achieved using a bi-level lift-off technique, to give good edge definition. A layer of PMMA resist is applied first, followed by a layer of AZ1450B resist. Windows corresponding to the gate metal pattern are opened in this top resist layer, using level 1 of the mask set.

The gates are aligned along the [110] direction, which can be found because it coincides with the direction of the major flat on the original wafer. Alignment in this direction gives a re-entrant rib profile, which is important because the effective gate length is then reduced.

The wafer is then flood exposed at a deep UV wavelength, to expose the PMMA resist layer which is now masked by the patterned AZ1450B layer. The PMMA is then slightly over-developed, so that the top resist layer is left overhanging, giving an ideal lift-off profile. The wafer is then ashed, and cleaned by dipping in 1:1:50, and is then suitable for metal deposition.

Gate metallisation is then thermally evaporated onto the wafer, which consists of titanium and gold in separate layers (Ti/Au). Titanium is widely used as a contact material, and has good adhesion to most semiconductors. Gold is also widely used due to its high conductivity and inert nature, which enables the pads to be contacted with low resistance. Gold also facilitates good thermo-compression bonding.
Excess metal is lifted off the wafer by soaking in acetone, which dissolves the resist from under this unwanted metal. At this stage, the wafer cross section is shown schematically in figure 3.7a.

This patterned gate metal is then used as a mask for the etching of the top p-type layer. The etchant used is 5:1:50, which has an etch rate of 0.45μm/min on InGaAs at 20°C. The
etching is done in stages, and after each etch step, the current flowing between two test pads is monitored. As the top p-type layer is gradually removed, the monitored current is reduced proportionally, until this layer is fully removed when the current falls to a level determined by the leakage current of the back to back pn junction so formed. In practice, the etching is continued until this current drops to less than 1μA, or until no further decrease is observed. As the channel layer is approached, the etch time is reduced, so that over-etching does not occur. Only three or four etch steps are usually required, to give a precision of better than 20nm.

Due to the undercutting of the mask and the crystallographic nature of the etching, a re-entrant rib profile is formed during this process. A schematic cross section of this structure is shown in figure 3.7b, and an SEM micrograph of this is shown in figure 3.8.

![SEM photograph of rib structure](image)

3.2.4 Source and drain formation

Patterning of the source/drain metal is accomplished by lift-off using Hunts W118 resist, and level 2 of the JFET mask set. This resist has been specially formulated to give good lift-off
profiles using a projection mask aligner. Although the metal edge definition using this resist is not as good as that achieved using the bi-level technique described above, the critical edges (along the gate edges), are formed by self alignment. The T-shaped gate acts as a shadow mask along these edges. Care must be taken not to deposit a layer of source/drain metal thicker than the rib height, otherwise the electrodes would short circuit one another.

This source/drain metallisation consists of 200nm of a eutectic alloy of germanium, and gold (Ge-Au), containing 12% germanium by weight. Lift-off of excess metal is again accomplished by soaking in acetone. The schematic cross section of the structure at this stage is shown in figure 3.7c.

To give a low contact resistance, these contacts are sintered at 300°C for 30 seconds, which has the effect of driving a small amount of germanium into the semiconductor to form a very thin highly doped n-type layer. A tunneling Schottky contact is then formed which has a lower resistance than the thermionic contact that existed prior to sintering.

3.2.5 Device isolation

Finally, the active area of the device is defined by mesa etching. Level 3 of the JFET mask set is used to pattern AZ1350 resist in a conventional photolithography procedure, so that device active areas are covered. The resist is hardened by baking, since it is to be used as an etch mask. The exposed InGaAs is then etched using 5:1:50 in steps to avoid overetching. If the resist has not been baked sufficiently, then severe undercut can occur, thereby ruining the wafer. When all of the exposed InGaAs has been removed, the mesa etching is complete since the InP substrate is not attacked by this etchant. However, etching is continued so that the gate bondpads can be isolated from the FET mesa, which occurs when the gap between the mesa and bondpad (which is bridged with gate metal) becomes fully undercut. A micrograph of the air bridge formed in this way is shown in section 3.1.3 (figure 3.4).

A photograph of the completed structure is shown in figure 3.9. The gate rib can be seen across the mesa (in the centre of the photograph), and separates the source and drain
electrodes. The gate bondpads are located at each side of this mesa, and the air bridges can be seen linking them to the gate rib.

Figure 3.9 Photograph of completed JFET
3.3 Optimisation

The fabrication procedures discussed in the previous section were optimised over a period of time. The first devices had very low yield, partly because the individual processes had not been sufficiently developed, and partly because they had not been combined in the best way to provide the optimum process schedule. Also, device performance was poor, because the design had not been optimised. The most important and relevant aspects of this optimisation are covered in more detail in this section.

3.3.1 Channel mobility

In the initial JFET design, the channel was grown directly onto the substrate. The devices made from these wafers showed low transconductance, and Hall effect measurements performed elsewhere at BTRL showed the mobility in the channel layers to be low. Typical mobility values of $3000 \text{cm}^2/\text{Vs}$ were recorded (good quality InGaAs layers should have electron mobilities of around $8000 \text{cm}^2/\text{Vs}$ at this doping level). Since thicker layers exhibited a higher average mobility, it was suspected that the mobility was being degraded near to the substrate. A mobility profile measurement through a channel layer showed this degradation with proximity to the substrate very clearly, and is shown in figure 3.10. Whitely and Ghandhi \cite{67}, observed a similar situation for MOVPE grown InGaAs, and attributed this degradation to space charge scattering caused by disorder at this interface. They observed some improvement by *in-situ* etching of the wafer prior to growth, but could not eliminate the problem.

In order to remove the channel from the deleterious effects of the substrate, a buffer layer was included in the device structure. A simple undoped buffer layer would have been unsuitable because the high residual doping level of greater than $10^{15} \text{cm}^{-3}$ would have given sufficient conduction in a buffer layer to adversely affect output conductance in the device characteristics.

InAlAs, was the initial choice of buffer because this material can be made semi-insulating when grown at low temperature. This is because the atomic mobility at the growing surface
is reduced at low temperature, leading to heavily defected growth. These defects act as electron traps, making the material semi-insulating. However, because an InAlAs buffer necessarily has poor crystallinity to be effective, the layers grown on top also have poor crystallinity. This results in severe degradation of device performance, characterised by very low transconductance and high gate leakage current.

The idea of using a p-type buffer layer was then attempted. The depletion region formed at the channel buffer interface effectively acts as an insulator, and the intention was to use this as a buffer. The simplest approach is a homojunction, and so structures incorporating a p-type InGaAs buffer layer were grown.

Initial results with this approach were disappointing. When buffers with a high doping level were grown (doped at $10^{19}\text{cm}^{-3}$ with beryllium), type conversion took place in the channel, and the whole structure became p-type. Buffer layers were also grown with lower doping levels ($10^{17}\text{cm}^{-3}$), and either type conversion or carrier compensation occurred in the
channel. Devices made from compensated material showed low transconductance, low threshold voltage, and non-ohmic source/drain contact behaviour.

The cause of this degradation was found to be beryllium diffusion during growth. Analysis by secondary ion mass spectroscopy (SIMS), clearly showed a significant diffusion profile extending from the buffer into the channel, and is shown in figure 3.11.

![SIMS profile of FET structure - beryllium concentration with distance into sample.](image)

The diffusion coefficient was assumed to be concentration dependent, according to the interstitial-substitutional model, similar to the case of zinc diffusion into GaAs [68] and InGaAs [69]. This hypothesis has been supported by some recent results of Enquist and coworkers [70] who observed anomalous redistribution of beryllium into GaAs during MBE growth. They explained some of their observations by assuming a concentration dependent diffusion coefficient.
Consequently, in order to alleviate this problem, the following steps were taken for material growth as part of this work,

i) the growth temperature and the beryllium concentration in the buffer were lowered, in order to reduce the diffusion coefficient. The growth temperature was reduced to a level where good quality epitaxy could still be maintained. More recent results showed no significant beryllium redistribution up to 530°C, even at a concentration of $10^{19}$ cm$^{-3}$. The revised growth temperature of 500°C was therefore well below this level.

The lowest practical beryllium concentration was approximately $10^{17}$ cm$^{-3}$ because the background donor level was variable, and could be greater than $10^{16}$ cm$^{-3}$ on occasions. Due to some anomalous doping profile measurements of layers grown at this level, there was doubt over the activation of beryllium at lower concentrations, which is thought to be due to oxygen incorporation.

ii) the doping level in the channel was slightly increased, to reduce any effects of carrier compensation. The silicon concentration in the channel was increased to $10^{17}$ cm$^{-3}$ from a previous design level of $5 \times 10^{16}$ cm$^{-3}$ in order to achieve this.

These adjustments to material growth and device design eventually lead to significant improvements to channel mobility and therefore device transconductance. Hall mobility measurements of this new structure showed more than a twofold increase, to values around 7000 cm$^2$/N s.

At a later stage in the course of this work, buffers were grown at high beryllium concentrations ($10^{19}$ cm$^{-3}$), in order to make the device easier to integrate with a PIN photodiode. Also, the channel doping level was reduced back to $5 \times 10^{16}$ cm$^{-3}$ because of the performance advantages to be gained. Neither of these changes degraded the channel, therefore only the reduction in growth temperature was important in this respect.
3.3.2 Contacts

3.3.2.1 Source and drain contact

The source and drain contact metallisation on the first sets of devices consisted of a layer of titanium followed by a layer of gold (Ti/Au). Because of the low Schottky barrier height formed by metals on n-InGaAs (0.2eV), it was considered that the ohmic contact produced by the thermionic emission of electrons over this barrier would be adequate for a low resistance contact. However, when the output characteristics of these devices were studied, it became apparent that current saturation occurred at a drain voltage that was higher than expected, due to the voltage dropped by the contact resistance.

Alloying is a common technique used to achieve low resistance contacts [71]. In this technique, a metal system such as gold, germanium, and nickel is patterned on the InGaAs surface. When this is heated to a temperature higher than the melting point of this combination (usually 450 - 500°C), alloying of the metal and semiconductor takes place. The nickel is added to provide 'wetting' for the metal pads. As the wafer is cooled, the alloy solidifies, and the germanium diffuses into the semiconductor creating a very high concentration of donors at the alloy-semiconductor interface. This produces a tunneling Schottky contact, since electrons can tunnel through the very thin depletion layer formed at this interface with ease.

However, this technique is unsuitable for the p-buffer JFET because the alloyed region extends up to 1μm into the semiconductor, and has a ragged 'front' with occasional spikes. Such a region would make contact with the buffer layer, and therefore lead to a parallel conduction path to the channel.

In an effort to achieve the same type of contact without alloying, contacts were formed using a eutectic mixture of germanium and gold (12% Ge) which were heated to temperatures below the alloy melting point for short periods. This process is termed 'sintering', and very low resistance contacts can be achieved in this manner. It is believed that some germanium diffuses into the semiconductor, forming a region of very high donor concentration, in a
similar fashion to alloyed contacts. The depletion region formed by such a highly doped semiconductor is extremely thin, leading to very large tunneling currents. Such a contact behaves in an ohmic manner, with very low resistance.

Figure 3.12 shows the variation in JFET 'end' resistance with sinter temperature for GeAu contacts, and a sinter time of 30 seconds. This resistance is the measured source-to-channel resistance, and is found by forward biasing the gate, and measuring the drain voltage required to give zero drain current. Since no voltage is dropped across the channel because no current flows, the drain voltage acts as a probe for the voltage at the source end of the channel. The end resistance is then given by this voltage divided by the gate current.

Each point in this figure is the average result taken from 29 devices, and the error bars show the extent of the deviation from the average value at each temperature.

Figure 3.12. Variation of JFET end resistance with sinter temperature.
The optimum sinter temperature was taken to be 300°C from this figure. The end resistance was found to be insensitive to sinter time in the range from 20 seconds to two or three minutes, therefore 30 seconds was adequate.

For the purpose of comparison, sets of Ti/Au and Ge-Au contacts were formed on the same wafer, and consisted of sets of two 100 by 100μm contact pads separated by 5μm. The resistance between sets of pads was measured for various times at a sinter temperature of 300°C. Table 3.1 shows the results of this work.

<table>
<thead>
<tr>
<th>Sinter time (s) at 300 °C</th>
<th>Contact resistance (Ω)</th>
<th>Ti/Au</th>
<th>Ge-Au</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>38.5</td>
<td>41.7</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>38.5</td>
<td>27.8</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>43.5</td>
<td>27.8</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>45.5</td>
<td>27.8</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1  Variation of contact resistance with sinter time at 300°C for Ti/Au and Ge-Au contact metallisations.

The resistance of the two contact metallisations are comparable prior to sintering. The Ge-Au contact resistance is substantially reduced after a 30 second sinter, whereas the Ti/Au contact resistance remains unchanged. This behaviour is expected from the model of the sinter process discussed above. However, upon longer periods of sintering, the Ti/Au contacts degrade, whilst the Ge-Au contact resistance remains at a low level.

As a result of this work, sintered Ge-Au contacts were used in the optimised JFET design.

3.3.2.2 Gate contact

The main requirements of a metal system when used for gate contact on the present JFET design are good ohmic contact to p-type InGaAs, resistance to the etchant used for rib
formation, and good mechanical strength for air bridge stability. Since the gate is voltage controlled, very low contact resistance is not essential, and therefore Ti/Au can be used if the acceptor level in the top layer exceeds $10^{19}$ cm$^{-3}$. This high concentration is necessary for a tunneling Schottky contact, otherwise the Schottky barrier height of 0.5 eV to p-InGaAs gives a rectifying characteristic. This metal system is resistant to hydrogen peroxide-sulphuric acid based etchants, and provides adequate mechanical properties, and has been used with success for this application.

Sintered zinc-gold eutectic (2% zinc) was tried as an alternative metal system, in order to dispense with the need for very high acceptor concentrations, thereby offering design flexibility (this type of contact is ohmic even at low semiconductor acceptor concentrations). Unfortunately, this metal alloy is attacked by hydrogen peroxide-sulphuric acid based etchants, and therefore proved unsuitable as gate contact metallisation.

Chromium and gold in separate layers, (Cr/Au), was also tried because it was expected to have better mechanical properties than Ti/Au, and therefore be more suitable for air bridge formation (the shear modulus of chromium is nearly three times that of titanium, and Young's modulus is more than double). The contact and etch resistance properties of this metal system are similar to those of Ti/Au. However, the air bridges formed were found to be too brittle in use, which lead to a high level of breakage.

Consequently, Ti/Au continued to be the preferred gate metal system, which meant that the use of highly doped top layers were necessary in the JFET design.

3.3.3 Rib profile
The optimum rib profile is trapezoidal, with the base narrower than the top, which produces a shorter gate length for a given gate metal dimension. This 're-entrant' profile also gives little chance of short circuiting during the self alignment process, partly because the gate metal is well supported and therefore less likely to sag, and partly because the base of the rib is well hidden from the source and drain metal.
Due to the crystallographic nature of the etching of III-V semiconductors, this type of profile can be obtained if the rib is aligned along the [110] direction, (a trapezoid with the base broader than the top is formed in the direction perpendicular to this). However, the actual profile formed is determined by the relative etch rates of the crystal planes. The trapezoids discussed above are formed because the (111) planes have a relatively slow etch rate, and are therefore preferentially exposed during etching. Therefore, the detailed profile is determined by the directional etch rate variation of the etchant used, which is a function of its composition and temperature [72].

A common perturbation to the perfect re-entrant trapezoid is a 'tail', as shown in figure 3.13. This tail is deleterious because it serves to increase the effective gate length of the device, and in severe cases can cause short circuiting between the rib and the source or drain contact. A measure of the degree of 'tailing' was characterised by the ratio of the depth of rib with tail to the total rib height, or the ratio $a/b$ in figure 3.13.

![Figure 3.13 Typical rib profile along [110] direction](image)

The rib etchant composition and temperature were varied in order to find an etchant that would give minimum tailing with a moderate etch rate, and leave a smooth surface. The results of this work are presented in table 3.2 below.
<table>
<thead>
<tr>
<th>Etchant</th>
<th>Temperature (°C)</th>
<th>Etch rate (µm/min)</th>
<th>Tail ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:1:50</td>
<td>20</td>
<td>0.11</td>
<td>0.53</td>
</tr>
<tr>
<td>5:1:50</td>
<td>20</td>
<td>0.45</td>
<td>0.36</td>
</tr>
<tr>
<td>10:1:50</td>
<td>20</td>
<td>0.92</td>
<td>0.24</td>
</tr>
<tr>
<td>10:1:50</td>
<td>10</td>
<td>0.45</td>
<td>0.25</td>
</tr>
<tr>
<td>20:1:50</td>
<td>10</td>
<td>0.68</td>
<td>0.17</td>
</tr>
<tr>
<td>20:1:50</td>
<td>5</td>
<td>0.60</td>
<td>0.16</td>
</tr>
<tr>
<td>10:2:50</td>
<td>10</td>
<td>0.51</td>
<td>0.19</td>
</tr>
<tr>
<td>10:2:50</td>
<td>5</td>
<td>0.40</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Table 3.2 Variation of etch rate and tail ratio with etchant composition and temperature

Unfortunately, the etchants that gave the lowest tail ratios gave a rough surface finish. Therefore the preferred rib etchant was 5:1:50, used at 20°C, which was chosen because it gave a smooth surface, and an acceptable rib profile.
Chapter 4. Discrete InGaAs FET performance - basic device

Results and analysis given in this chapter relate to the optimised device described in the previous section. The techniques, equipment, and software described below for current-voltage and capacitance-voltage measurements were devised, assembled and written as part of this work.

4.1 Measurement techniques

4.1.1 dc current-voltage (IV)

The basis of the dc current-voltage measurement system was a Hewlett-Packard HP4145 semiconductor parameter analyser. This instrument consists of four stimulus and measurement units (SMUs), each of which can be configured as a voltage source/current monitor or current source/voltage monitor. In addition to this, there are two voltage sources, and two voltage monitors.

The HP4145 can be operated as a stand-alone unit, using the front panel keyboard to control the menu driven custom software that forms the operating system of this machine. However, it is much more convenient to use a desktop computer to control the 4145 via the IEEE-488 bus. A HP9836 computer was used for this purpose. This computer was ideally suited for this application because of its high speed and large clear display for graphics presentation. In addition to this, the powerful BASIC language supplied gave simplified programming and editing.

All devices were measured on-wafer. To accomplish this, a Wentworth AWP-1000 automatic probe system was used, controlled by the computer via the IEEE-488 bus. A probe card was incorporated in this system to facilitate rapid measurement. This consists of a card with a fixed probe pattern, which is positioned to coincide with the contact pads of the devices. The card can then be stepped across the wafer, from device to device, if the pad layout is identical for each device.

The computer was programmed to instruct the prober to step from device to device, each time setting up the HP4145 to make the required measurements. To illustrate the speed of
this system, a typical device characterisation consisted of up to 15 spot measurements. Parameters extracted from these measurements were gate leakage current (at -1V and -2V), threshold voltage, saturated drain current at zero gate bias, transconductance (at zero gate bias and at half threshold voltage), and output conductance. A square centimetre slice would have a 24 x 15 array of devices, which could be fully characterised in less than an hour. Also, because it was fully automated, once the initial setting up was performed, the system could be left unattended.

Once the raw measurement data had been stored on disc, additional software was used to manipulate it into a form suitable for presentation and analysis.

4.1.2 Capacitance-voltage (CV)

The basis of the capacitance-voltage measurement system was a Hewlett-Packard HP4275 LCR meter, which is capable of operating at a choice of 10 spot frequencies between 10kHz and 10MHz. For all measurements, a frequency of 1MHz at a level of 10mV was chosen. For semiconductor device CV measurements in general, a low oscillator level is desirable to avoid problems due to non-linearity. However, a compromise must be found because accuracy is reduced if this level is too low. The measurement frequency of 1MHz has become a standard for this type of application, and was chosen for this reason.

The HP4275 uses a vector voltmeter to measure impedance, and a microprocessor converts this into the desired parameter, compatible with either a series or parallel resistance in the equivalent circuit. The most appropriate measurement parameters for this application were capacitance and conductance. If the conductance is too high then accuracy is reduced, and so it was essential to monitor this parameter during measurement.

To ensure accurate, stable readings, the device was connected to the LCR meter using high frequency co-axial probes, which were shielded to within a few millimetres of the probe tip. The probe system was positioned very close to the capacitance meter to minimise excess capacitance.

Setting up and operation of the HP4275 was accomplished using the HP9836 desktop computer, via the IEEE-488 bus. To improve accuracy, each measurement was formed from the average of ten, or sometimes fifty, successive readings. To perform CV measurements,
a Hewlett-Packard HP59501 programmable power supply was used, which was connected to the HP4275. To eliminate errors due to offset or non-linearity of this power supply, a Keithley 177 digital voltmeter (DVM) was used to measure the voltage from the power supply, and relay the measurement to the controlling computer via the IEEE-488 bus. A block diagram of this equipment is shown in figure 4.1.

![Block diagram of capacitance-voltage measurement system](image)

**Figure 4.1** Block diagram of capacitance-voltage measurement system

A computer program was written to perform these measurements automatically. Once the voltage limits had been set, the power supply was instructed to increment the voltage applied to the device. For each increment, the capacitance, conductance, and voltage measurements were passed to the computer, for display and storage.

The CV data obtained in this way can be manipulated to provide a carrier concentration profile of the channel layer\(^{[73]}\). The capacitance due to the depletion layer of a p\(^+\)n junction is inversely proportional to the depletion depth, d, and is given by the equation;
\[ C = \varepsilon_s A / d, \] ...... 4.1

where \( \varepsilon_s \) is the semiconductor permittivity, and \( A \) is the device area.

This depletion depth is related to the carrier concentration in the n-type layer, \( N_d \), (i.e. the FET channel layer), and the applied voltage, \( V \), by the following expression, which is derived from a one dimensional solution of Poisson's equation,

\[ d = \left( \frac{2 \varepsilon_s (V_{bi} - V) / q N_d}{1/2} \right) \] ...... 4.2

where \( V_{bi} \) is the junction built-in voltage. Substituting this relation into the previous equation, gives an expression for the depletion layer capacitance as follows;

\[ \frac{1}{C^2} = 2 (V_{bi} - V) / (q \varepsilon_s N_d A^2) \] ...... 4.3

This expression is differentiated with respect to \( V \), and then re-arranged to give

\[ N_d = \frac{2}{(q \varepsilon_s s A^2)}, \] ...... 4.4

where \( s \) is the differential \( d \left( 1/C^2 \right) / d V \).

This differential is simply the slope of a plot of inverse square capacitance with voltage, obtained from CV measurements. This expression also holds for a more general dopant distribution, and can be used to obtain a doping profile of the FET channel.

Software was written as part of this work to provide a doping profile from this CV data. The raw measurement data was taken, and a rolling 5-point least squares fit was performed on each data point to give a smoothed slope at each point (including the two points either side of the main point). This was used to compute the doping level at each point, and by calculating the incremental change in depletion depth, a plot of doping level with depth into the channel was produced.
To increase the accuracy of this technique, this sort of profile was usually obtained on large area test patterns. In this way, errors in the measurement of area were minimised, and also, effects due to fringing capacitance were made negligible.

4.1.3 Microwave measurements

In order to make microwave measurements, the wafers were diced, and the individual device chips were mounted onto alumina carrier substrates. These substrates used coplanar waveguide input/output transmission lines, and the JFET electrodes were connected to these transmission lines using short lengths of gold wire. This assembly was then inserted into a microwave transistor test fixture, made by Design Techniques, which had been characterised, (with the substrate) up to 26GHz. De-embedding software provided with the test fixture enabled the device characteristics to be evaluated up to this frequency.

S-parameters provide convenient characterisation for two-port devices at microwave frequencies. There are four S-parameters, which characterise reflection and transmission from a 50Ω system at the input and output ports. Using a software package known as Fetfitter, supplied by Cascade Microtech, JFET equivalent circuit values were extracted from S-parameter measurements. The S-parameter measurements were made using a Hewlett-Packard HP8510 network analyser, up to a frequency of 5GHz. A block diagram of this measurement system is shown in figure 4.2.

Noise figure measurements were made on similarly packaged devices using a Hewlett-Packard HP8970A noise figure meter with a HP346C calibrated noise source. The microwave transistor test fixture and bias tees were the same as used for S-parameter measurements described above. Noise figure (in a 50Ω system) was measured as a function of JFET bias conditions at a frequency of 1.5GHz. This measurement frequency was chosen because it was shown to be considerably higher than the break frequency where 1/f noise dominates (500MHz).
Figure 4.2 Block diagram of microwave measurement system
4.2 Standard device characteristics

The JFET produced from the standard mask set has nominal gate dimensions of 2.0μm x 275μm, and is fabricated by the self-alignment technique described in chapter 3. In the following, this device is characterised.

Figure 4.3 shows the output characteristics, and figure 4.4 shows the transfer characteristics of a typical device, which has gate metal dimensions of 2.6μm x 275μm. Maximum transconductance of this device (at zero gate bias) is 37mS - shown by the dashed line in figure 4.4 - which is equal to 135mS/mm when scaled for unit gate width. The highest figure achieved with this type of design was 215mS/mm, for a device with a shorter gate, and fabricated from material with superior electron transport properties. However, the characteristics shown in these figures represent a more typical situation.

Figure 4.3 Output characteristic of typical device.
The threshold voltage, $V_t$, of a JFET is determined by the channel thickness and doping by the following relation, obtained from a one-dimensional solution of Poisson's equation, (assuming an abrupt one-sided junction with uniform dopant density),

$$V_t = \left[ \frac{q}{(2\varepsilon_s) \cdot a^2 \cdot N_d} \right] - V_{bi} \quad \ldots \quad 4.5$$

For this work $V_t$ is defined as the voltage at which the saturated drain current is reduced to 5% of its value at zero gate bias. For this device, $V_t = 2.0\text{V}$, which gives $a^2 \cdot N_d = 3.7 \times 10^7 \text{cm}^{-1}$. The channel doping can be deduced independently of the transport measurements from a capacitance-voltage analysis. Figure 4.5 shows a plot of $1/C^2$ against voltage for a 100μm square test structure fabricated from the same wafer, and figure 4.6 shows the resultant doping profile from the analysis.

A threshold voltage can be found from this profile in a similar way to the transfer characteristic. The area enclosed by this doping profile is equal to the sheet charge in the channel. The threshold voltage is now the voltage at which the charge drops to 5% of the total charge from zero gate bias. Consistency in these measurements is preserved since the same voltage is observed in both techniques.
Figure 4.5 Inverse square capacitance variation with bias for test structure

Figure 4.6 Profile of carrier concentration with depth into channel layer derived from data in figure 4.5
A channel carrier concentration of $9.10^{16}$ cm$^{-3}$ is noted from figure 4.6, from which a channel thickness of 0.2μm is deduced. Having established the channel carrier concentration from CV measurements on the test structure, similar measurements on the actual device allow the calculation of the effective gate length. Re-arranging equation 4.4 and noting that $A = L \cdot Z$ gives the relation:

\[
L = \left\{ \frac{2}{(q \cdot \epsilon_s \cdot \varepsilon_s \cdot N_d \cdot Z^2)} \right\}^{1/2} \quad \ldots \quad 4.6
\]

Figure 4.7 shows the variation of gate-source capacitance and inverse square gate-source capacitance of this device with voltage. The slope, $s$, of the latter is, $s = 3.8$ pF$^2$/V, which gives $L = 2.0$ μm when inserted into the above equation.

![Figure 4.7 Variation of gate-source capacitance and inverse square gate-source capacitance with gate bias](image)
Since the gate metal length (and hence source drain spacing) is 2.6μm, and the device is symmetrical, a source-gate and gate-drain spacing of 0.3μm is deduced. This figure agrees with the undercut produced as a consequence of gate formation, since the undercut is roughly equivalent to the rib height.

The close source-gate spacing referred to above is one of the most important advantages in this JFET design. Without self-alignment, a spacing of 0.3μm would be very difficult to achieve with any consistency. This close spacing minimises the resistance between the source and the channel, which gives an increased extrinsic transconductance. To assess this series resistance, $R_s$, the end resistance, $R_e$, of this device was measured. The principle of this type of measurement is described in section 3.3.2, and figure 4.8 shows the results obtained for this device. The gate was forward biased to give $I_g = 10mA$, and the drain voltage was swept between 0 and 100mV. The intercept of this plot (i.e. at $I_d = 0$), gives the potential under the gate, since no voltage is dropped between this point and the drain.

![Figure 4.8 Variation of drain current with drain bias for forward biased gate (Igs = 10mA)](74)
The 'end' resistance, $R_e$, is then this intercept (35mV), divided by $I_g$ (10mA), and is therefore 3.5Ω. The total channel resistance, $R_t$, is equal to the inverse of the slope of this plot, and is equal to 7.0Ω for this device. These values include the excess resistance due to the probes and leads to the measurement system, which can be measured by probing to a highly conducting surface. This excess resistance was 1.5Ω in total, which gives $R_t = 5.5Ω$, and $R_e = 2.8Ω$, when subtracted. Due to the distributed nature of the current flow through the gate, the 'end' resistance is given by

$$R_e = R_s + R_{ch}/2,$$  \hspace{1cm} (4.7)

where $R_{ch}$ is the channel resistance\(^{[74]}\). Since the total resistance is,

$$R_t = R_s + R_{ch} + R_d,$$  \hspace{1cm} (4.8)

where $R_d$ is the resistance between the channel and the drain, and since the device is symmetrical (i.e. $R_s = R_d$), then it follows that,

$$R_t = 2 \cdot R_e,$$  \hspace{1cm} (4.9)

which is valid for this device. However, it is impossible to isolate $R_s$ directly from these measurements without making some assumptions regarding the channel.

An approximate calculation of the channel resistance can be made from the relation

$$R_{ch} = L / (N_d \cdot q \cdot \mu \cdot a \cdot Z)$$  \hspace{1cm} (4.10)

by assuming that negligible depletion exists due to the heavy forward bias of the measurement. A Hall mobility measurement on this layer gave a value of 8000cm\(^2/Vs\), which gives $R_{ch} = 3.2Ω$ when inserted into equation 4.10 with the other parameters. This gives $R_s = 1.2Ω$ when substituted into equation 4.7. The resistance of the semiconductor between the source and the channel can be estimated using an equation similar to 4.10, and
gives a value of 0.5Ω, which leaves a contribution due to the contact of 0.7Ω. The intrinsic transconductance of a FET, \( g_m^* \), is given by

\[
g_m^* = \frac{g_m}{(1 - g_m \cdot R_s)}
\]

which is equal to 40mS for this device, or 145mS/mm when scaled for gate width. The resistance added by the probes and leads has been included in this calculation.

An estimate of the average electron velocity in the channel can be obtained from the transfer characteristic if it is assumed that the channel is fully velocity saturated. Although this may not be strictly valid, the results obtained should give a qualitative assessment for devices with similar gate length. The drain current, \( I_d \), flowing in a velocity saturated channel is given by

\[
I_d = q \cdot N_d \cdot Z \cdot (a - h) \cdot u_s
\]

where \( h \) is the depletion depth, and \( u_s \) is the saturated electron velocity. If the drain current is differentiated with respect to the depletion depth, and then re-arranged, the following expression is found for \( u_s \)

\[
u_s = \frac{(d I_d/d h)}{(q \cdot N_d \cdot Z)}
\]

The depletion depth can be found as a function of gate bias from a one dimensional solution of Poisson’s equation:

\[
h = \{ \frac{2 \cdot \varepsilon_s \cdot V}{q \cdot N_d} \}^{1/2}
\]

where \( V \) is the potential across the depletion layer at the drain end of the channel at the onset of saturation. This is given by

\[
V = V_{dsat} + V_{bi} - V_{gs}
\]
if the voltage dropped between the channel and the drain is neglected. This is valid since $R_d$ has been measured to be very low resistance. Using $V_{ds}$ obtained from figure 4.3, and $V_{bi} = 0.6V$, a plot of drain current against depletion depth can be obtained from the transfer characteristic, and this is shown in figure 4.9.

The saturated velocity can be obtained from the slope of this plot, and a value of $v_s = 1.6 \times 10^7 \text{ cm/s}$ is found for this device.

![Image](image.png)

Figure 4.9 Variation of saturated drain current with depletion depth

Another important parameter to be considered is the output conductance of the device, which is the slope of the output characteristic beyond current saturation. This conductance must be low, or conversely, the output resistance, $R_{ds}$, must be high in order to have little effect on the amplifier gain. As a rough guide, a figure of $400 \Omega$, is considered adequate for this application. Figure 4.10 shows the variation of output resistance with gate bias at a drain bias of 1.2V.
It is apparent that the output resistance of this device is adequate over most of the range of gate bias, and is particularly good at operating bias (close to threshold). For instance, at $V_{gs} = -1.5\,\text{V}$, $R_{ds} = 1400\,\Omega$, (or $g_{ds} = 0.7\,\text{mS}$).

![Graph of output resistance vs gate bias](image)

Figure 4.10 Variation of output resistance with gate bias at a drain bias of 1.2V.

For low-noise operation, the gate leakage current of the FET must be low, since this adds shot noise to the receiver. Figure 4.11 shows the gate leakage current of this device, which indicates a leakage current of over 800nA at a typical operating bias of $V_{gs} = -1.5\,\text{V}$. This is considered to be a high value, and limits the usefulness of this transistor to high transmission rate systems, since the relative noise contribution decreases at high bit rates. However, lower leakage current has been achieved in other devices, and is known to increase during the contact sintering operation. Additional optimisation is necessary to reduce this leakage current further.
The dimensions and measured material properties of the channel layer of this device are listed in table 4.1, and the device parameters are summarised in table 4.2. As a demonstration of the applicability of the model discussed in appendix 1, the channel properties of table 4.1 were inserted into this model and compared to the measured results. Figure 4.12 shows the transfer characteristic from this model (lines), with the measurement points superimposed. Clearly, there is significant deviation between this model and experimental results which shows that caution must be shown when applying this model, even for qualitative use.
### Table 4.1 Channel properties of the standard device

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gate length, L</td>
<td>2.0μm</td>
</tr>
<tr>
<td>gate width, Z</td>
<td>275μm</td>
</tr>
<tr>
<td>thickness, a</td>
<td>0.20μm</td>
</tr>
<tr>
<td>carrier concentration, N\text{d}</td>
<td>(9 \times 10^{16}) cm(^{-3})</td>
</tr>
<tr>
<td>mobility, (\mu)</td>
<td>8000 cm(^2)/Vs</td>
</tr>
<tr>
<td>saturated velocity, (v_s)</td>
<td>1.6 \times 10^7 \text{cm/s}</td>
</tr>
</tbody>
</table>

### Table 4.2 Summary of measured parameters of the standard device

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>saturated drain current, (I_d)</td>
<td>46mA</td>
</tr>
<tr>
<td>transconductance, (g_m)</td>
<td>37mS</td>
</tr>
<tr>
<td>gate-source capacitance, (C_{gs})</td>
<td>0.60pF</td>
</tr>
<tr>
<td>output resistance, (R_{ds})</td>
<td>340Ω</td>
</tr>
<tr>
<td>gate-source leakage current, (I_{gs})</td>
<td>-</td>
</tr>
<tr>
<td>threshold voltage, (V_t)</td>
<td>2.1V</td>
</tr>
<tr>
<td>series resistance, (R_s)</td>
<td>1.2Ω</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(V_{gs} = 0) V</td>
<td>46mA</td>
</tr>
<tr>
<td>(V_{gs} = -1.5) V</td>
<td>5mA</td>
</tr>
<tr>
<td>(V_{gs} = 0) V</td>
<td>13mS</td>
</tr>
<tr>
<td>(V_{gs} = -1.5) V</td>
<td>0.33pF</td>
</tr>
<tr>
<td>(V_{gs} = 0) V</td>
<td>1400Ω</td>
</tr>
<tr>
<td>(V_{gs} = -1.5) V</td>
<td>800nA</td>
</tr>
</tbody>
</table>
The data obtained from this one device design is obviously limited, and it is difficult, therefore, to gain much insight into the device operation. A mask set was designed to overcome this limitation, which allowed devices to be produced with different gate lengths, and also different electrode spacings. Results from these devices are presented in the next chapter.

With the exception of the capacitance data, all the data presented in this chapter has been obtained under dc conditions. This is obviously easy to accomplish, because the data can be collected from each device whilst on the wafer. Although this data has been invaluable in the development of this JFET, a proper characterisation must include microwave measurements. These measurements serve to show the validity of the dc data, and also to evaluate some parameters that are difficult to measure under dc conditions.

Because this microwave characterisation was performed on devices with varying geometry, this data will be presented in the next chapter.
Chapter 5 Variation of JFET characteristics with channel properties

In order to gain a clearer insight into the behaviour of the InGaAs rib-JFET, it was necessary to analyse the variation of device characteristics with channel properties. However, due to inconsistency in the properties of the epitaxial material, this was difficult to achieve. Attempts to analyse device results from wafers grown consecutively with a range of channel properties were hindered by this inconsistency. It was necessary, therefore, to vary channel properties on the same wafer. A mask set was designed which enabled devices with different channel and gate dimensions to be fabricated. This mask set enabled devices to be produced with four different gate lengths, and three variations of electrode spacing, giving a cell of twelve different device types. On a standard wafer, which had a square centimetre of processed area, there were 30 examples of each device type per wafer. Table 5.1 below gives the design dimensions of each device type.

<table>
<thead>
<tr>
<th>gate type</th>
<th>nominal gate length</th>
</tr>
</thead>
<tbody>
<tr>
<td>very short gate (vsg)</td>
<td>1μm</td>
</tr>
<tr>
<td>short gate (sg)</td>
<td>2μm</td>
</tr>
<tr>
<td>medium gate (mg)</td>
<td>5μm</td>
</tr>
<tr>
<td>long gate (lg)</td>
<td>10μm</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>spacing type</th>
<th>source-gate spacing</th>
<th>gate-drain spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>self aligned (sa)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>narrow spaced (ns)</td>
<td>2μm</td>
<td>2μm</td>
</tr>
<tr>
<td>wide spaced (ws)</td>
<td>2μm</td>
<td>5μm</td>
</tr>
</tbody>
</table>

Table 5.1 Gate length and electrode spacing variation on revised mask set.
This mask set was used on several wafers, and the variation of several device parameters were analysed as a function of device type. A list of the parameters that were measured routinely for every device is given below:

- saturated drain current at zero gate bias \( I_{dss} \)
- threshold voltage \( V_t \)
- transconductance at zero gate bias \( g_m0 \)
- transconductance at half threshold \( g_{m1} \)
- output conductance at half threshold \( g_d \)
- gate leakage current at -1 V \( I_{gs} \)

Table 5.2 below gives the results from a typical wafer as a function of device type. The values given in this table have been averaged over the total number of working devices for each type. The variation of these parameters is discussed in the following sections.

<table>
<thead>
<tr>
<th>Device type</th>
<th>( I_{dss} ) (mA)</th>
<th>( V_t ) (V)</th>
<th>( g_m0 ) (mS)</th>
<th>( g_{m1} ) (mS)</th>
<th>( g_d ) (mS)</th>
<th>( I_{gs} ) (μA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>vsg sa</td>
<td>80.4</td>
<td>-2.97</td>
<td>28.1</td>
<td>24.4</td>
<td>15.4</td>
<td>4.15</td>
</tr>
<tr>
<td>sg sa</td>
<td>57.4</td>
<td>-2.62</td>
<td>28.0</td>
<td>20.5</td>
<td>7.9</td>
<td>4.56</td>
</tr>
<tr>
<td>mg sa</td>
<td>31.4</td>
<td>-2.53</td>
<td>20.9</td>
<td>10.4</td>
<td>1.5</td>
<td>5.53</td>
</tr>
<tr>
<td>lg sa</td>
<td>16.7</td>
<td>-2.43</td>
<td>12.5</td>
<td>5.4</td>
<td>0.4</td>
<td>6.28</td>
</tr>
<tr>
<td>vsg ns</td>
<td>48.4</td>
<td>-3.36</td>
<td>8.7</td>
<td>16.5</td>
<td>4.7</td>
<td>0.12</td>
</tr>
<tr>
<td>sg ns</td>
<td>41.7</td>
<td>-2.77</td>
<td>13.3</td>
<td>16.0</td>
<td>3.2</td>
<td>0.29</td>
</tr>
<tr>
<td>mg ns</td>
<td>26.0</td>
<td>-2.65</td>
<td>14.0</td>
<td>8.9</td>
<td>0.8</td>
<td>0.42</td>
</tr>
<tr>
<td>lg ns</td>
<td>15.2</td>
<td>-2.58</td>
<td>10.2</td>
<td>4.9</td>
<td>0.3</td>
<td>1.16</td>
</tr>
<tr>
<td>vsg ws</td>
<td>37.6</td>
<td>-3.41</td>
<td>3.8</td>
<td>14.9</td>
<td>3.4</td>
<td>0.05</td>
</tr>
<tr>
<td>sg ws</td>
<td>32.7</td>
<td>-2.81</td>
<td>6.2</td>
<td>15.1</td>
<td>3.0</td>
<td>0.24</td>
</tr>
<tr>
<td>mg ws</td>
<td>22.9</td>
<td>-2.74</td>
<td>9.0</td>
<td>8.6</td>
<td>0.7</td>
<td>0.53</td>
</tr>
<tr>
<td>lg ws</td>
<td>14.6</td>
<td>-2.64</td>
<td>8.5</td>
<td>4.8</td>
<td>0.3</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 5.2 Variation of parameters with device type on a typical wafer. Refer to table 5.1 for device type meaning.
5.1 Transconductance

5.1.1 Variation with gate bias

All simple JFET/MESFET models, including the model used for this work, predict that maximum transconductance occurs at zero gate bias in the saturation regime, and indeed, this situation is observed for the standard 2μm gate length, self-aligned device. From table 5.2 however, it is evident that this is not true for certain device types. For example, the very short gate length, wide spaced device type shows a very low transconductance at zero bias and a much higher transconductance at half threshold voltage. This transconductance compression at zero gate bias is very evident in some device types, while less so in others. In order to quantify this compression, a parameter $g_c$ was defined as

$$g_c = \frac{g_{m0} - g_{m1}}{g_{m1}}$$

Figure 5.1 shows the variation of this compression with gate length for different electrode spacings.

![Figure 5.1 Variation of transconductance compression with gate length for different electrode spacings](image-url)
Clearly, the greatest compression is observed in the very short gate, wide spaced device, and the least compression is observed in the long gate, self-aligned device. This effect is therefore more pronounced when the channel length is long compared to the gate length of the device.

This compression can be seen in more detail by comparing transfer characteristics of typical devices in each type group. Figure 5.2 shows the transfer characteristics of short gate devices at different electrode spacings. The transconductance is compressed above a bias of -0.6V in each case, and the compression is greatest in the wide spaced device.

![Figure 5.2](image)

**Figure 5.2** Variation of transconductance with gate bias for short gate length devices

Although the transconductance in the self-aligned device is at a maximum at zero bias, it is probably less than it would have been if this effect was absent. The convex curvature of the characteristic is evidence for this assumption.

Figure 5.3 shows the transfer characteristics of self-aligned devices at different gate lengths. Again, there are striking differences between the form of these plots. For the long
and medium gate length devices, the rate of increase of transconductance as zero gate bias is approached is positive, whereas for the two shorter gate length devices, this rate is negative. Transconductance compression is observed for the very short gate device above -0.6V.

![Graph showing variation of transconductance with gate bias for self-aligned devices](image)

**Figure 5.3** Variation of transconductance with gate bias for self-aligned devices (gate length given in μm)

To further emphasise this point, figure 5.4 shows the transfer characteristics of devices at the extremes of this effect. Transconductance compression is most evident in the very short gate, wide spaced device, whereas this effect probably has negligible influence on the long gate, self-aligned device. Although the transconductance of the very short gate device is expected to be far greater than that of the long gate device, the consequence of this transconductance compression is that the transconductance at zero gate bias of the very short gate device is approximately half that of the long gate device. However, at a gate bias of -2V, the transconductance of the very short gate device is approximately four or five times that of the long gate device, which is more in line with expectations.

As has already been mentioned, maximum transconductance compression occurs when the gate length is a small part of the total channel length. The inter-electrode spaces, therefore,
must dominate this effect. To discover which space was most influential (i.e. the source-gate spacing or the gate-drain spacing), a wide spaced device was measured normally (source nearest gate), and then reconfigured (drain nearest gate), and remeasured. Figure 5.5 shows that the transconductance compression is most evident for the configuration where the gate-drain spacing was greatest. This effect, therefore, is dominated by the region between the gate and drain electrodes.

Transconductance compression near zero gate bias is known to occur in the characteristics of dual gate FETs\textsuperscript{[75,76]}. This type of device is normally operated with the gate nearest the drain (gate 2) held at a fixed potential ($V_{g2s}$), while the bias on the gate nearest the source (gate 1) is varied ($V_{g1s}$). In the operating regime of interest, i.e., beyond current saturation and near zero gate 1 bias, transconductance compression occurs over a wide range of gate 2 bias. For example, according to the work referenced above, transconductance compression even occurs when $V_{g2s} = 0V$.

The operation of the dual gate FET can be thought of as two separate FETs in series, FET1 and FET2. In the operating regime discussed above, current flow in FET2 is saturated,
Figure 5.5  Variation of transconductance with gate bias for wide spaced device configured for different gate-drain separation (given in μm).

while in FET1 it remains unsaturated. The drain current is therefore determined by FET2, and as a result, $V_{g1s}$ has little influence on drain current, i.e. transconductance is low. However, as $V_{g1s}$ becomes more negative, the current in FET1 saturates, and then gate 1 regains control of the characteristic, i.e. transconductance becomes higher. If $V_{g2s}$ is made more negative, then $V_{g1s}$ must become more negative still to regain current control, i.e., transconductance compression is more pronounced.

According to Hariu, this effect has also been observed in enhancement-mode GaAs MESFETs\(^{[77,78]}\). These devices are operated with positive gate bias, and are therefore potentially sensitive to this effect. Hariu argued that the region between the channel and drain acts as a parasitic MESFET due to the surface depletion in this region, which is caused by the pinning of the Fermi level at the GaAs surface. Fermi level pinning at the GaAs surface is a well known phenomenon, and is caused by the high density of surface states in this material. The depletion region at the surface of GaAs caused by this pinning is equivalent to the depletion caused by the Schottky barrier in a MESFET. In fact, Fermi level
pinning is the reason why Schottky barriers formed using different metals have roughly the same barrier height on this material.

The depletion region of a GaAs MESFET covers the whole of the channel, therefore, and not just the region under the gate. The depletion region that exists between the gate and the drain will be modulated by variation of gate bias because the surface potential in this region will change with gate bias. This region under the free surface therefore acts in a similar manner to the second gate in the dual gate FET discussed earlier, and gives rise to transconductance compression. Using a simple model, where the potential of the free surface was assumed to vary linearly between the gate and drain, coupled with the analytical model proposed by Lehovec and Zuleeg for the intrinsic FET\textsuperscript{59}, transconductance compression near zero gate bias was shown to occur.

To investigate whether free surface depletion occurred in the InGaAs JFETs produced in this work, a simple low-field device model was devised, and then compared to experimental measurements. Low field in this sense refers to the longitudinal field. A low-field approach was chosen because of its simplicity and accuracy, and because it was considered sufficient to show whether or not free surface depletion was present in these devices.

In this model, the device is divided into three sections. Section 1 is the region between the source and gate, section 2 is the region under the gate, and section 3 is the region between the gate and drain. In all three sections, the longitudinal electric field is so low that the depletion region is assumed to be flat, and that Ohm's law is obeyed. The extent of the depletion layer in each section is simply determined from equation 4.14, by substituting the appropriate potentials across the depletion regions. In section 2, this is simply obtained from the gate bias. In the other sections it is assumed that the free surface potential is constant over the section, and is obtained by taking the average potential between the electrodes. In this way, the thickness of remaining channel in each section can be calculated as a function of gate bias. The resistance of each section can then be determined from its dimensions, if the carrier concentration and mobility of the channel are known. Using Ohm's law, the total drain current can then be calculated, and the transconductance can be found from the variation of this current with gate bias. This was performed at a drain bias of 0.1V, for
different values of built-in potential at the free surface, and compared to characteristics obtained from actual devices. Three narrow spaced devices, with different gate lengths were chosen for this purpose, and transfer characteristics were then measured for each device. Details of the geometry of each device were carefully measured, and channel properties were analysed from CV measurements as described previously. These details were used in this model, and then the value of the free surface built-in potential was varied until the best fit with experiment was obtained. Figures 5.6 and 5.7 show the results of this work, where the experimental data has been included as points, with the predictions from the model included as full curves.

Figure 5.6 Transfer characteristic at low field (drain bias = 0.1V). Theory (lines) and experiment (points) for the case where $\psi_s = 0.2$V. Gate length given in $\mu$m.
Figure 5.6, shows the fit for a free surface built-in potential ($\psi_b$) of 0.2V, while figure 5.7 shows the situation for no free surface depletion. Clearly, the best fit occurs for the situation where there is no free surface depletion.

\[ \text{drain current, mA} \]

\[ \begin{array}{c}
0.5 & 1 & 2 & 3 & 4 & 5 \\
0 & 1 & 2 & 3 & 4 & 5
\end{array} \]

\[ \text{transconductance, mS} \]

\[ \begin{array}{c}
0 & 1 & 2 & 3 & 4 & 5 \\
0 & 1 & 2 & 3 & 4 & 5
\end{array} \]

\[ \text{gate bias, V} \]

Figure 5.7 Transfer characteristic at low field (drain bias = 0.1V). Theory (lines) and experiment (points) for the case where $\psi_b = 0\text{V}$. Gate length given in $\mu\text{m}$.

From this evidence, it seems unlikely that the transconductance compression observed in the InGaAs JFETs produced in this work can be explained in the same way as for the GaAs MESFETs discussed earlier. Some insight into the surface effect can be gained from comparison of device performance before and after source/drain contact sintering.
Figure 5.8 shows the transfer characteristic of a medium gate length, self aligned device, before (a), and after (b) sintering at 300°C for 30s. The transconductance of this device is unchanged close to pinch-off, but the sintered device shows a significant increase in transconductance at zero gate bias. Although some increase is expected due to the lower contact resistance, such an increase suggests that another effect is partly responsible. The reduced transconductance compression in curve (b) shows that the sintering has had a significant effect on the inter-electrode spaces.

![Graph showing variation of transconductance with gate bias for medium gate, self-aligned device before (a) and after (b) sinter at 300°C (drain bias = 1.5V)](image)

To explain this effect, it is assumed that the as-etched semiconductor surface has a large surface state density, which pins the Fermi level at the surface, and gives rise to a space charge region. The surface depletion caused as a result of this space charge restricts the channel between gate and drain, and causes a large transconductance compression. However, when the device is sintered, the semiconductor surface becomes stabilised by the growth of a film (of unknown composition), and the surface state density is reduced. Consequently, pinning of the Fermi level no longer occurs, and the surface depletion
disappears. Transconductance compression is therefore reduced. A similar effect has been noted by Mullin[79], who showed that the surface state density of a InGaAs-Al$_2$O$_3$ interface could be reduced by annealing, and that the Fermi level could become unpinned by this action. As a test of this explanation, some sintered devices were given a very light etch to remove this surface film, and reversion to the original characteristic was noted.

However, in order to explain this transconductance compression effect for the sintered devices, it is apparent that current saturation near zero gate bias must occur in the region between the gate and drain before it occurs under the gate, even in the absence of surface depletion. Velocity saturated current flow in a semiconductor is given by the following expression, which is a more general case of equation 4.12,

\[ I_s = q \cdot \text{Nd} \cdot Z \cdot \nu_s \cdot b \]

\[ \text{...... 5.2} \]

where \( b \) is the thickness of material available for conduction. Under the gate of a JFET, this thickness is \( a-h \), where \( h \) is the depletion depth, and \( a \) is the full channel thickness when no depletion exists. At first sight, in the source-gate and gate-drain regions, the thickness available for conduction should equal this full channel thickness, i.e., \( b = a \). Substituting the relevant material properties and dimensions for the JFETs discussed above, one obtains a velocity saturated current of roughly 100mA. Since the maximum drain current observed in these devices was roughly 80mA, the source-gate and gate-drain regions should have had little influence on device characteristics. However, two effects combine to reduce the effective thickness of these regions, and especially the latter.

Firstly, the channel either side of the gate is physically thinner than the channel under the gate as a consequence of the way the gate is formed in this structure. The step-wise manner of the gate rib etching whilst checking for negligible conductivity inevitably leads to over-etching of the rib. This over-etching can be reduced by progressively shortening the step depth, but cannot be eliminated entirely. It is estimated that the gate rib could be over-etched by 0.05\( \mu \)m in the worst case. The actual degree of over-etching in any sample is difficult to
measure directly, due to the non-ideal rib profile, and because etch depth is influenced by the proximity of the rib (trenching). This means that the channel immediately adjacent to the rib can be thinner than it is away from the rib. These complications lead to an uncertain assessment of channel thickness either side of the gate.

The other factor which contributes to the thinning of the gate-drain region is the depletion layer formed by the channel-buffer interface. The potential difference at this interface is greatest near the drain since the buffer is at source potential. Therefore the associated depletion region will become wider towards the drain, thereby reducing the thickness of channel available for conduction in this region. Although the channel thickness is reduced by this depletion layer over the whole channel, this reduction is greatest in the gate-drain spacing.

It is thought that a combination of the two factors discussed above accounts for the observed transconductance compression near zero gate bias. However, it must be pointed out that this phenomenon is relatively unimportant in FETs that operate close to threshold voltage, since compression effects are negligible in this regime. For the application of concern here, i.e., for optical receivers, the bias condition that gives lowest noise and therefore highest sensitivity is found at $0.1 - 0.2 \times I_{dss}$, and consequently the transconductance remains unaffected by compression effects.

5.1.2 Variation with gate length and electrode spacing

Due to the transconductance compression effects discussed above, valid comparison between device geometries can only be performed near threshold voltage. For this reason, the transconductance values used in this section were measured at half threshold voltage. For some device geometries, where compression effects are particularly strong, these values are lower than would have otherwise been the case. For example, comparison of the transconductance at half threshold voltage of the very short gate and short gate devices with wide spaced electrodes shows that the device type with the longer gate has the higher transconductance. This is because transconductance compression effects are very marked in
the very short gate, wide spaced device type. For this reason, analysis of the variation of transconductance with gate length has been limited to self aligned devices to minimise the errors produced as a result of compression effects.

Figure 5.9 shows the variation of transconductance at half threshold voltage with gate length, for the self aligned device type. A curve has been added which fits the three longer gate length data points, and it can be seen that the data point for the very short gate device lies a little below this curve. This is probably a consequence of the degree of transconductance compression observed with this device type.

![Figure 5.9 Variation of transconductance with gate length for self aligned device](image)

The general trend shown in this figure is consistent with most FET models. Simple models in which velocity saturation does not occur show an inverse dependence of transconductance on gate length, whereas models that only consider velocity saturated current flow show no transconductance dependence on gate length\(^{[80]}\). In the latter case, the assumption is that the saturated electron velocity does not vary with gate length. For gate lengths where velocity saturation does not occur throughout the channel, this assumption is obviously invalid. The transconductance variation with gate length shown in the figure falls
between these two extremes, which indicates that the channel is partially velocity saturated.
The device model used for this work gives a similar variation, but quantitative comparison is
not considered valid for reasons explained later.

Figure 5.10 shows the variation of transconductance at half threshold voltage with
electrode spacing for the short gate length device type.

![Graph](image)

Figure 5.10 Variation of transconductance with channel length for short gate
length device

According to simple analysis, device transconductance is degraded by the extrinsic series
resistance between the source and channel, but not by the resistance between the channel and
drain. This is because the source-channel resistance gives rise to a voltage drop given by the
product of this resistance and the drain current, which means that not all of the voltage
applied to the gate electrode controls the gate depletion region.

The degradation of intrinsic transconductance by this resistance is given by equation 4.11.
To test the validity of this equation, it was applied to two specific device types, and
rearranged to give,
where the subscripts 1 and 2 refer to the particular device types. If this equation is applied to the self aligned and narrow spaced devices in figure 5.10, then a source-channel resistance difference of 13.7Ω is obtained. End resistance measurements performed on these devices at low longitudinal field gave this difference to be 0.9Ω, which is in total disagreement with the the calculation above. Therefore, the reduction of transconductance with increased gate-source spacing cannot be explained on the basis of the simple arguments discussed above.

Another way of approaching this discrepancy is to compare the intrinsic transconductance of each device. Although this is less accurate than the previous approach due to the difficulty in evaluating the absolute value of source-channel resistance, direct comparison of intrinsic transconductance gives a better insight into the reason for this discrepancy. Using equation 4.11, the intrinsic transconductance of the self aligned device is found to be approximately 22mS, while for the narrow spaced device, it is only 17mS. It is clear that the extra channel length in the narrow spaced device is responsible for a transconductance degradation far in excess of that which would be expected from simple ohmic voltage drop.

Channel length effects have been observed in GaAs MESFETs\(^1\). These devices had a gate length of 0.25μm, and it was found that the transconductance of a device with a channel length of 0.5μm was 70% higher than a device with a channel length of 2.1μm. This difference was shown to result from the variation of saturated electron velocity with channel length, as a consequence of velocity overshoot. If this was the explanation for the behaviour of the InGaAs JFETs discussed here, then a wide spaced device configured with gate nearest source should not have a significantly different transconductance from the same device configured with gate nearest drain. This is not the case, since a 40% increase in transconductance is observed for the configuration with gate nearest source.

Velocity overshoot effects seen in submicron GaAs devices probably have little effect on the relatively long devices discussed here, and it has been shown here that these effects are
unable to explain the observed characteristics. The most likely explanation is that the 'end' regions between the channel and source/drain electrodes are not ohmic beyond current saturation, and therefore the transconductance degradation due to these 'end' regions is in excess of that due to the simple theory whereby this is caused by ohmic voltage drop. The actual degradation from this effect is difficult to model, because of the uncertainty of the shape of the channel in these 'end' regions. This is also the reason why the device model used for the design of this InGaAs JFET was not used to predict or compare with observed characteristics, i.e., the effect of these 'end' regions were not included in this model, and obviously play a crucial role in determining device characteristics.

The conclusion from the variation of transconductance with electrode geometry, therefore, matches the conclusion obtained from the variation of transconductance with gate bias, i.e., that velocity saturation occurs in the 'end' regions beyond current saturation. This effect is a consequence of the design of this device, where the channel thickness either side of the gate is no greater than the channel thickness under the gate. An obvious solution is to adopt the recessed gate structure used for high performance GaAs MESFETs. In this structure, the gate region is recessed into the channel so that the channel is thicker at the 'end' regions. This is performed in order to reduce the resistance of these regions, but in this case it would have the effect of preventing velocity saturation from occurring. This situation is easy to accomplish for a MESFET, since the Schottky barrier is formed simply by deposition of the appropriate metal in the etched recess. For a JFET, a gate formed by the localised diffusion of an acceptor species has a similar effect, since the channel under the gate is thinner than the original thickness by the diffusion depth. However, due to the lateral spreading of the acceptor species during the diffusion process, the effective gate length of such a device would be lengthened. Another disadvantage of this approach is that self-alignment of source/drain electrodes would not be so straightforward.

Whilst it is recognised that the design of this device means that close electrode spacing is even more important than it is generally considered, the superior performance shown by the self-aligned design is clear from comparison with other reported structures (see section 2.1.2).
5.2 Gate-source capacitance

The depletion layer capacitance between gate and source electrodes has a square root dependence on channel carrier concentration as obtained from equations 4.1 - 4.3, i.e.,

\[
C = \left\{ \frac{q \cdot \varepsilon_s \cdot N_d \cdot a^2}{2 \cdot (V_{bi} - V)} \right\}^{1/2}
\]

... 5.4

Direct comparison of short-gate devices from wafers of different channel carrier concentration is difficult due to the large variation of effective gate length, (this variation is apparent even for devices on the same wafer). However, comparison of measurements of test structures, and also long-gate devices, shows this square root dependence.

To test whether this dependence can be extrapolated to short-gate devices, capacitance as a function of gate length was measured for four adjacent devices. This is plotted in figure 5.11, for zero gate bias.

![Figure 5.11 Variation of capacitance with gate length at zero gate bias](image)

99
Since these points fit a straight line, this extrapolation is shown to be valid. The intercept of this line is at the origin, which indicates that the parasitic capacitance at the gate is negligible. This parasitic capacitance consists of a combination of the gate bondpad capacitance and the fringing capacitance at the edges of the gate electrode. The maximum estimated error in the gate length of each device is \(-0.25\mu\text{m}\), which would give a maximum parasitic capacitance of \(60\text{fF}\). The capacitance contributed by the gate bondpads was measured by breaking the air bridge connection. The total excess capacitance from the two bondpads was found to be \(36\text{fF}\), therefore it is estimated that the maximum fringing capacitance of this structure is \(24\text{fF}\).

No significant difference in gate-source capacitance is observed for devices with different electrode spacings. Self aligned and wide spaced devices of fixed gate length showed almost identical capacitance and CV characteristics. Furthermore, wide spaced devices configured with source/drain interchanged showed no capacitance difference over the usual configuration. This observation lends further weight to the conclusion that the inter-electrode spaces have no surface depletion, since surface depletion in these regions would contribute excess capacitance.

5.3 Output conductance

The output conductance of each device was measured at a drain bias of 1.5V, and a gate bias of half threshold voltage. For all of the working devices, the average value for each device type was obtained. These results are shown in figure 5.12 below. The upper curve is for self-aligned devices, and the lower curve is for narrow spaced devices. Obviously, the output conductance is increased for short gate length and close electrode spacing. For optical receiver applications, an output conductance of less than \(2.5\text{mS}\) is desirable (see section 4.2). Comparison of the output conductance value of the short gate, self-aligned device with that obtained on the equivalent device discussed in section 4.2 shows that the general level of output conductance on this wafer is high.

It is apparent that the device geometry most suited to give high transconductance and low capacitance, also gives the highest output conductance.
A non-zero drain conductance is a consequence of the two-dimensional nature of the electron velocity. Due to the high transverse field at the drain end of the gate, the velocity vector is essentially vertical at this point, at the onset of current saturation. However, as the drain bias is increased, the longitudinal component of the field is increased, and therefore the longitudinal electron velocity is allowed to increase, even though the resultant electron velocity is saturated. Therefore the non-zero drain conductance is caused by a rotation of the electron velocity vector. To appreciate how this would affect drain conductance as a function of gate length is not clear, and would require device modelling using a two-dimensional computer simulation, which is outside the scope of this work.

Another effect which causes non-zero drain conductance is parallel conduction through the substrate or buffer layer, which can occur if the electrons have sufficient energy to overcome the potential barrier at this interface. Intuitively, one would expect that the higher electric fields associated with a short gate device would lead to higher electron energy and therefore give rise to more conduction through the buffer/substrate.
5.4 Gate leakage current

Gate leakage current at a bias of -1V was measured for each device, and the mean value for each device type is given in table 5.2. However, considerably more care is needed when comparing gate leakage current variation as a function of device geometry and channel doping than with other measured parameters. This difference can be illustrated by comparing the distributions of the data populations. Data obtained from semiconductor device parameter measurement usually follow lognormal distributions, and a good example of this can be seen in figure 5.13, which shows a plot of the distribution of transconductance at half threshold voltage for the long gate length, self-aligned device type. The Y-axis of this plot is the cumulative probability that the population will have less than a given value, and is arranged in such a way that a normal distribution gives rise to a straight line. Since the X-axis has a

![Cumulative probability vs. transconductance](image-url)

Figure 5.13 Distribution of transconductance at half threshold voltage data for long gate length, self-aligned device
logarithmic scale, a straight line is therefore evidence of a lognormal distribution. It can be seen from the figure that this data follows such a distribution.

However, if the leakage current data is plotted in a similar way, then this ideal behaviour is not observed, and a good example of this is shown in figure 5.14.

Figure 5.14 Distribution of gate leakage current at -1V data for short gate length, narrow spaced device

This distribution shows bi-modal behaviour, which is assumed to be a consequence of different gate leakage current mechanisms for each mode. The mean (or median) of the total population is therefore no longer a valid representation of the sample. In order to obtain a valid comparison between types of device, the mean of the mode of interest must be chosen. It is because the distributions of the other device parameters are essentially single-moded that a mean of the whole population can be used successfully in these cases.
Low leakage current operation is of interest here, therefore the mean of the low current mode was evaluated for each device type, and has been plotted as a function of gate length in figure 5.15, for the narrow spaced device types.

![Graph showing the variation of mean gate leakage current with gate length for narrow spaced device](image)

**Figure 5.15** Variation of mean gate leakage current with gate length for narrow spaced device

Leakage current in pn junctions can arise through various mechanisms. Bulk leakage can occur through diffusion, recombination-generation, or tunneling processes, and, in addition, surface leakage at the edges of the junction can add significantly to the total. Bulk and surface components can be isolated since bulk leakage is proportional to the device area, and surface leakage is proportional to the device perimeter. In the present case, the gate area is proportional to the gate length, while the device perimeter remains effectively constant due to the large gate width to length ratio. Therefore, referring to figure 5.15, the bulk leakage current can be obtained from the slope of the plot, and the surface leakage current component is given by the intercept. A bulk current density of $10^{-3} \text{A/cm}^2$, and a surface leakage current of 10nA are obtained from this figure.
Concentrating on the bulk leakage current, the contributions from the three processes mentioned earlier are assessed.

Diffusion current is caused by thermally generated minority carriers diffusing into the depletion region, and an approximate expression for this (for a one-sided p+n abrupt junction) is given by \[^{[80]}\],

\[
J_d = q \cdot \left( \frac{n_i^2}{N_d} \right) \cdot \left( \frac{D_p}{\tau_p} \right)^{1/2}
\]  \(...... 5.5\)

where \(n_i\) is the intrinsic carrier concentration, \(D_p\) is the hole diffusion coefficient, and \(\tau_p\) is the hole lifetime. These quantities have been tabulated by Forrest, \[^{[82]}\] and are given below,

\[
\begin{align*}
n_i &= 5.4 \times 10^{11} \text{ cm}^{-3}, \\
D_p &= 13.0 \text{ cm}^2/\text{s}, \\
\tau_p &= 200 \text{ ps}.
\end{align*}
\]

When these quantities are substituted into equation 5.5, then for a carrier concentration of \(5 \times 10^{16} \text{ cm}^{-3}\), and a device area of \(5 \times 10^{-6} \text{ cm}^2\) (gate length of 2 \(\mu\)m) gives a diffusion current of,

\[
I_d = 1.2 \text{ pA}.
\]

Clearly, the diffusion current component to the gate leakage is insignificant.

Leakage current caused by band-to-band tunneling can occur at high reverse bias due to the proximity of the bands in the depletion region under this condition. The tunneling current density is given by \[^{[80]}\],

\[
J_t = \gamma \cdot \exp \left( \frac{4}{(2 \cdot m^*) \cdot E_g^{3/2}} \right) \cdot \left( \frac{3 \cdot q \cdot E_m \cdot \hbar}{E_g} \right)^{-1/2}
\]  \(...... 5.6\)

where \(m^*\) is the electron effective mass, \(E_g\) is the bandgap, \(E_m\) is the maximum electric field, \(\hbar\) is the reduced Planck constant, and the prefactor \(\gamma\) is given by,
\[
\gamma = \frac{(2 \cdot m^*) \cdot q^2 \cdot E_m \cdot V}{4 \cdot \pi^2 \cdot h^2 \cdot E_g^{1/2}}
\] 

where \( V \) is the applied voltage. The maximum electric field at the junction is given by,

\[
E_m = -2 \left( V_{bi} - V \right) / W
\]

where \( W \) is the depletion width, given by,

\[
W = \sqrt{\frac{2 \cdot \varepsilon_s \cdot [ V_{bi} - V ]}{(q \cdot N_d)}}
\]

The effective mass ratio, \( m_e \), used in this work was 0.041 \[^3\]. Using the above expressions, tunneling current was found as a function of carrier concentration in the channel for different applied bias levels, and is shown in figure 5.16 for a 2\( \mu \)m gate length.

![Figure 5.16](image)

**Figure 5.16** Variation of tunneling current with carrier concentration as a function of bias
At a bias of -1 V, significant tunneling current (> 10nA) does not occur until the channel doping exceeds $3 \times 10^{17}$ cm$^{-3}$. Since the channel carrier concentration in these devices is considerably lower than this level, the conclusion is that band-to-band tunneling current is insignificant in this case. However, it is possible that tunneling via mid-band states contributes to the current, but this has been left for future work.

Since diffusion and tunneling currents have been shown to be insignificant in these devices, then generation-recombination must be the mechanism for the bulk leakage currents observed. This can be verified by measurement of forward bias current flow. The current density in the forward direction can be represented by the empirical relation,

$$J_f = \exp \left\{ \frac{(q \cdot V)}{(n \cdot k \cdot T)} \right\} \quad \ldots 5.10$$

where the factor $n$ is close to unity when diffusion current dominates, and is close to two when generation-recombination current dominates. This relation can be re-arranged to give,

$$n = \frac{q}{(k \cdot T)} \cdot \left( \frac{\ln J_f}{V} \right)^{-1} \quad \ldots 5.11$$

Figure 5.17 shows the variation of the logarithm of the forward current of a typical narrow spaced, short gate length device as a function of bias. The slope of this variation is used in the relation above, to give $n = 1.85$, which validates the assumption that generation-recombination is the main leakage mechanism.

Generation and recombination of electron-hole pairs in the depletion region are responsible for this type of leakage current, and the rates of these processes are greatly increased by a concentration of mid-band trapping centres. Generation-recombination current density is given by,

$$J_{rg} = q \cdot n_i \cdot W / \tau_{eff} \quad \ldots 5.12$$

where $\tau_{eff}$ is the effective carrier lifetime. This quantity is inversely proportional to the density, $N_t$, and cross section, $\sigma_t$, of the traps present in the material, and is therefore
heavily dependent on the material quality. If the bulk leakage current shown in figure 5.15 is due to generation-recombination, then $J_{rg} = 10^{-3}$ A/cm$^2$. Re-arranging equation 5.12, and substituting this value of $J_{rg}$ gives,

$$\tau_{\text{eff}} = 1.8 \text{ nS}$$

Takanashi and Horikoshi have collated effective lifetime data from several sources, including their own, and fitted the relation, $\tau_{\text{eff}} = 4.10^8 / N_d$ to the data$^{[83]}$. For a carrier concentration of $5.10^{16}$ cm$^{-3}$, this leads to $\tau_{\text{eff}} = 8.0$ nS, which is within an order of magnitude of the value obtained from the present work. One conclusion from their work was that the density of the trap responsible for the generation-recombination current was proportional to the doping density, and the ratio $N_t / N_d$ was of the order of 0.001. The origin of this trap is not known.

The intercept in figure 5.15 of 10nA is a consequence of surface leakage. To measure the leakage current from the gate bondpads, the air bridge was broken on some devices. At a
bias of -1V, a bondpad leakage current of less than 1nA was measured. This current is a consequence of the finite resistivity of the semi-insulating substrate, which allows the bondpads to contribute to the leakage current up to a limit imposed by the resistance of the substrate between the gate and source electrodes.

The remainder of the surface leakage current arises from conduction across the region between the gate rib and source contact. Surface states can act as recombination centres, which can give rise to surface conduction. Also, adsorbed species on the semiconductor surface can lead to ionic conduction, and will modify the surface state density. However, the precise nature of this conduction mechanism is unclear.

The observations above relate to the narrow spaced device types. Clearly, the surface leakage current will depend on the spacing between electrodes. No significant difference was measured between wide spaced devices configured for either a 2µm or 5µm spacing. However, the leakage current measured for the self aligned devices showed an order of magnitude increase on this level.

Also, since the transconductance improvement after sintering was attributed to a change in surface condition in section 5.1.1, some variation in the surface current would be expected as a consequence of this process. Taking the short gate length devices to illustrate this, the mean leakage current of the narrow spaced devices was reduced after sintering, while the mean leakage current of the self aligned devices showed an order of magnitude increase. These results are summarised in table 5.3 below.

<table>
<thead>
<tr>
<th>Spacing</th>
<th>Mean leakage current</th>
<th>post-sinter</th>
</tr>
</thead>
<tbody>
<tr>
<td>narrow spaced</td>
<td>32 nA</td>
<td>15 nA</td>
</tr>
<tr>
<td>self aligned</td>
<td>200 nA</td>
<td>4.94 µA</td>
</tr>
</tbody>
</table>

Table 5.3 Variation of mean leakage current with electrode spacing and sinter for short gate length devices
If the mechanism for surface conduction is via surface recombination, then the reduction of surface state density after sintering is a possible explanation for the reduction in leakage current observed for the narrow spaced devices. However, the extremely high leakage current observed for the self aligned devices cannot be explained so simply. The vast majority of this current is due to surface conduction, since there is not much variation in the level of this current with gate area. One tentative explanation for this behaviour is that the magnitude of the current conduction via the surface states is critically dependent on the electric field, which is highest for the self aligned devices. However, if the surface state density is reduced by the sintering process, then the large leakage current increase measured in this device type is anomalous, and an additional mechanism is required for explanation.

Since the leakage current behaviour of these devices is not a majority carrier phenomenon, some degree of comparison between wafers was valid. Therefore, the variation of the mean value of this parameter was plotted as a function of carrier concentration in the channel, and is shown in figure 5.18 for the short gate, narrow spaced device type.

![Graph showing gate leakage current vs carrier concentration](image)

**Figure 5.18** Variation of gate leakage current at -1V with carrier concentration for short gate length, narrow spaced device
An increase in leakage current with carrier concentration is to be expected if recombination-generation current is dominant. This is because the effective carrier lifetime has been shown to be inversely proportional to carrier concentration, as noted previously. If the surface leakage current contribution is assumed to be independent of carrier concentration, then the increase in recombination-generation current with carrier concentration can be estimated from the effective lifetime obtained from the measurements discussed previously, i.e., 1.8nS for a carrier concentration of $5 \times 10^{16}\text{cm}^{-3}$. Using this data to extrapolate for a carrier concentration of $3 \times 10^{17}\text{cm}^{-3}$, then a recombination-generation current of 13nA is deduced. From figure 5.18, it is clear that the increase in leakage current exceeds that from recombination-generation current. Therefore, surface leakage current must also increase with carrier concentration.
5.5 Microwave measurements

Microwave measurements were performed on two of the short gate length device types - those with self-aligned and narrow-spaced source/drain electrodes - as described in section 4.1.3.

The equivalent circuit values extracted from the S-parameter measurements are shown in table 5.4 for the self-aligned device and in table 5.5 for the narrow-spaced device, as a function of gate bias.

<table>
<thead>
<tr>
<th>$V_{gg}$ (V)</th>
<th>$g_m$ (mS)</th>
<th>$C_{gs}$ (pF)</th>
<th>$r_{ds}$ (Ω)</th>
<th>$f_T$ (GHz)</th>
<th>NF$_{50}$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>30.9</td>
<td>0.45</td>
<td>70</td>
<td>9.5</td>
<td>8.0</td>
</tr>
<tr>
<td>-0.5</td>
<td>28.3</td>
<td>0.40</td>
<td>87</td>
<td>10.2</td>
<td>8.2</td>
</tr>
<tr>
<td>-1.0</td>
<td>24.2</td>
<td>0.36</td>
<td>93</td>
<td>9.7</td>
<td>8.7</td>
</tr>
<tr>
<td>-1.5</td>
<td>20.0</td>
<td>0.33</td>
<td>97</td>
<td>8.6</td>
<td>9.3</td>
</tr>
<tr>
<td>-2.0</td>
<td>15.5</td>
<td>0.31</td>
<td>107</td>
<td>7.4</td>
<td>10.0</td>
</tr>
<tr>
<td>-2.5</td>
<td>10.3</td>
<td>0.29</td>
<td>132</td>
<td>5.5</td>
<td>10.9</td>
</tr>
<tr>
<td>-3.0</td>
<td>4.7</td>
<td>0.29</td>
<td>186</td>
<td>2.6</td>
<td>12.6</td>
</tr>
</tbody>
</table>

Table 5.4 Variation of microwave parameters with gate bias for self-aligned device at a drain bias of 1.5V

<table>
<thead>
<tr>
<th>$V_{gg}$ (V)</th>
<th>$g_m$ (mS)</th>
<th>$C_{gs}$ (pF)</th>
<th>$r_{ds}$ (Ω)</th>
<th>$f_T$ (GHz)</th>
<th>NF$_{50}$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>11.0</td>
<td>0.40</td>
<td>53</td>
<td>3.8</td>
<td>11.6</td>
</tr>
<tr>
<td>-0.5</td>
<td>15.6</td>
<td>0.35</td>
<td>68</td>
<td>6.1</td>
<td>8.8</td>
</tr>
<tr>
<td>-1.0</td>
<td>18.7</td>
<td>0.32</td>
<td>92</td>
<td>8.3</td>
<td>8.4</td>
</tr>
<tr>
<td>-1.5</td>
<td>16.9</td>
<td>0.30</td>
<td>104</td>
<td>8.2</td>
<td>9.0</td>
</tr>
<tr>
<td>-2.0</td>
<td>13.5</td>
<td>0.28</td>
<td>113</td>
<td>7.2</td>
<td>9.7</td>
</tr>
<tr>
<td>-2.5</td>
<td>9.2</td>
<td>0.27</td>
<td>133</td>
<td>5.4</td>
<td>10.8</td>
</tr>
<tr>
<td>-3.0</td>
<td>3.9</td>
<td>0.27</td>
<td>180</td>
<td>2.6</td>
<td>12.4</td>
</tr>
</tbody>
</table>

Table 5.5 Variation of microwave parameters with gate bias for narrow-spaced device at a drain bias of 1.5V
The drain bias used for these measurements was 1.5V. Also included in these tables are noise figure measurements (in a 50Ω system) obtained at the same bias conditions.

These measurements were useful in two respects. Firstly, they were used to validate the dc measurements, and secondly, they were used to obtain device parameters that were difficult to measure under dc conditions.

To check the validity of the parameters derived from dc measurements, the transconductance values measured at dc and rf were compared for the self-aligned device. This is shown in figure 5.19, where the rf data points are superimposed on the dc data line.

![Graph showing transconductance vs gate bias](image)

**Figure 5.19** Variation of transconductance with gate bias for self-aligned device. rf data as points, dc data as line.

There is obviously good agreement between these data sets, which gives confidence that the parameters measured at dc are valid.

Measurement of the drain-source resistance beyond saturation, $r_{ds}$, is extremely sensitive to the bias conditions. A small change in the gate bias can produce large changes in $r_{ds}$. The values of $r_{ds}$ shown in tables 5.4 and 5.5 are lower than are measured at dc, which could be
attributed to this bias sensitivity. However, GaAs MESFETs also show this behaviour. The transition frequency at which low values of \( r_{ds} \) are measured in GaAs MESFETs is known to be a function of substrate quality, i.e. the higher the substrate quality, the higher the transition frequency, which is usually in the range 10kHz-1MHz. The \( r_{ds} \) measurements on these InGaAs JFETs are therefore not anomalous in this respect.

Gate-source capacitance, \( C_{gs} \), is not easy to evaluate by direct measurement due to the contribution that the gate-drain capacitance, \( C_{gd} \), makes to the measurement. The capacitance measurements discussed earlier were taken with an unconnected drain electrode, in order to minimise the effects of \( C_{gd} \). The \( C_{gs} \) values given in tables 5.4 and 5.5 are independent of \( C_{gd} \) and are therefore more realistic. However, the directly measured \( C_{gs} \) values are close to those derived from S-parameters, which shows that \( C_{gd} \) has only a small effect.

Also included in tables 5.4 and 5.5 are values of \( f_T \), the frequency at which the current through \( C_{gs} \) is equal to the current generator in the intrinsic FET, and is given by

\[
f_T = \frac{g_m}{2 \pi C_{gs}}.\]

This is in excess of 10GHz for the self-aligned device at optimum bias. This is a useful benchmark for conventional microwave transistors, however, as stated previously, the figure of merit for transistors used in optical receivers is inversely proportional to the square of the gate-source capacitance. Therefore \( f_T \) is not the most appropriate figure of merit for this application.

The 50Ω noise figure, \( NF_{50} \), in dB, is plotted in figure 5.20 for both device types as a function of gate bias. This relates to the coefficient \( P \) in the Pucel noise model (appendix 1) by the relation,

\[
NF_{50} = 1 + \left( \frac{R_g + R_s + P}{g_m} \right) / 50 \]

...... 5.13

if the channel noise is dominant, especially if the measurement is performed below the induced gate noise break frequency (=4GHz). Also included in figure 5.20 is the expected variation of \( NF_{50} \) with gate bias according to the model using this relation.
The noise figure of the self-aligned device has a minimum at zero gate bias of 8 dB, and is comparable to the noise figure observed in GaAs MESFETs of similar dimensions. However, the noise figure increases with gate bias, which is contrary to expectations. Normally, in a GaAs MESFET, minimum noise occurs close to threshold voltage, and this is backed by the modelled variation in figure 5.20. The noise figure variation of the narrow-spaced device shows a minimum at a gate bias of -1 V. It is clear from these measurements that minimum 50 Ω noise figure occurs at maximum transconductance, for each device type. The reason for this difference in behaviour with GaAs MESFETs is not understood. The high noise figure near threshold voltage has important implications for the use of this device in optical receiver applications, and should be the subject of further investigation.
Chapter 6. Monolithic Integration of InGaAs JFET and PIN-photodiode

Having successfully developed a discrete InGaAs JFET, the next stage was to demonstrate its potential for monolithic integration. In section 1, the design of the integration scheme with an InGaAs PIN-photodiode is illustrated, and the performance of this structure is discussed in section 2.

6.1 Design

The PIN-photodiode is used extensively in receivers for optical communication systems because it is simple to use and construct, and gives very good performance when used with a good low noise front end amplifier.

This device consists of a thick intrinsic layer, bounded by a p-type layer and an n-type layer. The intrinsic region has a carrier concentration equal to the background level of the growth technique, and is usually n-type. The intrinsic layer is therefore partially depleted, due to the pn junction at the p-type layer interface, and the device is designed so that this depletion region extends over the whole of the intrinsic layer when biased. When photons are absorbed in this layer, the electron-hole pairs produced are separated by the high electric field, and give rise to photocurrent in an external circuit.

The simplest approach to PIN-photodiode design is the mesa-style device. A schematic cross section of a typical design is shown in figure 6.1. This 'mesa' type of device is simple to construct, especially when the pn junction is grown-in by epitaxy, but has the disadvantage of having the pn junction exposed to the atmosphere. This gives unstable operation unless the device is well passivated, or hermetically sealed in a suitable ambient. However, for the purposes of this work, it was regarded as a suitable device for demonstration of monolithic integration.

The top-entry style of device was adopted for this work, where the incident photons enter through the top of the photodiode. Although this style of device has relatively low quantum efficiency due to the absorption of photons in the p-InGaAs top layer, this approach gives ease of coupling for incident light.
Of the design parameters that affect device performance, those pertaining to the intrinsic region are most important. This layer must be relatively thick (several microns), so that a high proportion of the incident photons are absorbed. Also, the carrier concentration must be very low, so that this layer can be fully depleted. This ensures low device capacitance, high speed (because of minimal absorption in low field regions), and low bias operation, (which gives lower leakage current).

The PIN-photodiode described above is not very compatible with the rib-JFET described in chapter 3. The most obvious difference is that the JFET structure is grown on a semi-insulating substrate for low parasitics and minimum device interaction, while the PIN is grown on a highly conducting substrate for ease of contact. Another substantial difference is that the FET channel must be thin and moderately doped to give high transconductance, and practical threshold voltage, while the PIN intrinsic layer must be thick and undoped to give high quantum efficiency, low capacitance, and low bias operation.

These incompatibilities impose severe constraints on design, if high performance is to be achieved. In order that each device could be optimised individually, the structure was designed so that separate layers were grown for each device. To achieve this, the PIN structure was modified to make it more compatible with the FET requirements, in the following ways,
i) due to the difficulties encountered when more than one growth technique is used for device fabrication (hybrid growth), the whole structure was grown by MBE, although this technique had shown no aptitude for the growth of good PIN structures,

ii) the PIN-photodiode structure was modified by incorporation of an epitaxial n-type contact layer, so that growth on semi-insulating substrates could be used,

iii) this design used a grown-in junction for the formation of the photodiode. This enabled simplified fabrication, and was well suited to the integration scheme.

By these means, the structure was designed so that the FET layers were grown directly on top of the PIN layers. This orientation was preferable because the FET layers, being thinner than the PIN layers, allowed easier interconnection due to the lesser step between devices. Also, the PIN p-type layer and the FET p-type buffer layer were combined in one dual-purpose layer.

Another important innovation was the use of a remote bondpad for the PIN-photodiode, which takes advantage of the isolation provided by the semi-insulating substrate. This bondpad contributes negligible parasitic capacitance or leakage current because of this isolation, and provides simpler and higher yield wire-bonding because it can have large area, and is remote from sensitive device regions. Also, the PIN-photodiode area can be potentially very small, since the bondpad was not included in the active area, which gives low capacitance and low leakage current, and therefore high performance.

A schematic cross section of the PIN-FET structure is shown in figure 6.2. The FET gate is connected to the PIN anode in this design, for relative ease of construction. Details of the fabrication scheme will not be given here. The FET is constructed in the same way as the discrete device presented in chapter 3, and if a cross section is taken perpendicular to the one in figure 6.2, the structure of the rib JFET would be evident. The PIN-photodiode was defined by mesa etching, and the device isolation used the air bridge technique discussed in chapter 3.

The layout of this circuit can be seen in the photograph of a completed wafer, and this is shown in figure 6.3.
The circuit configuration consists of two PIN-photodiodes and two JFETs, connected so that the JFETs have a common source. In the photograph of this structure, the two FET gate ribs can be seen across the large central mesa that incorporates the two drain electrodes (top and bottom) and the common source (centre). The gate/PIN p-contact bondpads can be seen to the left of the gate ribs, connected by air bridges. The PIN-photodiodes are located to the left of these bondpads, and their n-contacts are adjacent, and can be seen at the top and bottom of the photograph.
6.2 Performance

6.2.1 FET performance

The concept of using a p-type buffer layer for the JFET was proven to be effective when device characteristics for the integrated structure showed little difference to those of the discrete counterpart. Because of this similarity, detailed characterisation of the integrated JFET will not be presented in this section. Instead, a simple presentation of results is given, with special emphasis on any differences that have been observed.

Output and transfer characteristics of one of the best integrated JFETs are shown in figures 6.4 and 6.5 respectively. The device had gate dimensions of 1.5 x 200μm.

![Figure 6.4 Output characteristic of integrated JFET](image)

Comparison of these characteristics with those of a discrete device (chapter 4) shows that no degradation has occurred, which indicates that the p-buffer layer is effective in providing isolation between the JFET and the parasitic effects of the underlying PIN layers. Maximum transconductance occurs at zero gate bias, and is 34mS. When scaled for unit width, a value of 170mS/mm is obtained, which is the highest reported figure for a monolithic InGaAs FET to date.
Refering to table 2.3 it is evident that the transconductance of the monolithically integrated JFET described in this work represents a threefold improvement on similar devices reported by other workers. This is partly due to the high quality of epitaxy at BTRL, but the major factor in this significant improvement is due to the 'vertical' integration design, which allows the JFET to be optimised independently.

Gate leakage current of the as-processed device is shown in figure 6.6 (curve a). This level of leakage current is much greater than the level normally found with the discrete device. The cause of this excess current was attributed to the more involved post rib-etch processing of the integrated device. The oxide layer formed on the exposed rib/channel surface which results from this processing, gives rise to a surface leakage current component. If this oxide layer is removed using a dilute etchant, then the leakage current is reduced to a level similar to that of the discrete device. Figure 6.6 also shows the effect of a 3 second dip in 1:1:50 on the leakage current (curve b). At -0.5V, the level of leakage current is reduced from 1μA to 25nA.
Figure 6.6 Gate leakage current variation with bias. Curve a - as processed; curve b - after light etch.

However, because the source-channel and drain-channel regions are eroded by this process, the channel series resistance is increased, which leads to a decrease in transconductance of approximately 10%.

Capacitance values measured at 1MHz and an oscillator voltage of 10mV are shown as a function of gate bias in figure 6.7. The main point to note from this figure is that no excess capacitance is introduced as a result of the underlying PIN layers. Neglecting the fringing capacitance, the expected gate-source capacitance can be calculated from simple theory (see section 4.1.2). For a channel doping level of $10^{17}$cm$^{-3}$, and a built-in potential of 0.6V, a depletion width of 0.093μm is calculated. A total gate area of 4.10$^{-6}$cm$^2$ (which consists of 3.10$^{-6}$cm$^2$ for a 1.5 x 200μm gate added to a total bridge support area of 1.10$^{-6}$cm$^2$), gives a value of gate capacitance equal to 0.50pF at zero bias. The measured zero bias capacitance of 0.51pF agrees well with this calculation, suggesting that there is no excess capacitance associated with the design.
6.2.2 PIN performance

The current-voltage, capacitance-voltage, and optical characteristics of the PIN-photodiode integrated monolithically in this structure show considerable deviation from ideal behaviour. However, since this device was intended as an illustration of the integration capabilities of the JFET, this structure was regarded as a suitable end-point for this work. Details of this anomalous behaviour are discussed below.

The current-voltage characteristics of a typical integrated PIN-photodiode are shown in figure 6.8. The reverse bias characteristic shows a high level of leakage current, which could either be due to a large surface current caused by processing, or a large bulk current caused by material inadequacies. The dominant source is uncertain, however. The forward characteristic shows a very sharp turn-on at an unusually high bias, instead of the normal exponential turn-on at the built-in potential. These anomalies have been interpreted by assuming non-ohmic behaviour at either or both contacts, resulting in a back-to-back Schottky contact and PIN in series.
It is well known that contact to p-InGaAs gives a barrier height of 0.5V, but that ohmic contact can be achieved by having a high acceptor concentration at the contact to give tunneling behaviour. However, measurements of Au-Ge contacts to the undoped InGaAs layer in the PIN-FET structure also showed rectifying behaviour. If the assumption of a series combination of a Schottky contact and PIN is correct, then the sharp turn-on is due to premature breakdown of the Schottky contact.

Figure 6.9 shows the variation of PIN capacitance with forward and reverse bias. It is clear from this figure that there is considerable deviation from ideal behaviour. Assuming the interpretation discussed above, then one would expect a V-shaped characteristic in this figure, due to the dominance of the junction under reverse bias.

The flatness of the characteristic under reverse bias suggests that there is an additional effect to consider. One possible explanation for this behaviour could be that the p-type layer is heavily compensated, giving semi-insulating properties. This layer would then determine the capacitance of the structure when the Schottky contact is forward biased, and would not vary with bias. When the Schottky contact is reverse biased, the capacitance would be
determined by the fixed p-type layer capacitance in series with the Schottky contact capacitance, which will vary with bias. The combination of these two situations would give rise to the type of characteristic observed. Further evidence to support this hypothesis comes from the optical properties of the device.

In order to determine the quantum efficiency of the photodetector, it was packaged onto a TO-5 header and placed on a micromanipulator mounted on an optical bench. Light from a 1.53μm semiconductor laser was chopped and focussed on the photodetector, and the short circuit photocurrent was recorded as a function of bias. Figure 6.10 shows the variation of photocurrent with bias for an incident optical power of 910μW.

There are two points to note from this figure. Firstly, at a bias of -6V, the responsivity of the photodetector is 0.90A/W, which leads to an external quantum efficiency of 73%. Because the device has no anti-reflection coating, this figure infers an internal quantum efficiency of greater than 100%. Therefore this device must have gain.
Secondly, the variation of photocurrent with bias is linear. If the device acted as a PIN photodiode, the photocurrent would show an exponential saturation with bias. However, a photoconductor shows a linear photocurrent dependence on bias due to the linear increase in transit time of the carriers. Also, a photoconductor can show gain if the carrier transit time is less than the carrier lifetime.

The inference from these observations is that the device is behaving like a photoconductor. A photoconductor consists of a layer of undoped or semi-insulating semiconductor with ohmic contact to the ends. Photogenerated carriers can travel around the closed circuit many times before they recombine, thereby producing gain. In this instance, the compensated p-type layer acts as the photoconductor, which effectively has ohmic contact on both sides when the structure is reverse biased.

All of the PIN-photodiode measurements discussed above show anomalous behaviour, which can be interpreted as being due to non-ohmic contacts and compensated p-type region. A simple solution to this problem would be to increase the carrier concentration in this layer.
to avoid compensation and to facilitate ohmic contact, which has been shown not to adversely affect the characteristics of the JFET.

6.2.3 PIN-FET characterisation

To demonstrate the operation of the monolithic PIN-FET as a photoreceiver, the structure was packaged in a TO-5 header and mounted on an optical bench. Due to the limitations of the photodetector discussed above, dc measurements of the photoreceiver were considered suitable for demonstrating that the components were connected and behaving as photodetector and amplifier.

A gate bias resistor of 1000Ω, and a drain load resistor of 270Ω, were connected in a common source arrangement. To obtain maximum transconductance, the JFET was operated at zero gate bias, and the drain bias was adjusted to give a drain current of 80mA. Light from a 1.53μm laser was chopped and focussed onto the photodetector, at an incident power level of approximately 50μW. The voltage at the JFET gate resulting from the photocurrent flowing through the bias resistor, and the output voltage at the drain were monitored on an oscilloscope. A photograph of the resulting traces is shown in figure 6.11, where the lower trace shows the voltage at the input, and the upper trace shows the voltage at the output.

The voltage gain of the JFET measured from these traces is 1.7. The transconductance of the device used in this instance was 25mS, which should have given a gain of more than 6 with a load resistor of 270Ω. However, the low output impedance of the device (100Ω) gave an effective load resistance of only 70Ω, which gave rise to the measured gain of 1.7. Under normal operating conditions, i.e., near pinch-off, the output impedance of the JFET is much higher, and will present less of a problem in this respect.

To summarise, the characteristics of the JFET have been shown not to be significantly degraded by transfer from the discrete to the integrated structure. The PIN photodiode requires further work to achieve optimum performance. However, the monolithic integration of two very different styles of device has been demonstrated, which has been achieved through novel device design.
Figure 6.11 Photoresponse of PIN-FET structure. Upper trace shows voltage at the output, and lower trace shows voltage at FET gate as a result of 50μW incident optical power.
Chapter 7. Conclusions

The InGaAs JFET described in this thesis has demonstrated high performance and an aptitude for monolithic integration, and therefore meets the objectives of this work. In terms of transconductance, this JFET greatly surpasses all other published work on this type of device, as discussed in chapter 2. Furthermore, the integrated device showed negligible difference to its discrete counterpart. The design features which account for these achievements are re-iterated below:

* The rib-gate design allows the fabrication of devices with short gate length because of the undercutting of the gate metal. The benefits of short gate length are increased transconductance, reduced gate-source capacitance, and reduced noise. Because of this undercut, the effective gate length is significantly shorter than the photolithographically-defined gate metal stripe. This is in contrast to the more usual method of gate formation, where a pn junction is produced by zinc diffusion, which leads to a gate dimension which is lengthened by lateral diffusion. The gate-rib technique therefore can produce short gate length devices without the need for fine-line photolithography. This is an important factor since mask aligners needed for small dimension work may not be available.

* The self-alignment of source and drain contacts, (facilitated by the gate-rib structure), maximises the extrinsic transconductance by reducing the parasitic resistance between source and channel.

* The use of a p-type buffer layer allows good electron transport in the channel by reducing the deleterious effects of the substrate. The effectiveness of this type of buffer layer is not impaired by significant excess conduction in this structure, i.e. this buffer layer shows good electron confinement. Another advantageous feature of the p-type buffer layer was the possibility of a simple integration scheme with the photodetector

* The air bridge design gives a stable, reproducible method of interconnection between mesas. This improves device yield and produces a simple but effective means of
integration. The simplified processing allowed by this technique gives higher yield of finished wafers, and the reproducibility ensures high yield of devices on each wafer.

These design features combine to give an easily fabricated, high performance device with suitability for monolithic integration. Although the gate-rib technique had been published previously \[35\], the self-alignment of source and drain metal was a novel aspect of this work at the time. The air bridge formed by the full undercut of semiconductor from below a metal track was also a novel feature of this work. These features have since been used by other workers \[38,87\]. The p-type buffer layer concept had been used experimentally in some GaAs MESFETs prior to this work, and had been used successfully by Yamasaki et al \[88\] to suppress substrate current in sub-micron gate length GaAs MESFETs. When this idea was introduced into the present work, it had not been used for InGaAs FETs. However, due to the difficulties encountered with dopant redistribution (see section 3.3.1), the first publication of this approach was made elsewhere \[38\].

Although the gate-rib approach can lead to short gate length devices, a consequence of this design is that the 'end' regions between source and gate and between gate and drain are necessarily no thicker than the channel thickness. This leads to problems of increased 'end' resistance, which degrades extrinsic transconductance unless the electrodes can be self-aligned, as in this design. Another problem caused by the thin 'end' regions with this design is premature electron velocity saturation in these regions, which can dominate the electrical characteristics of this device and give rise to transconductance compression near zero gate bias. Sensitivity to surface effects is also a potential difficulty in this type of device.

To overcome these problems, a recessed gate or buried channel design must be implemented, which, in practice, must use either diffusion or ion implantation to produce the gate. Both of these techniques, however, lead to gate length broadening, and consequently, the intrinsic performance of these device types will be inferior to the rib-JFET.

Therefore a trade-off between intrinsic and extrinsic performance results. The precise nature of this trade-off depends on the operating point of the device, which will depend on
the application. For the optical receiver, the operating point is usually chosen close to threshold voltage, where the degradation in extrinsic performance is not significant in the rib--JFET. In this situation, the advantage gained in having high intrinsic performance is not diminished appreciably. Even for operation close to zero gate bias, the self-aligned rib--JFET shows acceptable extrinsic performance.

Maximum transconductance is used as the figure of merit in most of the InGaAs FET literature. Although the device discussed here is superior in this respect, a proper comparison is difficult due to the lack of detailed characterisation in the literature. A more worthwhile comparison is obtained from observation of GaAs MESFET characteristics, since this type of device forms the basis of the hybrid receiver. A 2µm gate length GaAs MESFET was obtained from elsewhere in BTRL for this purpose. Comparison of characteristics could be made at more appropriate bias conditions by this means, i.e. close to threshold. A typical standard device in this work has a maximum transconductance of 135mS/mm, and a transconductance at half threshold voltage of 80mS/mm (although much higher values have been obtained). This compares with 100mS/mm and 70mS/mm for the GaAs MESFET at these bias conditions. Gate-source capacitance for the GaAs device was approximately 50% higher, giving this InGaAs JFET a considerable performance advantage.

It has been demonstrated that the device discussed in this thesis has considerably better performance than other published InGaAs JFETs/MESFETs, and also has better performance than GaAs MESFETs of similar dimensions. However, commercially available GaAs MESFETs have gate lengths less than 0.5µm, which gives greatly enhanced performance, which cannot be matched by the relatively long gate length devices produced in this work.

Although the transconductance obtained with the self-aligned device is greater than that obtained with the spaced-electrode device, some parameters of secondary importance were degraded by self-alignment. In particular, gate leakage current and output conductance are relatively high in this type of device. In some cases, the preferred device may be the spaced-
electrode type, especially since the transconductance advantage of the self-aligned device is reduced when operation is close to threshold voltage.

Clearly, the only means of obtaining significant performance improvement with this device is to reduce the gate length. The main advantage of this approach is the reduction of input capacitance, but significant improvements in transconductance and noise will be achieved in this way. In the sub-micron gate length regime, effects of velocity overshoot may become important. Therefore, one of the primary objectives of any future work will be to implement this length reduction.

Another objective of future work on this device will be to investigate the anomalously high noise figure close to threshold voltage. This has important implications for high sensitivity optical receiver applications, and must be resolved. The source of this excess noise is uncertain, but it can be shown not to be associated with high gate leakage current or 1/f type noise.

The first steps towards monolithic integration of FET and photodiode have been taken in this work, but further work is required in order to produce a fully functional PIN-FET combination. Although the JFET has been fabricated with success, the PIN-photodiode is far from optimum, and needs more development. In particular, the contacts to this device need to be improved, and some material development may be necessary to produce uncompensated layers to facilitate this. Also, considerable material and process development may be required to produce devices with low values of dark current.

To conclude, this work has so far produced an InP-based FET with high performance which has been shown to be suitable for monolithic integration. However, further work will be required to resolve some anomalies in its behaviour, and to fully realise a monolithic, high performance PIN-FET optical receiver.
REFERENCES


**Acknowledgements**

The work for this thesis was carried out as a collaborative degree, under the auspices of the University of Surrey and British Telecom Research Laboratories. I am indebted to my supervisors, Dr. K. Foulds (University of Surrey), and Dr. S. Ritchie (BTRL) for their advice and encouragement, especially for their helpful comments during the preparation of this thesis. Thanks are also due to Dr. I. Henning, Dr. A. Livingstone, and to Dr. A. Steventon, who supervised this work at BTRL at various times. Acknowledgement is made to the Director of Research, British Telecom for permission to publish the work contained in this thesis.

I am also indebted to many colleagues at BTRL for their assistance during the course of this work. In particular, I would like to thank Mr. S. Judge, and Mr. M. Learmouth for their help with device fabrication, and also Dr. G. Scott, and Mr. D. Andrews for their close co-operation in the supply of epitaxial material (MBE). Thanks are due to colleagues for various measurements; to Dr. G. Spiller for SIMS, to Mr. C. Allen for Hall mobility, to Dr. L. Westbrook for quantum efficiency, and to Mr. P. Smyth for his assistance with S-parameter measurements. I am especially grateful to Dr. G. Scott and Mr. P. Smyth for their constructive comments during the preparation of this thesis. I would also like to thank Dr. M. O'Mahony for helpful discussions on circuit modelling.

Finally, I must express my gratitude to my wife for her support and encouragement during the preparation of this thesis.
Appendix 1  Device model

The device model used for preliminary device design was based on the work of Pucel, Haus, and Statz [56]. The full mathematical treatment of this treatise will not be replicated here, but for clarity, an explanation of the underlying principles, and the essential expressions used in this work are given here.

I  Basic model

The FET channel is divided into two regions. In region I, of length L1, the electric field is below the level required for velocity saturation, and consequently Ohm’s law is obeyed. In region II, of length L2, the field is greater than the threshold field for velocity saturation (E_s), and therefore electrons travel at constant velocity, independent of the field.

For simplicity, the following normalised potentials are used:

\[ s = \left( \frac{V_s}{V_{po}} \right)^{1/2} \]  \hspace{0.5cm} \text{..... A1.1} \\

\[ p = \left( \frac{V_p}{V_{po}} \right)^{1/2} \]  \hspace{0.5cm} \text{..... A1.2} \\

where \( V_s \) is the channel potential at the source, \\
\( V_p \) is the channel potential at the joining point of the two regions, and \( V_{po} \) is the pinch-off potential.

In region I, the channel potential, \( V(x) \), is obtained by integrating the one-dimensional Poisson equation in the depletion region, assuming that the edge of the depletion region is abrupt, and that the longitudinal electric field is negligible compared to the transverse field. This gives,

\[ V(x) = V_{po} \cdot \frac{h^2}{a^2} \]  \hspace{0.5cm} \text{..... A1.3} \\

where \( h \) is the depletion depth at \( x \), and \( a \) is the total channel thickness.
Ohms law is used to obtain the drain current, $I_d$,

$$I_d = q \cdot \mu \cdot Z \cdot N_d \cdot (a - h) \cdot \frac{dV}{dx} \quad \text{...... A1.4}$$

The differential $dV$ is obtained from equation A1.3,

$$dV = 2 \cdot V_{po} h \cdot dh \quad \text{...... A1.5}$$

Substituting A1.5 into A1.4, and integrating from $x = 0$ to $x = L_1$ gives, (using normalised potentials),

$$I_d = \frac{a \cdot q \cdot \mu \cdot N_d \cdot Z \cdot V_{po} \cdot f_1}{2 \cdot L_1} \quad \text{...... A1.6}$$

where the function $f_1$ is given by,

$$f_1 = p^2 - s^2 - \frac{2}{3} (p^3 - s^3) \quad \text{...... A1.7}$$

$L_1$ can be determined from current continuity between regions I and II. In region II carriers travel at their saturation velocity and therefore,

$$I_d = q \cdot N_d \cdot Z \cdot u_s \cdot (1 - p) \quad \text{...... A1.8}$$

Equating A1.6 and A1.8 gives

$$L_1 = \frac{V_{po} \cdot f_1}{E_s \cdot (1 - p)} \quad \text{...... A1.9}$$

For a given drain current, the potential drop across the channel can be obtained by integrating the longitudinal electric field. In region I this gives,

$$V \text{ (region I)} = V_{po} \cdot (p^2 - s^2) \quad \text{...... A1.10}$$
In region II, the longitudinal electric field becomes significant, therefore the potential distribution must be determined by solving the two-dimensional Poisson's equation in the space charge region. The shape of this region is approximated by a rectangle, and appropriate boundary conditions are imposed. A solution can be found by a superposition of Poisson's and Laplace's equations in this region. The solution of Laplace's equation with the appropriate boundary conditions is given by a Fourier series, the lowest space harmonic of which is given by the following functional form,

\[
V_{\text{region II}} = \frac{2}{\pi} a E_s \cdot \cos \left( \frac{\pi}{2} \cdot \frac{y}{a} \right) \cdot \sinh \left( \frac{\pi}{2} \cdot \frac{x - L_1}{a} \right)
\]

...... A1.11

which further reduces to (at \( y = a \) and \( x = L \)),

\[
V_{\text{region II}} = \frac{2}{\pi} a E_s \cdot \sinh \left( \frac{\pi}{2} \cdot \frac{L_2}{a} \right)
\]

...... A1.12

Adding this to the potential drop in region I gives the source-drain potential,

\[
V_{sd} = V_{po} \cdot (p^2 - s^2) + \frac{2}{\pi} a E_s \cdot \sinh \left( \frac{\pi}{2} \cdot \frac{L_2}{a} \right)
\]

...... A1.13

If this equation is combined with equation A1.9, then the unknown \( L_1 \) can be eliminated, which allows \( p \) to be found for a given set of bias conditions. The drain current can then be determined using equation A1.8. Solution of these non-linear equations is performed using a desk-top computer.
II Small signal parameters

Expressions for small signal parameters can also be obtained. Transconductance is defined as \( g_m = dI_d/dV_{gs} \) at a fixed drain bias. Using equation A1.8,

\[
dI_d/dV_{gs} = dI_d/dp \cdot dp/dV_{gs} = q \cdot N_d \cdot Z \cdot u_s \cdot dp/dV_{gs} \quad \text{...... A1.14}
\]

Therefore,

\[
g_m = q \cdot N_d \cdot Z \cdot u_s/2 \cdot s \cdot V_{po} \cdot dp/ds \quad \text{...... A1.15}
\]

Differentiating equations A1.9 and A1.13, enables \( dp/ds \) to be found, which leads to the expression,

\[
g_m = q \cdot N_d \cdot Z \cdot a \cdot u_s \cdot f_g \quad \text{...... A1.16}
\]

where \( f_g \) is a function given by,

\[
f_g = \frac{(1-s) \cdot \cosh (\pi \cdot L2/2 \cdot a) - (1-p)}{[2 \cdot p \cdot (1-p) + E_s \cdot L1/V_{po}] \cdot \cosh (\pi \cdot L2/2 \cdot a) - 2 \cdot p \cdot (1-p)}
\quad \text{...... A1.17}
\]

The gate source capacitance is defined as, \( C_{gs} = dQ_g/dV_{gs} \) at a fixed drain bias, where \( Q_g \) is the free charge on the gate electrode. This charge can be found by integration of the normal field component over the gate electrode. In region I, this is the field produced by the depletion region space charge. In region II, there is an additional component produced by the Laplacian potential (equation A1.11). Integrating these field components over the gate electrode gives,
\[ Q_g = q N_d a Z \left( \frac{(f2/f1) L1 + p L2 + E_s a^2 (\cosh \frac{\pi L2}{2a} - 1)}{\pi V_{po}} \right) \] ...... A1.18

where the function \( f2 \) is given by,

\[ f2 = \frac{2}{3} (p^3 - s^3) - \frac{1}{2} (p^4 - s^4) \] ...... A1.19

Differentiation of this charge with respect to \( V_{gs} \) gives,

\[ C_{gs} = \varepsilon_s Z f_c \] ...... A1.20

where the function \( f_c = f_{c1} + f_{c2} \), with,

\[ f_{c1} = \frac{2 L1}{f1 a} \left[ f_g \left( \frac{2 p^2 (1 - p)^2 + f2}{1 - p} \right) - s (1 - s) \right] \] ...... A1.21

\[ f_{c2} = 2 L2 f_g/a + (1 - 2 p f_g) \left[ \frac{2 p V_{po}}{a E_s \cosh (\pi L2/2 a)} + \tanh (\pi L2/2 a) \right] \] ...... A1.22

In this expression, \( f_{c1} \) and \( f_{c2} \) represent the contributions of regions I and II respectively. Fringing capacitance has been neglected in this analysis, and must be added. This becomes more important for very short gate length devices.

III Noise analysis

The intrinsic noise in a FET channel arises from two basic mechanisms. In region I, in which Ohms law is applicable, thermal noise is produced, and the treatment is based on the work of van der Ziel \(^{[84,85]} \). The effects of a field dependent electron temperature are included, which is based on the work of Baechtold \(^{[86]} \).

In region II, Ohms law no longer holds, and therefore thermal noise is absent. The noise source in this region is attributed to random charge displacements which result in dipole formation, and is known as diffusion noise. Thermal equilibrium statistics are assumed to
apply, since the velocity saturation processes produce random electron velocity, which maintains a spherical velocity distribution.

The mean square fluctuation in drain current is given by,

\[ \overline{i_d^2} = 4kT \Delta f g_m P \]  \hspace{1cm} \text{A1.23}

where \( \Delta f \) is the bandwidth, and \( P \) is a noise coefficient composed of a contribution from region I (\( P_1 \)) and from region II (\( P_2 \)) such that \( P = P_1 + P_2 \), where,

\[ P_1 = \frac{(1-p)(P_0 + P_d)}{f_1 f_g \gamma^2} \]  \hspace{1cm} \text{A1.24}

\[ P_2 = \frac{(1-p)V_{po} f_3}{E_s f_r^2 f_g} \]  \hspace{1cm} \text{A1.25}

\( P_0 \) represents the contribution from simple thermal noise considerations, while \( P_d \) represents the hot electron contribution. These terms are given by,

\[ P_0 = [(p^2 - s^2)^{-4/3}(p^3 - s^3) + 1/2(p^4 - s^4)]/f_1 \]  \hspace{1cm} \text{A1.26}

\[ P_d = [(s - p) + \ln \frac{1-s}{1-p} \cdot 2 \delta (1-p)^3]/f_1 \]  \hspace{1cm} \text{A1.27}

\( \delta \) is an empirical constant based on experimental results. \( f_r, \gamma \), and \( f_3 \) are given by the following expressions,

\[ f_r = \frac{[2p(1-p) + E_s L_1/V_{po}] \cosh \pi L_2/2a - 2p(1-p)]}{(1-p)} \]  \hspace{1cm} \text{A1.28}

\[ \gamma = \frac{(1-p)^2 f_r}{f_1 \cosh (\pi/2a) L_2} \]  \hspace{1cm} \text{A1.29}
\[
\frac{f^3}{\pi^3 D_0} \left[ \sin \left( \frac{\pi}{2} \right) \left( 1 - p \right) \right]^2 \cdot \left( \frac{\exp \left( \frac{\pi L_2}{a} - 4 \exp \frac{\pi L_2}{a} + 3 + \frac{\pi L_2}{a} \right)}{2a} \right)
\]

\[
D = \mu_0 \cdot k T / q
\]

In the latter expression, \( D \) is the high field diffusion constant, and \( D_0 \) is the low field diffusion constant obtained from Einstein's relation,

\[
D_0 = \mu_0 \cdot k T / q
\]

The fluctuations in drain current induce noise currents in the gate circuit due to capacitative coupling. The mean square gate noise current is given by,

\[
\overline{i_g^2} = (4 \cdot k T \cdot \Delta f \cdot \omega^2 \cdot C_{gs}^2 / g_m) \cdot R
\]

where \( R \) is a noise coefficient composed of a contribution from region I (\( R_1 \)) and from region II (\( R_2 \)) such that \( R = R_1 + R_2 \), where,

\[
R_1 = \frac{4 \cdot V_{po}^2 \cdot f_1^3 \cdot f_g \cdot (R_0 + R_d)}{(a \cdot E_s \cdot \gamma f_c)^2 \cdot (1 - p)^3}
\]

\[
R_2 = \frac{4 \cdot V_{po}^3 \cdot f_1^2 \cdot f_g \cdot k^2 \cdot f_3}{[a \cdot E_s \cdot (1 - p)]^3 f_r^2 f_c^2}
\]

\( R_0 \) represents the contribution from simple thermal noise considerations, while \( R_d \) represents the hot electron contribution. These terms are given by,

\[
R_0 = \left\{ k^2 \cdot (p^2 - s^2) - \frac{4}{3} k \cdot (k + \gamma) \cdot (p^3 - s^3) + \frac{1}{2} (k^2 + 4k \cdot \gamma + \gamma^2) \cdot (p^4 - s^4) - \frac{4}{5} (k \cdot \gamma + \gamma^2) \cdot (p^5 - s^5) + \gamma^2 / \beta (p^6 - s^6) \right\} / f_1^3
\]
\[ R_d = [-2(k - y)^2 \cdot (p - s + \ln 1 - p/1 - s) + (2k y - y^2) \cdot (p^2 - s^2) \]
\[-2/3 y^2 (p^3 - s^3) \cdot \delta (1 - p)^3 / f 1^3 \]

\[ k = [- (p^3 - s^3) / 3 + (p^4 - s^4) / 6 + (s^2 - 2/3 s^3) \cdot (p - s)] / f 1 \]
\[ + L_2 / L_1 (1 - p) \]

The parameter \( k \) is given by,

Since the original noise voltages in the channel are responsible for both the drain noise current and the gate noise current, then some degree of correlation must exist. The correlation coefficient, \( C \), is defined as,

\[ jC = \frac{i_g * i_d}{(i_g^2 i_d^2)^{1/2}} \]

and is given by,

\[ C = \frac{S_o + S_d}{[R_o + R_d \cdot (P_o + P_d)]^{1/2}} \cdot (P_{1,R1} / P_{R1})^{1/2} + (P_{2,R2} / P_{R2})^{1/2} \]

Again, the subscripts \( o \) and \( d \) refer to the contribution from simple thermal noise considerations and from the hot electron contribution respectively. \( S_o \) and \( S_d \) are given by,

\[ S_o = [k [(p^2 - s^2) - 4/3 (p^3 - s^3) + 1/2 (p^4 - s^4)] \]
\[ + y [-2/3 (p^3 - s^3) + (p^4 - s^4) - 2/5 (p^5 - s^5)] / f 1^2 \]

\[ \]
\[ S_d = \left[ \left( k - \gamma \right) \left( s - p + \ln \frac{1 - s}{1 - p} \right) + \gamma \frac{p^2 - s^2}{2} \right] \]
\[ \times 2 \delta (1 - p)^3 / f 1^2 \]

The coefficients \( P, R, \) and \( C, \) produced above are used for obtaining an overall noise parameter for the FET. This parameter, \( \Gamma, \) is defined such that the total intrinsic mean square noise current at the FET output is given by,

\[ \overline{i_{\text{c}}^2} = 4 k T \Delta f g_m \Gamma \]

In terms of the noise parameters \( P, R, \) and \( C, \) produced above, \( \Gamma \) is given by,

\[ \Gamma = P + R \left( \frac{C_{gs}}{C_t} \right)^2 - 2 C \left( PR \right)^{1/2} \left( \frac{C_{gs}}{C_t} \right) \]

where \( C_t \) is the total input capacitance, including the PIN capacitance and stray capacitance.
Appendix 2 Derivation of figure of merit for PIN-FET design

In appendix 1, an expression was derived for the intrinsic FET noise coefficient, \( \Gamma \), which is used to obtain the total intrinsic mean square noise current at the FET output, i.e.,

\[
i_c^2 = 4kT \Delta f g_m \Gamma \quad \text{...... A2.1}
\]

To derive a figure of merit for the PIN-FET based on maximising receiver sensitivity, the total noise current must be compared with the signal current (the PIN photocurrent) at the FET input. To transform the noise current at the FET output into a noise voltage at the FET input, it must be divided by the transconductance, i.e.,

\[
e_n^2 = \frac{4kT \Delta f \Gamma}{g_m} \quad \text{...... A2.2}
\]

Extrinsic noise sources in a FET arise from parasitic resistances, and from gate leakage current. The parasitic resistances (source-channel resistance, \( R_s \), and gate metallisation resistance, \( R_m \)) give rise to a thermal noise voltage, given by,

\[
e_p^2 = 4kT \Delta f ( R_m + R_s ) \left( \frac{ ( C_s + C_p )}{C_t} \right) \quad \text{...... A2.3}
\]

where \( C_s \) is the stray capacitance, and \( C_p \) is the PIN capacitance. The total mean square FET noise voltage at the input is the sum of these components, i.e.,

\[
e_n^2 = \frac{4kT \Delta f}{g_m} \left( g_m ( R_m + R_s ) \left( \frac{ ( C_s + C_p )}{C_t} \right) + \Gamma \right) \quad \text{...... A2.4}
\]

Let

\[
\psi = g_m ( R_m + R_s ) \left( \frac{ ( C_s + C_p )}{C_t} \right) + \Gamma \quad \text{...... A2.5}
\]
then,

\[ \overline{e_t^2} = \frac{4kT \Delta f}{g_m} \psi \]  \hspace{1cm} \ldots \ldots \text{A2.6}

To clarify the relevance of these noise coefficients, \( \Gamma \) is a measure of the intrinsic FET noise, and \( \psi \) is a measure of the extrinsic FET noise, referred to the input. To convert this mean square noise voltage into an equivalent mean square noise current, it must be divided by the square of the FET input impedance, i.e.,

\[ \overline{i_f^2} = \frac{4kT \Delta f}{g_m} \psi \left( \omega C_t \right)^2 \]  \hspace{1cm} \ldots \ldots \text{A2.7}

Obviously, low noise operation is desirable, and therefore a simplified figure of merit for the FET, \( F \), is derived from the inverse of this expression,

\[ F = \frac{g_m}{\left( C_t \psi \right)^2} \]  \hspace{1cm} \ldots \ldots \text{A2.8}

Shot noise from the FET gate leakage current and thermal noise from the FET load resistor also contribute to the total noise, and should be included for a more complete approach. However, in a well designed receiver, the load resistor noise will be negligible. Although gate leakage current can be high in these devices, and can therefore contribute significantly to the total noise, the level of leakage current is not an easily controllable parameter. Consequently, it was not included in this simplified analysis.
Appendix 3. Papers published as part of this work.


Acknowledgement is made to the following, for permission to reproduce these papers here:

Institute of Electrical and Electronic Engineers, 345 East 47th St., New York, U.S.A., for papers 1 and 3,

American Vacuum Society, 335 East 45th St., New York, U.S.A., for paper 2, and

Institute of Electrical Engineers, 8 Southgate House, Stevenage, UK, for paper 4.
A Self-Aligned In$_{0.53}$Ga$_{0.47}$As Junction Field-Effect Transistor Grown by Molecular Beam Epitaxy

D. WAKE, A. W. LIVINGSTONE, D. A. ANDREWS, AND G. J. DAVIES

Abstract—An In$_{0.53}$Ga$_{0.47}$As field-effect transistor has been fabricated on MBE-grown material, using a novel self-alignment technique. This device has a dc transconductance of 60 mS/mm for a 3-μm gate length, one of the highest reported figures for such a length, and a very low gate leakage of 100 nA at -3-V gate bias.

In$_{0.53}$Ga$_{0.47}$As is a very useful semiconductor for long-wavelength optical communication systems. This composition is lattice-matched to InP, and has sufficiently narrow band gap (0.75 eV) to be used as a photodetector for systems operating up to a wavelength of 1.67 μm. Very high electron mobility is predicted in this material, and a value of 9,000 cm$^2$/V·s has been reported [1] for an LPE-grown layer with a carrier concentration of $10^{17}$ cm$^{-3}$. In addition, the low effective mass, 0.041$m_e$ [2] and the large $T_D$ conduction valley gap (0.55 eV) [3] give In$_{0.53}$Ga$_{0.47}$As a high peak electron velocity (2.9 $10^7$ cm/s has been reported [4]). These electronic properties give In$_{0.53}$Ga$_{0.47}$As distinct advantages over GaAs and InP for high-performance high-speed electronics devices. We have fabricated field-effect transistors using In$_{0.53}$Ga$_{0.47}$As primarily intended for connection to In$_{0.53}$Ga$_{0.47}$As p-i-n photodiodes to form p-i-n FET optical receivers [5], as a first step towards the monolithic integration of the two devices. For low-noise operation, these devices must have low source contact-to-channel resistance. In order to minimize this resistance, a novel self-aligning source/drain contact method has been used to form submicrometer source-gate spacings. This has the added advantage of increased yield, since there is no critical alignment step.

The Schottky gate technology used for GaAs transistors is unsuitable for In$_{0.53}$Ga$_{0.47}$As because of the low Schottky barrier height formed by metals on this semiconductor (0.2 eV) [6]. Various approaches to this problem have been tried by other workers. A thin layer of an insulator [7] or larger band-gap semiconductor [8] have been placed between the metal and In$_{0.53}$Ga$_{0.47}$As layer in an effort to increase the apparent barrier height. These approaches have been reasonably successful, but have not given very low leakage gates. which is also important for low-noise operation. With the device reported here, this problem is overcome by using the large potential barrier formed by the p-n junction of a junction field-effect transistor (JFET).

A cross section of this device is shown in Fig. 1. The material growth was carried out in a Vacuum Generators MB298...
system which has been described previously [9]. The semi-insulating Fe-doped (100) InP substrates were prepared by a novel technique. After solvent degreasing, they were etched in a 1 percent Br2/methanol solution to remove surface damage. They were then ozone cleaned in a discharge from a UV lamp. This has been shown to remove very effectively carbonaceous deposits, and leave a thin well-characterized oxide layer [10]. After loading, the substrates were outgassed in the preparation chamber at 600 K until the pressure fell to less than $10^{-9}$ torr. The substrates were then transferred to the growth chamber, where they were heated to 790 K in an arsenic flux to desorb the surface oxide layer [11]. This provided an As-stabilized surface, reconstructed to give a streaked (2 X 4) RHEED pattern. This in situ cleaning left the InP surface free of C and O contamination whilst maintaining the substrates chemical integrity. The FET structure consists of 0.3 μm of n In0.53Ga0.47As (Si-doped at $5 \times 10^{18}$ cm$^{-3}$) and 0.3 μm of p+ In0.53Ga0.47As (Be-doped at $10^{20}$ cm$^{-3}$). The substrate temperature was kept at 790 K for the initiation of growth and then raised to 850 K for the remainder of the growth sequence.

The first step in the fabrication of this device is to define the gate metal (Ti/Au) photolithographically. The p⁺ In0.53Ga0.47As layer is then completely etched using the gate metal as a mask, so that a rib of p⁺ material is formed. This is achieved by successively etching in small steps and monitoring the current flowing between two adjacent contacts on the slice until a back-to-back p-n junction characteristic is found. An etchant with a moderately slow etch rate is used to make this process controllable. (5 H2O2 : 1 H2SO4 : 50 H2O kept at 20°C has an etch rate of 0.65 μm/min. If the gates are aligned along the [110] direction then the profile shown in Fig. 1 is obtained. A second photolith step defines the mesas area, and the structure is etched until the mesa is formed, with then etched further until the link between the gate bondpad and the gate is fully undercut. The air bridge formed (shown in Fig. 2) isolates the gate bondpad mesa from the FET mesa, ensuring that the high bondpad mesa capacitance and leakage current do not affect device performance. This method ensures a planar gate metallization, overcoming the disadvantages of running a gate metal track down a mesa wall, i.e., the possibility of discontinuity, or shorting of the p+n junction. The next step is to evaporate source/drain metal (Ti/Au). The overhanging gate metal acts as a shadow mask, and a self-aligned structure is formed, By using this technique, submicrometer gate-source and gate-drain spacings are achieved (without any critical photolith step) and gate-metal thickness is increased thereby reducing parasitic resistances. Also, since the p⁺ rib has angled walls, the actual gate dimension can be much less than the gate-metal length.

Fig. 3 shows the output characteristic of a typical device fabricated in this way, with a gate width of 280 μm. This device has a dc transconductance of 60 mS/mm at $V_{ds} = 2$ V and $V_{ds} = 0$ V, for a gate length of 3 μm (implying a mobility of 6000 cm²/V·s), and the gate-source capacitance near pinchoff is 0.4 pF. As expected, the parasitic resistances of this device are low, i.e., the channel-to-source resistance is 6 Ω (without alloyed contacts), and the gate metallization resistance is 4.5 Ω. On a similar slice, values as low as 3.7 Ω have been measured for these resistances. Also, gate leakage is very low in this device (100 nA at -3 V). These values make this device attractive for low-noise applications, particularly for the intended monolithic integration with a p-n photodiode, to form a high sensitivity optical receiver.
In summary, this FET design requires no critical processing, giving high yield, and low parasitic resistances because of sub-micrometer source-gate spacing and thick gate metal. This design should be well suited to achieving sub-micrometer gate-length devices, in order to further improve the trans-conductance and gate-source capacitance.

ACKNOWLEDGMENT

The authors would like to thank S. N. Judge and M. D. Learmouth for help with device processing, E. G. Scott for assistance with growing the material, and the Director of Research, British Telecommunications, for permission to publish this letter.

REFERENCES


Factors affecting the growth of an integrated Ga\textsubscript{1-x}In\textsubscript{x}As/InP PIN-FET by molecular beam epitaxy\textsuperscript{a}

E. G. Scott, D. Wake, A. W. Livingstone, D. A. Andrews, and G. J. Davies

British Telecom Research Laboratories, Martlesham Heath, Ipswich, IP5 2BE, United Kingdom

(Received 27 January 1985; accepted 28 January 1985)

A combination of electrical and optical properties makes GaInAs lattice matched to InP a suitable material for the fabrication of an integrated PIN-FET receiver for use in telecommunications applications at 1.55 $\mu$m. As steps towards integration, the individual devices in such a receiver have been fabricated using molecular beam epitaxy (MBE) material. With the use of careful substrate cleaning it has been demonstrated that significant improvements in the epitaxial quality of GaInAs layers can be produced. JFETs have been fabricated from this material and using a novel self-aligning source/drain contact method, source contact to channel resistance has been minimized by the formation of submicron source to gate spacings. These devices had a dc transconductance of 60 mS/mm for a 3 $\mu$m gate length. Similarly PIN diodes have been fabricated. The high field region of the device contains low doped GaInAs, Nd:Na approximately $5 \times 10^{17}$ $\text{cm}^{-3}$ with corresponding $\mu_{n0} > 9000$ $\text{cm}^{2} \text{V}^{-1} \text{s}^{-1}$. For monolithic integration, the two components require an intervening isolation region. The choice of isolation material as well as the optimization of growth conditions for the integrated structure together with preliminary device results are discussed.

I. INTRODUCTION

Long wavelength optical communications have been made possible by the recent development of very low-loss, low-dispersion silica fibers. These fibers have loss minima in the 1.3-1.55 $\mu$m wavelength range and this has initiated considerable effort to develop high-sensitivity detectors for use in this spectral region. Ga\textsubscript{1-x}In\textsubscript{x}As, lattice matched to InP has a band gap (0.7 eV) suitable for use out to 1.67 $\mu$m. In addition, very high electron mobility is predicted in this material, and values in excess of 10 000 $\text{cm}^{2} \text{V}^{-1} \text{s}^{-1}$ have been reported.\textsuperscript{1} Also, the low effective mass (0.04$m_{e}$) together with the larger $\Gamma - L$ conduction valley gap (0.55 eV) give Ga\textsubscript{1-x}In\textsubscript{x}As a high peak electron velocity—a value of 2.9 $\times 10^{7}$ $\text{cm} \text{s}^{-1}$ has been reported.\textsuperscript{2} The combination of electrical and optical properties make Ga\textsubscript{1-x}In\textsubscript{x}As, lattice matched to InP, a most suitable material for the fabrication of receivers for use in the 1.55 $\mu$m wavelength region.

As far as specific devices are concerned Ga\textsubscript{1-x}In\textsubscript{x}As PIN-FET detectors offer advantages over avalanche photodiodes (APDs) and hybrid devices, including lower operating voltages, lower noise levels, reduced capacitance and parasitic resistances, and a less stringent specification.\textsuperscript{3} This paper reports the MBE conditions required to meet the specifications of Ga\textsubscript{1-x}In\textsubscript{x}As FET and PIN devices and their subsequent integration.

II. EXPERIMENTAL

The MBE material growth was carried out in a Vacuum Generators MB288 system which has been described previously.\textsuperscript{4} The semi-insulating Fe-doped (100)InP substrates were prepared by either of two methods—ozone cleaning or by conventional etches as described in Ref. 4 but with careful attention to detail. In the former, after solvent degreasing the substrates were etched in 1%Br\textsubscript{2}/methanol solution to remove surface damage. They were then ozone cleaned in a discharge from a UV lamp under which a trickle of oxygen was allowed to flow. This has been shown to remove carbonaceous deposits very effectively, and leave a thin, well-characterized oxide layer.\textsuperscript{5} After washing in dilute HCl and deionized water the substrates were loaded into the MBE equipment and outgassed in the preparation chamber at 600 K. The substrates were then transferred to the growth chamber, where they were heated to 790 K in an arsenic flux to desorb the surface oxide layer.\textsuperscript{6} This provided an arsenic stabilized surface, reconstructed to give a streaked (2x4) RHEED pattern. p- and n-type doping was achieved using Be- and Si-Knudsen cells, respectively. The substrate temperature was kept at 790 K for the initiation of growth and then raised to 850 K for the remainder of the growth sequence. Ga\textsubscript{1-x}In\textsubscript{x}As growth was performed using separate Ga and In crucibles and without substrate rotation.

![Electron Mobility vs Composition](image)

FIG. 1. Effect of In composition $x$ on the electron mobility of Ga\textsubscript{1-x}In\textsubscript{x}As as epitaxial layers.
III. Ga$_{1-x}$In$_x$As FET

In the Ga$_{1-x}$In$_x$As FET the active region, $N_D - N_A = 5 \times 10^{16}$ cm$^{-3}$, can be situated adjacent either to the InP semi-insulating substrate or to a lattice matched, semi-insulating Al$_{1-x}$In$_x$As layer. As will be seen in a later section this latter option, though attractive from a device point of view, has proven to be less than satisfactory in practice. In this section only the direct growth of the JFET structure on (100) Fe doped InP is considered.

As the source to substrate distance is large $>$ 18 cm compositional uniformity across the wafer is good. In = 52 ± 3%, over a 2 in. wafer. However, a grading still exists and the effect of In composition on measured Sheet Hall mobility is shown in Fig. 1 for a layer doped $N_D - N_A = 1 \times 10^{16}$ cm$^{-3}$. For $x < 0.50$ the mobility degrades sharply with composition as does the morphology. However, for $x > 0.53$ the mobility plateaus over the region are measured and there is also an absence of gross topographical features. This is in accord with the earlier work of Whitely and Ghandi$^7$ on Ga$_{1-x}$In$_x$As growth by MOCVD, who showed the greater tolerance of the epilayer to compression than to tension.$^{12}$ Thus only that part of the layer with $x > 0.52$ should be processed for FET technology.

Similarly because the active device forms part of a heterojunction the integrity of the junction has a profound effect on the electrical characteristics of the active layer. This is particularly so for a thin device layer adjacent to the substrate interface. It has been shown by Whitely$^7$ that the mobility of a Ga$_{1-x}$In$_x$As layer grown by MOCVD decreases dramatically as one approaches the InP interface and it is not until one is $>1 \mu$m from the interface that bulk-like mobilities are obtained. We have observed a similar situation in MBE grown Ga$_{1-x}$In$_x$As but with a much steeper decrease (Fig. 2) so that bulk like mobilities are obtained after 0.25 $\mu$m. It was believed that impurities gathered at the interface generated dislocations which then propagated through the layer and that the high dislocation density contributed to the poor mobility close to the interface. This was especially severe in the 0.3 $\mu$m active region required for the JFET as shown in Fig. 3 in which 5 X 10$^7$ defects cm$^{-2}$ were detected. The micrograph also shows a layer close to the interface from which the defects arise and which also has a "strain" region associated with it. It was believed that the defects originated at impurities at the interface and a TEM plane view of the specimen adds support to that argument (Fig. 4). The defects are shown to exist preferentially in linear arrays roughly orientated along the (110) direction. There are no arrays running orthogonal to these so misfit arrays are unlikely. The arrays originate at the interface but their true

![Fig. 1. TEM cross sectional micrograph across the Ga$_{1-x}$In$_x$As/InP substrate interface showing defects emanating from the interface as well as a strained layer at the interface.](image1)

![Fig. 2. Variation of electron mobility with distance from the Ga$_{1-x}$In$_x$As/InP substrate interface. () continuous profile sheet measurement = 493 cm$^2$/V·s$^{-1}$.](image2)

![Fig. 3. TEM plan view of a Ga$_{1-x}$In$_x$As layer close to the InP substrate. The linear arrays of defects are roughly aligned along the (110) direction.](image3)

![Fig. 4. TEM plan view of a Ga$_{1-x}$In$_x$As layer close to the InP substrate. The linear arrays of defects are roughly aligned along the (110) direction.](image4)

![Fig. 5. TEM cross sectional micrograph across the Ga$_{1-x}$In$_x$As/InP substrate interface after ozone cleaning.](image5)

nature is still unknown at this instant. However, it is believed that they are impurity based or else gettered impurities. On this basis careful attention to substrate preparation together with ozone cleaning was maintained as described in an earlier section. This resulted in layers free from dislocations as shown in Fig. 5 and with x-ray double crystal rocking curves with extremely narrow halfwidths (Fig. 6) indicating a high crystalline quality in the epitaxial layer. However, careful examination of the TEM micrograph in Fig. 5 indicates that there still exists a region of strain at the InP/Ga1-xInxAs interface. This is believed to be caused by the transients in flux produced when first opening the Group III cell shutters, forming a thin mismatched layer close to the interface. This mismatched layer probably also contributed to the poor mobilities found close to the interface (see Fig. 2) as the “average” sheet mobility increased from \(-3300\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}\) to \(>6000\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}\) in the active region when the improved cleaning and etching techniques were used. However, this is still below the \(9000\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}\) mobility found in layers \(\sim 1\ \mu\text{m}\) thickness.

In the fabrication of the JFET structures the gate metallization (Ti/Au) was defined photolithographically and the \(p^+\) Ga1-xInxAs layer completely etched using the gate metal as a mask so that a rib of \(p^+\) material is formed.12 A second photolith step defines the mesa area and the structure was etched until the mesa formed. Further etching resulted in the line between the gate bond pad and the gate being fully undercut. The airbridge formed (shown in Fig. 7), isolates the gate bond pad mesa from the FET mesa, ensuring that the high bond pad mesa capacitance and leakage current do not affect device performance. This technique ensures a planar gate metallization, overcoming the disadvantage of a gate metal track running down the mesa wall. Further Ti/Au metal evaporation over the overhanging gate metal as a self-aligning metal mask formed the source and drain contacts. In this way submicron gate-source and gate-drain spacings are produced and gate metal thickness are increased thereby reducing parasitic resistances. Devices produced in this way had dc transconductances of 60 ms/mm at \(V_D = 2\ \text{V}\) and \(V_{GS} = 0\ \text{V}\), for a gate length of 3 \(\mu\text{m}\), and a gate-source capacitance near pinch off of 0.4 \(p\text{F}\). Parasitic resistances are low, \(-\frac{\Omega}{}\) for gate metallization while the gate leakage is acceptably low (100 \(n\text{A}\) at \(-3\ \text{V}\)), which makes the device very attractive for integration in a high sensitivity receiver.

**IV. Ga1-xInx As PIN DIODES**

Basic materials requirements for III-V PIN diodes are layers > 3 \(\mu\text{m}\) in thickness with \(N_D - N_A < 5 \times 10^{16} \text{cm}^{-3}\). Background doping levels were typically, \(N_P - N_A = 5 - 10 \times 10^{16} \text{cm}^{-3}\) which are slightly too high to make effective PIN diodes. The residual \(n\)-type dopant is thought to be Si, an impurity in the indium source. Cheng et al.9 have reported that for optimized growth of Ga1-xInxAs, high substrate temperatures and high arsenic overpressures need to be maintained. However, Davies et al.10 have shown that for Al1-xInx As, these conditions can result in material with high residual doping (\(N_P - N_A > 10^7 \text{cm}^{-3}\)) which was shown by SIMS to be doped with one of the chalcogens, S, Se, or Te depending on the As source material. This may again be a problem in Ga1-xInxAs, as so attention to the purity of both source materials is under study and should lead to epitaxial material with lower residual impurities.

PIN diode mesa structures 35 × 40 \(\mu\text{m}\) were made from this material using an MOVPE overgrowth of \(p\)-type InP for the junction layer. The device had leakage currents in the range 4-50 \(n\text{A}\) at \(-10\ \text{V}\) with an associated capacitance of approximately 0.1 \(\text{pF}\) at \(-10\ \text{V}\).

**V. INTEGRATED PIN-FET**

An initial structure for an integrated PIN-FET is shown in Fig. 8. It is evident that as well as optimizing the growth conditions for the PIN and FET separately the isolating Al1-xInxAs layer also has to be characterized. We have shown recently9 that the growth temperature is critical in determining the electrical properties of Al1-xInxAs, as lattice matched to InP. In that study growth temperatures < 500 \(^\circ\text{C}\) resulted in semi-insulating Al1-xInxAs while for \(T_g > 520^\circ\text{C}\) nominally undoped layers were \(n\) type (\(N_D - N_A \sim 1 \times 10^{16} \text{cm}^{-3}\)). It was postulated in that work that the semi-insulating nature of the material arose either from oxygen incorporation into the layer or from an in-
Fig. 8. Schematic diagrams of integrated PINFET structures with isolating Al\(_{1-x}\)In\(_x\)As layer.

increased defect population caused by lack of surface mobility of atoms at the lower growth temperature. In our initial attempts at growing this integrated structure we have found that layers grown on this isolating, semi-insulating Al\(_{1-x}\)In\(_x\)As had poor crystal quality and hence poor electrical characteristics. This led us to believe that the JFET layers were defected from growth on top of an imperfect matrix.

As a result of this work other PINFET device structures are now being considered. The major difference is that the whole structure is now Ga\(_{1-x}\)In\(_x\)As and utilizes junction isolation, similar to that used in silicon bipolar technology. A complete presentation of the device characteristics of this structure is in preparation and will be presented elsewhere.

VI. CONCLUSION

In summary, it has been shown that with the use of careful substrate cleaning, significant improvements in the epitaxial quality of Ga\(_{1-x}\)In\(_x\)As layers lattice matched in InP can be produced. JFETs have been fabricated from this material and using a novel self-aligning source/drain contact method, source contact to channel resistance has been minimized by the formation of submicron source to gate spacings. These devices had transconductances of 60 ms/mm for 3 \(\mu\)m gate lengths. PIN diodes have also been prepared but their characteristics have suffered owing to the high residual impurity level, \(N_d - N_a = 5 \times 10^{16} \text{ cm}^{-3}\) in the undoped region. Similarly semi-insulating Al\(_{1-x}\)In\(_x\)As has been shown to be nonideal for the isolating region in the integrated PINFET and an alternative structure utilizing junction isolation have been designed for growth by MBE as a result of this work. Performance characterization measurements are in progress.

ACKNOWLEDGMENTS

The authors wish to thank M. Taylor and M. Hockley for the TEM studies, and R. Heckingbottom and A. Stevenson for useful discussions. Acknowledgment is made to the Director of British Telecom Research Laboratories for permission to publish this paper.

*This paper was presented at the 3rd International Conference on Molecular Beam Epitaxy, San Francisco (1984).


5. A. W. Nelson (private communication).


InGaAs/InP Junction Field-Effect Transistors with High Transconductance Made Using Metal Organic Vapor Phase Epitaxy

D. WAKE, A. W. NELSON, S. COLE, S. WONG, I. D. HENNING, AND E. G. SCOTT

Abstract—The growth and fabrication of a InGaAs/InP junction field-effect transistor using metal organic vapor phase epitaxy is reported for the first time. Very high extrinsic transconductance has been achieved (210 mS/mm for a gate length of 1.5 μm), by the use of a p-InP buffer layer which allows close to maximum electron velocity in the channel, and by using a self-alignment technique to give very low values of access resistance, typically 0.5 Ω-mm.

InGaAs, lattice-matched to InP, is an important material for optoelectronic devices. Excellent electronic properties (high mobility and high peak electron velocity), make this material suitable for high performance field-effect transistors, while the absorption edge at a wavelength of 1.67 μm means that it is an ideal candidate for photodetectors used in optical communication systems operating in the low-loss windows of silica-based fibers at wavelengths of 1.3 and 1.55 μm. Monolithic integration of photodetector and amplifying electronics is an important step towards very high sensitivity, or low-cost optical receivers. The JFET described in this letter has been designed to facilitate vertical integration with a p-i-n photodiode by incorporating a p-type buffer layer, which acts as an isolation layer between the two devices [1]. Recently, results on a JFET using a p+ InGaAs buffer layer were reported [2], but relatively low values of transconductance were obtained, (90 mS/mm). Our work on similar structures has shown that a gate length of 1.5 μm, and a 0.3-pm n-InGaAs channel, and a 0.3-pm p-InP gate contact layer, sulphur doped at 7.10×10²⁰ cm⁻³, and a 0.3-μm p-InP gate contact layer, sulphur doped at 3.10×10²⁰ cm⁻³. The structure was grown by atmospheric pressure MOVPE in a horizontal reactor arrangement, using the conventional alkyls TMIn and TMMG as the group III sources, and 5 percent AsH₃ in H₂, and 3 percent PH₃ in H₂ as the group V sources. The growth temperature was 650°C and typical growth rates were 0.13 μm/min, for InGaAs and 0.08 μm/min for InP. Other growth conditions were adjusted to give lattice-matched InGaAs having double crystal X ray rocking curve halfwidths of 50­/80 arc seconds, and van der Pauw measurements on thick (2-4 μm) undoped layers gave room temperature Hall mobility values of 9000 cm²/Vs.

Dimethyl cadmium was used as the p-type dopant source, which further aided the realization of abrupt p-n heterojunctions since DMCD is a very inefficient dopant [3], and the possibility of unintentional p doping during growth of the n-type channel from residual DMCD remaining in the reactor chamber was, therefore, substantially reduced. The n-type doping of the channel was obtained using H₂S as the dopant source. After growth, doping levels and layer thicknesses were measured using the British Telecom Electrochemical Profiler (Polaron Ltd).

Fabrication of the devices was similar to that described previously [4]. Ti/Pt/Au gate metal was defined photolithographically, which was then used as an etch mask for removal of the top layer, leaving a rib of p-InP. This was done using a material selective etchant, until the required degree of undercut was achieved, which gave control of the gate length. Source/drain metal (12 percent Ge/88 percent Au) was then evaporated directly over this undercut structure, which ensured self alignment due to the shadow effect of the overhang gate metal. The source/drain metal was then sintered at 300°C for 20 s to ensure low resistance contacts (the Pt layer in the gate metallization acted as a barrier to the diffusion of Ge or Au during this process). Source-to-channel and channel-to-drain spacings of less than 0.5 μm are easily achieved, giving values of access resistance of the order of 0.5 Ω-mm after sintering. Mesa isolation and air bridge formation (to...
isolate individual devices and to isolate FET mesa from gate bond pad mesa, respectively), complete the fabrication.

Typical device characteristics are shown in Figs. 2 and 3. The transconductance at $V_{ds} = 3.0$ V and $V_{gs} = 0$ V is 58 mS for a gate length of 1.5 µm and a gate width of 270 µm. This gives a transconductance per gate width in excess of 210 mS/mm. The gate voltage required to pinch off the channel is -1.8 V, taken from a transfer characteristic at $V_{ds} = 0.1$ V.

The saturation velocity of electrons in the channel may be calculated using the expression:

$$v_{sat} = -1/(q \cdot N \cdot Z) \cdot \left(\frac{dI_{ds}}{dH}\right)$$

where $N$ is the channel carrier concentration, $Z$ is the device width, and $H$ is the depletion layer thickness, which is calculated from the effective potential across the p-n junction at the drain end of the gate. Fig. 4 shows the variation of saturated drain current with depletion layer thickness, from which a saturation velocity of $2.7 \times 10^6$ cm/s is obtained, which is close to the maximum electron velocity in this material. This result shows that the buffer layer does not degrade the channel layer, and is successful in reducing any deleterious effects due to the substrate.

The gate-source capacitance of this device is 0.50 pF at $V_{gs} = 0$ V (measured at 1 MHz) which gives a value of $f_t = 18.5$ GHz at this bias. The gate leakage current is ~15 µA at -1.0 V, with a break down voltage of ~2.0 V measured when $I_{ds} = -100$ µA. This high level of leakage is thought to be due to a large number of interface states in the high-field region of the gate.

In summary, InGaAs/InP JFET's showing high transconductance have been produced from MOVPE grown material for the first time, with extrinsic values as high as 210 mS/mm for a gate length of 1.5 µm. This figure reflects the high saturation electron velocity attained, which is due to the abrupt nature of p-n heterojunctions in the device. This has been achieved by using cadmium as the p-type dopant, which has a very low diffusion coefficient in InP.

ACKNOWLEDGMENT

The authors would like to thank S. N. Judge for help with device processing, and the Director of Research, British Telecommunications, for permission to publish this letter.

REFERENCES

MONOLITHICALLY INTEGRATED InGaAs/InP PIN-JFET PHOTORECEIVER

Indexing terms: Optical communications, Photodetectors, Integrated optics

A vertically integrated InGaAs/InP PIN-JFET has been fabricated, in which separate layers are used for the FET channel and PIN intrinsic region. The maximum transconductance was 170 mA/mm, which is the highest reported figure for an integrated structure, and the photodiode quantum efficiency was 64% at -5 V and 1-55 μm wavelength, without antireflection-coating.

The PINFET receiver is used extensively in optical communication systems operating in the low-loss windows of silica-based fibres at wavelengths of 1.3 and 1.55 μm. InGaAs, lattice-matched to InP, is used to construct the PIN photodiode, because it absorbs radiation in both of these windows with high quantum efficiency. Generally, the PIN photodiode is coupled to a low-noise GaAs MESFET in a hybrid configuration. However, InGaAs also has favourable electronic properties for a high-performance FET, and promises to be superior to the equivalent GaAs device. Therefore, monolithic integration of a PIN photodiode and an FET using InGaAs would combine high-performance devices with a minimum of parasitic losses.

Previous attempts at an integrated structure using InGaAs have used one layer for both devices. This approach compromises the overall performance, since the FET channel must be moderately doped to give high transconductance, and the PIN photodiode intrinsic region must be undoped to allow full depletion and therefore high speed. Alternatively, monolithic PINFETs have been made by using InP MISFETs to overcome this problem. However, to date the performance of the InP MISFET has been relatively poor.

The structure described in this letter overcomes these limitations by vertically integrating the two types of device, i.e. by growing one device on top of the other, in one epitaxial process. By using this arrangement, the performance of each device can be optimised separately. The structure was grown first, and consisted of a 0.6 μm p-contact layer, a 0.3 μm n-type layer (10^16 cm^-3) channel layer, and a 0.3 μm p-type layer for the FET gate. A schematic cross-section of this structure is shown in Fig. 1.

In this design, the PIN p-contact layer is used as the FET buffer layer. This type of buffer has been shown to be effective in maintaining high electron velocity in the channel region of the FET while adding negligible excess capacitance, and we have achieved transconductance as high as 220 mA/mm by using this technique with discrete JFETs.

Discrete JFETs with an InP n-type buffer layer made using MOVPE material gave very similar results. The JFETs used in this integrated structure were formed using the undercut gate rib and self-aligned source-drain metal technique described previously. In this technique, the gate metal acts as an etch mask to fabricate a rib of n-type material, which forms the FET gate. Owing to underetching, a T-shaped structure is produced, which acts as a shadow mask when source and drain metal is evaporated. This self-alignment gives a source-gate spacing of less than 0.5 μm, and therefore very low parasitic resistance. Gate length and width were 1 μm and 200 μm, respectively. The PINs were formed by mesa-etching in a two-stage process to gain access to the n-contact layer. Devices were interconnected using the air bridge technique, formed by etching away semiconductor from underneath a metal track. This method of interconnection provides good, reliable contact without excess capacitance or short-circuiting. The separation of the PIN from its bondpad gives the opportunity of making the PIN area very small, thereby minimising capacitance and leakage. This also makes the device easier to bond to, since the bondpad can have a large area and is away from any critical part of the PIN. In this case, the PIN dimensions are 100 x 100 μm, for easy optical alignment, but in principle can be made as small as the incoming optical spot size.

Output and transfer characteristics for the FET are shown in Fig. 2. The maximum transconductance is 34 mA, or 170 mA/mm when scaled for gate width, the highest so far reported in any integrated structure. The gate-source capacitance is 0.50 pF at Vgs = 0 V, and 0.29 pF at Vgs = -2 V. Gate-source leakage is 10 mA at -2 V, which is higher than the level generally experienced with the equivalent discrete FET. This is probably due to the more involved processing in the presence of the exposed gate rib with the integrated structure. The leakage can be reduced by over an order of magnitude by dipping the slice in a very dilute etch to remove a thin layer of semiconductor between the gate and source/drain.

The PIN photodiode has a dark current of 400 nA at a bias of -5 V. The capacitance of the PIN photodiode is 0.5 pF; the total input capacitance is therefore 0.8 pF at the operating bias. The quantum efficiency of the photodiode is 64% at -5 V and a wavelength of 1-55 μm without antireflection-coating.

Reprinted from ELECTRONICS LETTERS 19th June 1986 Vol. 22 No. 13 pp. 719-721
at the FET input (lower trace) and at the output (upper trace). The low voltage gain (1.7) was due to the low output impedance of the FET (100 Ω), which gave an effective load resistance of 70 Ω.

Acknowledgments: The authors would like to acknowledge the hard work of S. N. Judge for device processing, L. D. Westbrook for the optical measurements, and G. J. Davies and D. A. Andrews for help with material development. We would also like to thank C. J. Stanghan for bonding the devices, our colleagues in the Materials Division for semiconductor analysis, and the Director of Research, British Telecom, for permission to publish this letter.

Fig. 3: Photocurrent of integrated structure

Lower trace shows input voltage caused by chopped 50 µW incident laser light at 1.53 µm; upper trace shows output voltage at the FET input (lower trace) and at the output (upper trace). The low voltage gain (1.7) was due to the low output impedance of the FET (100 Ω), which gave an effective load resistance of 70 Ω.

Acknowledgments: The authors would like to acknowledge the hard work of S. N. Judge for device processing, L. D. Westbrook for the optical measurements, and G. J. Davies and D. A. Andrews for help with material development. We would also like to thank C. J. Stanghan for bonding the devices, our colleagues in the Materials Division for semiconductor analysis, and the Director of Research, British Telecom, for permission to publish this letter.

References


