IMPACT IONIZATION AND ELECTRICAL TRANSPORT
IN MULTIPLE QUANTUM WELL STRUCTURES

BY

JEREMY ALLAM

A thesis submitted to the Faculty of Science
at the University of Surrey for the Degree
of Doctor of Philosophy.

Department of Physics,
University of Surrey,
ABSTRACT

This thesis describes experimental studies of perpendicular transport in multi-layer III-V semiconductor structures. Impact ionization, charge trapping at heterointerfaces and resonant tunnelling in multiple quantum well photodiodes were studied.

Long-wavelength high-speed multiple quantum well avalanche photodiodes were studied. Ga$_{0.47}$In$_{0.53}$As/InP structures grown by trichloride, and metal-organic, vapour phase epitaxy exhibited avalanche multiplication at room temperature. The pulse response at high fields was not limited by carrier pile-up at the heterointerfaces. High quality Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As photodiodes, grown by molecular beam epitaxy, exhibited multiplication of $\approx 70$, microwave gain of $\approx 12$ and an intrinsic response time of $\approx 100$ ps at room temperature and high electric fields, indicating no carrier pile-up despite the very abrupt heterointerfaces.

In multiple quantum well photodiodes which exhibited large reverse-bias leakage currents, a new carrier multiplication effect was observed. This was attributed to impact ionization across the heterojunction discontinuity of carriers "stored" in the quantum wells. Experimental evidence for this effect is presented for several material systems. Single carrier-type multiplication was demonstrated in a multilayer photodiode with compositionally-graded interfaces.

Sequential resonant tunnelling and resonant Zener tunnelling were studied in high quality Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As structures with low background doping. Carrier transport in undepleted quantum wells was studied in Ga$_{0.47}$In$_{0.53}$As/InP structures grown by gas source molecular beam epitaxy.
ACKNOWLEDGEMENTS

The experimental work presented in this thesis was carried out during a 21 month stay at the Solid State Electronics Research Laboratory, AT&T Bell Laboratories, Murray Hill. I would like to thank AT&T Bell Laboratories for their hospitality and financial support during my visit. I also wish to acknowledge the Science and Engineering Research Council for funding and for their flexibility in supporting this arrangement.

I particularly wish to thank my supervisor at Murray Hill, Dr. Federico Capasso, for his tireless enthusiasm and interest in the work. I acknowledge the many co-workers at Bell Laboratories who made this work possible, especially the crystal growers (mentioned in the text) for providing the high quality structures. Thanks are due to A.L Hutchinson for patient instruction in the art of device fabrication, and for processing most of the devices studied. C.G. Bethea, N.A. Olsson and J.E. Bowers are thanked for help with the high frequency measurements and providing semiconductor lasers, and J.C. Campbell for help in the calibration of a reference detector. Collaboration with Khalid Mohammed, Fabio Beltram and Susantha Sen was much appreciated. Other co-workers are acknowledged in the publications.

Prof. A.R. Adams, of the University of Surrey, is thanked for illuminating discussions and useful suggestions, as well as for reading this thesis. Discussions with Dr. E.O'Reilly have been very useful.

Finally, I would like to thank my family for encouragement and help during illness, and the Harris family for their kindness and Karen for her support whilst I was in the States.
LIST OF PUBLICATIONS.

The following papers based on the work reported in this thesis have been published in the scientific literature.

1. "New high speed long wavelength Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As multiple quantum well avalanche photodiodes", K.Mohammed, F.Capasso, J.Allam, A.Y.Cho and A.L.Hutchinson, Applied Physics Letters 47(6), 15 September 1985, pp. 597-599.


#$ CONTENTS.

Abstract ......................................................... 1
Acknowledgements ............................................. 11
List of publications .......................................... 111
Contents .......................................................... v
CHAPTER 1: INTRODUCTION ........................................ 1
  1.1 Development of low-loss optical fibres ..................... 1
  1.2 Photodetectors for optical fibre communications ........... 2
  1.3 Organization of thesis ...................................... 5
CHAPTER 2: PROPERTIES OF BULK III-V SEMICONDUCTORS .......... 7
  2.1 Band-structure of III-V semiconductors .................... 7
  2.2 Properties of a reverse-bias P-N junction .................. 12
    2.2.1 Depletion capacitance ................................ 14
    2.2.2 Dark current mechanisms .............................. 15
  2.3 Impact ionization in bulk materials ....................... 17
    2.3.1 Multiplication factors and ionization rates ........... 17
    2.3.2 Signal-to-noise ratio for an APD ..................... 21
    2.3.3 McIntyre noise theory ................................ 22
    2.3.4 Frequency response of bulk APD's .................... 23
    2.3.5 Threshold energies .................................... 25
    2.3.6 Theoretical calculation of impact ionization rates 28
    2.3.7 Survey of experimental results ....................... 31
CHAPTER 3: PROPERTIES OF QUANTUM WELL STRUCTURES ............ 34
  3.1 Heterojunctions and band discontinuities .................. 34
  3.2 Carrier transport across a heterojunction ................. 35
    3.2.1 Quantum efficiency of Separate Absorption and ......... 37
       Multiplication region APD's .............................
    3.2.2 Pulse response of SAM-APD'S .......................... 39
### CHAPTER 3: QUANTUM WELLS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3 Energy levels in quantum wells</td>
<td>40</td>
</tr>
<tr>
<td>3.4 Optical properties of quantum wells:</td>
<td>44</td>
</tr>
<tr>
<td>excitons and electroabsorption</td>
<td></td>
</tr>
<tr>
<td>3.5 Electrical properties of quantum wells:</td>
<td>46</td>
</tr>
<tr>
<td>resonant tunnelling</td>
<td></td>
</tr>
<tr>
<td>3.6 Carrier transport across a quantum well</td>
<td>50</td>
</tr>
<tr>
<td>3.7 Enhancement of ionization rate ratios in multilayer APD's</td>
<td>54</td>
</tr>
</tbody>
</table>

### CHAPTER 4: LATTICE-MATCHED III-V HETEROSTRUCTURES:

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Properties of III-V semiconductors</td>
<td>62</td>
</tr>
<tr>
<td>4.1.1 The binary III-V's</td>
<td>62</td>
</tr>
<tr>
<td>4.1.2 The ternary alloys and lattice matching</td>
<td>65</td>
</tr>
<tr>
<td>4.1.3 Quaternary alloys</td>
<td>68</td>
</tr>
<tr>
<td>4.2 Heterojunction systems</td>
<td>70</td>
</tr>
<tr>
<td>4.2.1 Measurement methods for band-offsets in GaAs/AlGaAs</td>
<td>71</td>
</tr>
<tr>
<td>4.2.2 Band offsets in Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As</td>
<td>72</td>
</tr>
<tr>
<td>4.2.3 Band offsets in Ga_{0.47}In_{0.53}As/InP</td>
<td>74</td>
</tr>
<tr>
<td>4.3 Growth techniques for heterojunctions</td>
<td>76</td>
</tr>
<tr>
<td>4.3.1 Liquid phase epitaxy</td>
<td>76</td>
</tr>
<tr>
<td>4.3.2 Vapour phase epitaxy</td>
<td>79</td>
</tr>
<tr>
<td>4.3.3 Metal organic chemical vapour deposition</td>
<td>79</td>
</tr>
<tr>
<td>4.3.4 Molecular beam epitaxy</td>
<td>81</td>
</tr>
<tr>
<td>4.3.5 Gas-source molecular beam epitaxy</td>
<td>86</td>
</tr>
<tr>
<td>4.3.6 Chemical beam epitaxy</td>
<td>87</td>
</tr>
</tbody>
</table>

### CHAPTER 5: EXPERIMENTAL TECHNIQUES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Sample design for measurement of ionization rates</td>
<td>92</td>
</tr>
<tr>
<td>5.2 Device fabrication procedure</td>
<td>96</td>
</tr>
<tr>
<td>5.3 Electrical and optical characterization</td>
<td>97</td>
</tr>
<tr>
<td>5.4 Apparatus for measurement of ionization rates</td>
<td>99</td>
</tr>
</tbody>
</table>
5.5 Measurement of the pulse response of photodiodes 102
5.6. Low temperature techniques 105

CHAPTER 6: Ga_{0.47}In_{0.53}As/InP MULTIPLE QUANTUM WELLS APD'S GROWN BY VPE AND MOCVD 105

6.1 Ga_{0.47}In_{0.53}As/InP MQW APD's grown by continuous trichloride vapour phase epitaxy 105
6.1.1 Material growth 105
6.1.2 Characterization of layers 106
6.1.3 Device fabrication and characteristics 107
6.1.4 High speed response 112

6.2 High speed Ga_{0.47}In_{0.53}As/InP MQW APD's grown by MOCVD 117
6.2.1 Material growth 117
6.2.2 Device characteristics 117
6.2.3 High speed behaviour 121

6.3 Ionization rates in Ga_{0.47}In_{0.53}As/InP MQW APD's 124

CHAPTER 7: Al_{0.46}In_{0.52}As/Ga_{0.47}In_{0.53}As MQW APD'S GROWN BY MBE 126

7.1 Device structure and growth 128
7.2 Device fabrication and characteristics 136
7.3 Impact ionization rates 143
7.4 High frequency response 143

CHAPTER 8: IMPACT IONIZATION ACROSS THE BAND EDGE DISCONTINUITY IN MULTIPLE QUANTUM WELL STRUCTURES 148

8.1 Introduction 148
8.2 MQW's with doped wells: theoretical calculation of ionization rates 151
8.3 MQW's with undoped wells: experimental evidence 152
8.4 Single carrier type multiplication in multiple graded-well samples 173
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>PERPENDICULAR TRANSPORT IN MULTIPLE QUANTUM WELL STRUCTURES</td>
<td>188</td>
</tr>
<tr>
<td>9.1</td>
<td>Sequential resonant tunnelling</td>
<td>188</td>
</tr>
<tr>
<td>9.2</td>
<td>Resonant Zener tunnelling</td>
<td>206</td>
</tr>
<tr>
<td>9.3</td>
<td>Transport in an undepleted Ga$<em>{0.47}$In$</em>{0.53}$As/InP structure</td>
<td>214</td>
</tr>
<tr>
<td>10</td>
<td>CONCLUSIONS AND RECOMMENDATIONS</td>
<td>225</td>
</tr>
<tr>
<td>10.1</td>
<td>Performance of long-wavelength MQW APD's</td>
<td>225</td>
</tr>
<tr>
<td>10.2</td>
<td>Measurement of ionization rates in a MQW</td>
<td>226</td>
</tr>
<tr>
<td>10.3</td>
<td>Impact ionization across the band-edge discontinuity</td>
<td>227</td>
</tr>
<tr>
<td>10.4</td>
<td>Tunnelling in MQW 's</td>
<td>228</td>
</tr>
<tr>
<td>10.5</td>
<td>Transport in undepleted wells</td>
<td>229</td>
</tr>
</tbody>
</table>
This thesis reports experimental studies of electrical transport in multiple quantum well (MQW) semiconductor structures, under an electric field applied perpendicular to the heterojunction interfaces. In particular impact ionization phenomena in long-wavelength MQW avalanche photodiodes (APD's) were studied. Work on such structures has been motivated by the requirement for low-noise, long-wavelength photodetectors for use in optical fibre communications.

1.1 Development of low-loss optical fibres.

Early optical-fibre communications systems operated at a wavelength (\(\lambda\)) of 0.85 \(\mu\)m, using multimode fibres with transmission losses of the order of 4 dB-Km\(^{-1}\) and Si APD detectors. Attention switched to longer wavelengths as low-loss, low-dispersion silica fibres and semiconductor lasers operating CW at these wavelengths became available.

The transmission of silica fibres is limited at short wavelengths by Rayleigh scattering, which is proportional to \(\lambda^{-4}\), and at long wavelengths by multiphonon scattering which causes a sharp reduction in transmission for \(\lambda \approx 1.6 \mu m\). Absorption associated with the O-H bond causes reduction in the transmission at \(\lambda \approx 1.24\) and \(\lambda \approx 1.4 \mu m\).

Zero material dispersion in silica fibres occurs at \(\lambda \approx 1.3 \mu m\) (Payne and Gambling, 1975). Dispersion-free fibres with losses at 1.3 \(\mu m\) of 0.5 dB-km\(^{-1}\) were demonstrated by Kawacki et al. (1977). Miya et al. (1979) demonstrated low-loss single-mode fibres with losses of \(\approx 0.2 \text{ dB-km}^{-1}\) at 1.55 \(\mu m\) limited only by Rayleigh scattering, with
dispersion sufficiently small for use with narrow band-width sources such as lasers. The transmission loss of a typical single-mode silica fibre is shown in figure 1.1

1.2 Photodetectors for optical fibre communications.

The principle requirements for photodetectors for optical fibre communications are high sensitivity, low noise and fast response time. High sensitivity is obtained by using a material with a high absorption coefficient \( \alpha_0 \) in the wavelength region of interest. Absorption coefficients are typically \( \geq 10^4 \text{ cm}^{-1} \) for above-bandgap radiation, so incident light is absorbed within a few microns. The internal quantum efficiency, \( \eta_{\text{int}} \), for an absorbing region of width \( d \), is

\[
\eta_{\text{int}} = 1 - \exp(-\alpha_0 d) \tag{1.1}
\]

if all photogenerated carriers are collected. The external quantum efficiency \( \eta_{\text{ext}} \) is reduced by reflections at the semiconductor-air interface. The reflectivity \( R \) is

\[
R = \frac{(n_{\text{semi}} - n_{\text{air}})^2}{(n_{\text{semi}} + n_{\text{air}})^2} \tag{1.2}
\]

where \( n \) is the refractive index, and is typically \( \approx 30 \% \).

The photocurrent, \( I_{\text{ph}} \), generated by illumination with light of energy \( h\nu \) and intensity \( P_0 \), in the absence of any current multiplication process, is

\[
I_{\text{ph}} = \frac{\eta_{\text{ext}} P_0}{h\nu}. \tag{1.3}
\]

Four detector types are most commonly considered for use with optical fibres: unity gain or "P-I-N" photodiodes (see chapter 2), APD's, phototransistors and photoconductors. These have been compared
Figure 1.1. Transmission loss and dispersion in a typical silica optical fibre.
in detail by a number of authors (Brain and Smith, 1983; Milano et al, 1982, Forrest, 1987). APD's, phototransistors and photoconductors have internal gain mechanisms to increase the quantum efficiency above unity. P-I-N's and APD's operate as reverse-bias PN junctions, and thus have low dark current and dark-current noise. Although photoconductors and phototransistors display a number of advantages such as low-voltage operation and gain stability, most success has been achieved with P-I-N's and APD's.

P-I-N photodiodes are operated at low bias and unity gain, thus the shot noise is small and there is no excess noise associated with current multiplication. High efficiency and high speed response can be obtained when all the light is absorbed in the depleted region. The photogenerated carriers are swept apart by the applied field and collected by the contacts. The intrinsic "I" region is several microns wide to ensure high internal efficiency and is low doped so that it is depleted at low bias. The pulse response is limited by the transit time of the device and the capacitance of the depletion region. Ga_{0.47}In_{0.53}As P-I-N diodes with unity quantum efficiency, low leakage current (<10 nA) and response times of a few tens of picoseconds are commercially available.

APD's employ impact ionization to give photocurrent multiplication at high bias. This leads to increased sensitivity, but also to increased shot noise due to the higher dark current and excess noise from the multiplication process itself. The response time is also increased by the time required for the avalanche build-up which leads to a gain-bandwidth product limitation. The noise and high speed performance of APD's are determined by the ionization rate ratio, as described in chapter 2.

The silicon APD offers high performance for photodetectors at
wavelengths less than about 1 μm, due to its low dark current and small excess noise. Germanium, although it has a high responsivity at wavelengths up to ≈ 1.8 μm, has an ionization rate ratio close to unity leading to large excess noise. Thus attention has focussed on alloys of the III-V semiconductors, some of which have direct band-gaps in the wavelength region of interest. As with Ge however, the excess noise performance of most III-V APD's is poor.

Multilayer semiconductor structures offer the possibility of tailoring material characteristics to suit specific device applications ("band-structure engineering"). Heterostructure and multiple quantum well APD's have shown improved performance over bulk devices, although a long-wavelength APD with noise performance comparable to Si has yet to be demonstrated.

1.3 Organization of thesis.

Electrical transport in multiple quantum wells and implications for the design of long-wavelength APD's are discussed in this thesis in the following format. Chapter 2 discusses the relationship between multiplication, impact ionization rates and excess noise in APD's. Theories of impact ionization and the role of band-structure are treated briefly, and previous measurements of ionization rates in some III-V semiconductors are surveyed.

Electrical and optical properties of heterostructures and quantum wells are discussed briefly in chapter 3. Theoretical and experimental work on impact ionization in multiple quantum well APD's is reviewed.

Chapter 4 summarizes the relevant material parameters of the semiconductors considered in this thesis, and describes the growth techniques used. Chapter 5 describes the sample design and fabrication, and the experimental techniques used to characterize the photodiodes.
In chapter 6, the performance of a Ga$_{0.47}$In$_{0.53}$As/InP multiple quantum well (MQW) APD is discussed, including the high speed performance. Chapter 7 reports high-speed long-wavelength MQW APD's fabricated from Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As grown by molecular beam epitaxy. Avalanche multiplication at high frequency and a possible enhancement of the ionization rates were investigated.

In chapter 8, a new impact ionization phenomenon in MQW APD's, impact ionization across the heterojunction discontinuity, is described. Photocurrent multiplication effects in quantum well structures fabricated in several different material systems showed experimental evidence for this process. Single carrier-type multiplication in a III-V APD is reported in a multilayer APD with compositionally-graded interfaces.

In chapter 9, some aspects of perpendicular transport in MQW's are discussed. Sequential resonant tunnelling and resonant Zener tunnelling were studied in Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW's with low background doping. Ga$_{0.47}$In$_{0.53}$As/InP structures grown by gas source molecular beam epitaxy, which were not fully depleted at zero bias, are also discussed. Carrier transport in this case can be dominated by localisation in the undepleted wells.

Chapter 10 presents conclusions and suggests areas of further work.
CHAPTER 2

PROPERTIES OF BULK III-V SEMICONDUCTORS

This chapter describes properties relevant to APD's in III-V bulk (as opposed to multilayer) semiconductors.

Section 2.1 summarizes important elements of the band-structure of III-V semiconductors. Properties of reverse-biased P-N junctions are discussed in section 2.2.

Impact ionization in bulk III-V's is reviewed in section 2.3. Multiplication factors, ionization rates, noise factors, ionization thresholds and theories of impact ionization are discussed, and a brief survey of experimental results in some III-V materials is presented.

2.1 Band-structure of III-V semiconductors.

The electrical properties of crystalline materials are largely determined by their energy band-structure, the general features of which are determined by the crystal structure. The periodic potential of the crystal lattice gives rise to a band structure with regions of allowed energy separated by forbidden regions or "band-gaps". Semiconductors are characterised by band-gaps \( E_g \) in the range \( 0 < E_g \leq 2 \text{ eV} \), giving rise to electrical properties strongly dependent on temperature and on impurity concentration.

The crystal structure of III-V compounds is that of zinc-blende (see figure 2.1(a)), with each atom surrounded by four nearest neighbours of different valency. The regular tetrahedron arrangement is equivalent to two interpenetrating face-centred cubic lattices. The corresponding Brillouin zone is body-centred cubic (figure 2.1(b)). The principle symmetry points are \( \Gamma \) at the zone centre, \( X \) at the (100)
Figure 2.1(a). Zincblende structure.
(b). Brillouin zone and symmetry points of face-centred cubic lattice.
zone boundary and L at the (111) zone boundary. The degeneracy of band-edge features is determined by the crystal symmetry for the zinc-blende structure there are 8 equivalent L points and 6 equivalent X points.

The band-structure of GaAs is shown in figure 2.2. The band-gap at the Γ-point is $E_0$, at the L-point is $E_1$ and at the X-point is $E_2$. The ordering of band minima in a given semiconductor is determined by material parameters. The valence bands have a single maxima at the Γ point. The III-V materials of interest for opto-electronic devices have the lowest conduction-band minima at the Γ-point. $E_0$ is then referred to as the direct band-gap, $E_g$.

The band-structure in the vicinity of the zone centre is shown in figure 2.3. The valence bands consist of a heavy-hole band and a light-hole band which are degenerate at the zone centre, and a spin-orbit band separated from the heavy- and light-hole bands at the zone centre by the spin-orbit splitting $\Delta_0$. The bands are approximately parabolic and are characterized by different values of the effective mass $m^*$, defined as

$$\frac{1}{m^*} = \frac{1}{n^2} \cdot \frac{d^2E(k)}{dk^2}$$

(2.1)

The effective mass of a parabolic band is thus constant. Deviations from parabolicity, and the presence of higher-lying conduction bands, become important at the high carrier energies required for impact ionization.

Energy gaps can be determined experimentally by optical measurements. Effective masses can be obtained from magneto-transport measurements or Raman scattering.

A number of methods exist for calculating band-structures given certain material parameters. These include the quantum dielectric
Figure 2.2. Band-structure of GaAs calculated by pseudopotential method (after Chelikowsky and Cohen, 1976).
Figure 2.3. Energy band-structure at the zone centre.
theory (Philips, 1968; van Vechten, 1969), Keldysh perturbation theory
(Kane, 1966) and pseudopotential techniques (Cohen and Bergstrasser,
1966; Chelikowsky and Cohen, 1976).

2.2 Properties of a reverse-biased P-N junction.

P-N junctions are formed at the interface between regions of a
semiconductor doped with acceptors (P-type) and with donors (N-type).
In this section the capacitance-voltage and current-voltage
characteristics of reverse-bias P-N junctions are summarised.

As a P and an N region are brought into contact, majority carriers
diffuse from the doped regions. In a region around the junction the
carriers recombine leading to depletion of free carriers. The space
charge from the ionized impurities leads to an internal field which
opposes the diffusion current. The steady state is obtained when the
diffusion of majority carriers is balanced by the drift of
carriers in this internal field, which occurs when the Fermi levels in
the P and N regions are coincident (see, for example, chapter 2 of Sze,
1981). The built-in potential for zero applied bias is

$$V_{bi} = \frac{k_B T}{e} \ln \left( \frac{N_A N_D}{n_i^2} \right),$$

where $N_A$ is the acceptor concentration in the P region, $N_D$ is the donor
concentration in the N region, $n_i$ is the intrinsic carrier
concentration at temperature $T$ and $k_B$ is the Boltzmann constant.

Three types of P-N junction are relevant to this thesis: the P$^+$-N
abrupt single-sided junction, the P$^+$-N-N$^+$ ("punch-through") and the
P$^+$-I-N$^+$ ("PIN") junction. The charge density, field and potential in
these devices are shown in figure 2.4.
Figure 2.4. Charge density, field and potential profile in P\textsuperscript{+}-N, P\textsuperscript{+}-N-N\textsuperscript{+} and P\textsuperscript{+}-I-N\textsuperscript{+} diodes.
2.2.1 Depletion capacitance.

The width \( W(V) \) of the depletion layer can be obtained from Poisson's equation:

\[
\frac{d^2V}{dx^2} = \frac{dF}{dx} = \frac{\rho(x)}{\varepsilon_s}
\]  

(2.3)

where \( \rho(x) \) is the charge density and \( \varepsilon_s \) is the dielectric constant of the semiconductor. Integration under appropriate boundary conditions leads to

\[
W(V) = \sqrt{\frac{2\varepsilon_s}{e} \left[ \frac{N_A + N_D}{N_A N_D} \right] (V + V_{bi})}.
\]  

(2.4)

For a P⁺-N-N⁺ punch-through structure

\[
W(V) = \left\{ \begin{array}{ll}
\sqrt{\frac{2\varepsilon_s}{e} N_D} (V + V_{bi}) & (V < V_{pt}) \\
L & (V \geq V_{pt})
\end{array} \right.
\]  

(2.5)

where \( L \) is the width of the N region and \( V_{pt} \) is the punch-through bias

\[
V_{pt} = \frac{L^2 e N_D}{2\varepsilon_s} - V_{bi}.
\]  

(2.6)

The capacitance of a P⁻N junction is equivalent to that of a parallel plate capacitor with plate separation \( W(V) \) and dielectric constant \( \varepsilon_s \). Thus for a P⁺-N diode

\[
C(V) = \frac{\varepsilon_s A}{W(V)} = A \sqrt{\frac{\varepsilon_s e N_D}{2(V + V_{bi})}}
\]  

(2.7)
and for a P⁺-N-N⁺ with $V > V_{pt}$

$$C(V) = \frac{\epsilon_s A}{L}$$ (2.8)

### 2.2.2 Dark current mechanisms.

P-N junctions have high resistance in reverse bias below breakdown, since diffusion of majority carriers from the neutral regions is prevented in reverse bias by the field. Dark currents arise through three mechanisms: minority-carrier diffusion from the neutral regions, thermal generation within the depletion region and band-to-band tunnelling.

Minority carriers thermally generated in the neutral region, within a diffusion length of the depletion edge, give rise to a leakage current (Shockley, 1949) given by

$$I_{diff} = I_S [\exp(eV/kT) - 1]$$ (2.9)

where

$$I_S = eN_i^2 \left[ \frac{D_n}{\tau_n} \frac{A}{N_A} + \frac{D_p}{\tau_p} \frac{A}{N_D} \right].$$ (2.10)

$D$ is the minority-carrier diffusion constant and $\tau$ is the minority-carrier diffusion lifetime.

Thermal generation of electron-hole pairs within the depletion region, via Shockley-Read centres in the band-gap (Shockley and Read, 1952), can be an important leakage current mechanism in reverse bias. The generation rate is governed by four processes: electron and hole capture and emission (Sah, Noyce and Shockley, 1957). The generation
current, for the simplified case in which electron and hole capture rates are equal \( \sigma_n = \sigma_p = \sigma \) is given by

\[
I_{g-r} = \frac{e \chi_{AW}(V)}{\tau_{\text{eff}}} [\exp(eV/2k_BT) - 1],
\]

where \( \tau_{\text{eff}} \) is an effective lifetime given by

\[
\tau_{\text{eff}} = \frac{\cosh [(E_t - E_i)/k_BT]}{\sigma v_{th} N_t}
\]

and \( v_{th} \) is the thermal velocity, \( N_t \) is the density of centres at energy \( E_t \) and \( E_i \) is the intrinsic Fermi level. For neutral material \( \tau_{\text{eff}} = \tau_i = 1/(\sigma v_{th} N_t) \), typically of the order of \( \approx 10^{-9} \) s for III-V's.

Zener tunnelling of electrons from the valence band to the conduction band is significant for narrow gap semiconductors at high reverse fields (Forrest et al., 1980). Tunnelling can occur from band-to-band or via a band-gap state. The tunnelling current for a P-I-N diode is given by (Moll, 1964)

\[
I_{\text{tun}} = \gamma A \exp \left[ - \frac{c m^* \eta_{\text{tun}}^{3/2}}{eF} \right],
\]

where \( F \) is the field, and

\[
\gamma = \left[ \frac{2m^*}{E_g} \right]^{3/2} \left[ \frac{e^2 F V}{4\pi^2 \hbar^2} \right]
\]

for band-to-band tunnelling. \( E_{\text{tun}} \) is the tunnelling barrier height which is equal to \( E_g \) for band-to-band tunnelling, and \( c \) is a parameter reflecting the shape of the potential barrier (1.11 for a parabolic barrier, 1.89 for a triangular barrier). Evidence for both band-to-band and defect-assisted tunnelling has been observed in a number of narrow-
2.3 Impact ionization in bulk materials.

In this section the process of impact ionization in bulk materials is discussed. More details can be found in the reviews of Stillman and Wolfe (1977) and Capasso (1985), and the references contained therein.

Avalanche multiplication due to impact ionisation occurs in reverse-biased P-N junctions at high electric fields. When a carrier is heated by the field to sufficiently high energy, a collision with an electron in the valence band may impart sufficient energy to that electron to ionize it across the band-gap into the conduction band, giving rise to an additional electron-hole pair. These carriers may also be heated and impact ionize, giving rise to an avalanche. The dark current is multiplied by this process.

Similarly, carriers photogenerated in a photodiode will be multiplied at high fields. This internal gain mechanism in an avalanche photodiode (APD) leads to increased responsivity.

2.3.1 Multiplication factors and ionization rates.

The multiplication at a given field due to impact ionization is determined by the average distance travelled by a carrier before it impact ionizes. The inverse of this distance is the ionization coefficient, $\alpha(F)$ for electrons and $\beta(F)$ for holes, expressed in units of cm$^{-1}$. The coefficients for electrons and holes may be quite different since both the scattering rates and threshold energies required for ionization to occur will differ.

The multiplication rates for electrons and holes can be directly related to the ionization coefficients in a given device structure. The electron current $J_e(x)$ and hole current $J_h(x)$ within a high field
depletion region of width $x$ are governed by the rate equations

$$\frac{d}{dx} J_e(x) = \alpha(x)J_e(x) + \beta(x)J_h(x) + eG(x) ,$$

(2.15)

$$- \frac{d}{dx} J_h(x) = \alpha(x)J_e(x) + \beta(x)J_h(x) + eG(x) .$$

(2.16)

where $G(x)$ is the generation rate of carriers at $x$. Multiplication by an integrating factor $\exp[-\phi(x)] = \exp[-\int_0^x (\alpha - \beta)dx']$ and integration from $x=0$ to $x=W$ leads to the following expression for the total current, $J$:

$$J = \frac{J_h(W) + J_e(0)\exp[\phi(W)] + e\exp[\phi(W)] \int_0^W G(x)\exp[-\phi(x)]dx}{1 - \int_0^W \beta(x)\exp\left[\int_0^x (\alpha - \beta)dx'\right]dx}$$

(2.16)

It is easy to calculate the current from the above equation given a knowledge of the ionization coefficients. However, the usual experimental problem is to calculate $\alpha$ and $\beta$ from the measured photocurrent. The charge injection conditions are simplified by ensuring that no carriers are generated inside the depletion region, and considering separately the cases when holes are injected at $x=W$ or electrons are injected at $x=0$ (referred to as pure hole and pure electron injection, respectively).

$$M_h = \frac{J}{J_h(W)} = \frac{\exp[-\int_0^W (\alpha - \beta)dx']}{1 - \int_0^W \alpha \exp[-\int_0^x (\alpha - \beta)dx']dx}$$

(2.17)

$$M_e = \frac{J}{J_e(0)} = \frac{1}{1 - \int_0^W \alpha \exp[-\int_0^x (\alpha - \beta)dx']dx}$$

The ionization rates derived from the above rate equations are not
generally identical to the microscopic rates calculated theoretically as the reciprocal of the average distance travelled by a carrier before it reaches threshold. If there is significant variation in the field over an ionization distance, the ionization rates become explicitly position dependent, and a spatial steady state distribution is not reached. Beni and Capasso (1979) related the measured quantities $\alpha'$ and $\beta'$ to the microscopic rates $\alpha$ and $\beta$

\[
\frac{\alpha - \alpha'}{\alpha - \beta} = \frac{1}{v_e} \frac{dV_e}{dF} \frac{dF}{dx}
\]

(2.18)

\[
\frac{\beta - \beta'}{\alpha - \beta} = \frac{1}{v_h} \frac{dV_h}{dF} \frac{dF}{dx}
\]

Thus the measured rates are only identical to the microscopic rates when $\frac{dF}{dx}$ is zero, i.e. in a P-I-N diode. In this device structure the field and thus the ionization rates are constant throughout the depletion region. This also leads to a simplification of the above equations

\[
M_h = \frac{\alpha - \beta}{\alpha - \beta \exp[(\alpha-\beta)W]}
\]

(2.19)

\[
M_e = \frac{(\alpha - \beta) \exp[(\alpha-\beta)W]}{\alpha - \beta \exp[(\alpha-\beta)W]}
\]

where $M_h$, $M_e$, $\alpha$ and $\beta$ are functions of the field.

Two extreme cases exist. If the ionization rates are equal ($\alpha=\beta$), then

\[
M_h = M_e = \frac{1}{1 - \alpha W} .
\]

(2.20)

There is a true breakdown ($M = \infty$) when $\alpha W = 1$, when each injected carrier experiences one impact ionization in traversing the depletion width. In this case, the hole generated by an ionizing electron will be accelerated back across the depletion region and impact ionize, giving
rise to a positive feedback which causes the rapid breakdown with increasing field. For a large multiplication there need only be one impact ionization per traversal of the depletion region, due to the positive feedback. This implies that for a large gain the current pulse width due to a single injected electron will be long, due to the large number of traversals made. Thus there is a gain-bandwidth product limit. Finally, again since there is only one ionization per traversal, at any given time there are few carriers within the depletion region; thus a small random variation of the number of carriers will be significant, leading to a large excess noise.

However, if one of the ionization rates is zero (e.g. \( \alpha = 0 \)) then \( M_e = 1 \) and \( M_h = \exp(\beta W) \). Thus there is an exponential increase in multiplication with field, leading to good gain stability. Since for a large gain there are many carriers within the depletion region, this is the optimum situation in terms of noise performance and gain-bandwidth product (the current pulse width is simply the sum of the electron transit time and the hole transit time).

The maximum multiplication obtained in an APD may be limited by a number of effects which occur at high current density, such as voltage drop across the load resistor and contacts and resistance heating of the junction causing an increase in the breakdown field. Thus the dark current and photocurrent should be small to achieve maximum gain.

The maximum useful multiplication is also limited by spatial non-uniformity in the photoresponse. This can be caused by doping fluctuations or, in particular, by microplasmas. These are regions of local high electric field, often associated with recombination centres and macroscopic clusters of dislocations (Magnea et al., 1985).
Noise in an APD arises from three sources: the shot-noise associated with generation of the primary carriers, the excess noise associated with the multiplication process and the thermal noise of the external circuit. The multiplied rms photocurrent for an APD with multiplication $M$ for a sinusoidal optical signal

$$P(\omega) = P_0 [1 + \cos(\omega t)]$$

is

$$i_p = e \eta P_0 M / h \nu/2 , \quad (2.21)$$

where $\eta$ is the quantum efficiency. The total current is the sum of the average photocurrent

$$I_p = e \eta P_0 M / h \nu , \quad (2.22)$$

and the dark current ($I_d$).

The primary carriers are generated by random processes governed by Poisson statistics. Thus they contribute a mean-square shot-noise current $\langle i_s^2 \rangle$ given by

$$\langle i_s^2 \rangle = 2eIB , \quad (2.23)$$

in a bandwidth $B$, in the absence of multiplication. If the multiplication was noise-free, the shot-noise would be multiplied by $M^2$. However, since impact ionization is also a random process there will be an excess noise associated with the multiplication, expressed as a noise factor $F$

$$\langle i_s^2 \rangle = 2eIM^2FB . \quad (2.24)$$

The noise figure is dependent on the diode structure and the properties of the material. The total shot-noise is thus

$$\langle i_s^2 \rangle = 2eB(I_pF_p + I_dF_D)M^2$$

(2.25)
where $F$ is the excess noise factor for each of the shot-noise sources.

Thermal noise arises from the interaction of the detector with the external circuit, in the simplest case an external load resistor $R_L$ at temperature $T_L$. The thermal noise current is given by

$$
\langle i_L^2 \rangle = 4k_B T_L (1/R_L) B .
$$

Thus the signal-to-noise ratio is

$$
\frac{S}{N} = \frac{i_p^2 (\lambda, \omega) R_L}{\langle i_N^2 \rangle R_L} = \frac{\frac{1}{2} (e^2 P_0 / \hbar v)^2 M^2}{\langle i_S^2 \rangle + \langle i_L^2 \rangle} \frac{\hbar (e^2 P_0 / \hbar v)^2 / B}{2e(I_p F_p + I_p F_D) + (4k_B R_L / R_L M^2)}
$$

At high incident intensity, the signal-to-noise ratio is limited by the shot-noise produced by the photocurrent (the quantum noise limit). At low incident intensities the noise is dominated by the thermal noise currents, and $S/N$ increases with increasing $R_L$. However, this leads to frequency limitations because of the roll-off of the load resistor-junction capacitance circuit. For high bandwidth devices, $R_L$ is always sufficiently small that the thermal noise of the following circuit dominates the noise current. Thus multiplication decreases $S/N$ by decreasing the importance of this term. The optimum multiplication is determined by the relative size of the excess noise and the thermal noise.

2.3.3 McIntyre noise theory.

The shot-noise current due to an average current $I$ for a current multiplication $M$ was given above by

$$
\langle i_S^2 \rangle = 2eIM^2FB,
$$

for an excess noise factor $F$ and bandwidth $B$. The presence of positive feedback leads to an increase in the noise due to multiplication, and for $\alpha=\beta$ the noise current is $\langle i_S^2 \rangle = 2eIM^3B$ (Tager, 1956).
The noise factor for arbitrary $\alpha, \beta$ was calculated by McIntyre (1966). The noise factors for pure injection of electrons ($F_e$) or holes ($F_h$) are given by

$$F_e = M_e \left[ \left( 1 - \frac{1}{k} \right) \left( \frac{M_e - 1}{M_e} \right)^2 \right]$$

$$F_h = M_h \left[ \left( 1 - \frac{1}{k} \right) \left( \frac{M_h - 1}{M_h} \right)^2 \right].$$

where $k (= \beta/\alpha)$ is independent of position for the case of a P-I-N diode. For low excess noise at useful multiplications, the ionization rates for electrons and holes must be very different. For the best case ($\alpha = 0$) $F = 2 - 1/M$ and for the worst case ($\alpha = \beta$) $F = M$. Thus for large gain the shot noise is at best twice that for noise-free multiplication. This is due to the spatial indeterminancy of the ionization process. For low noise operation the carrier with the highest ionization rate should be injected into the depletion region.

### 2.3.4 Frequency response of bulk APD's.

The response of bulk APD's to high speed light pulses is limited by a number of factors: the RC time constant of the depletion capacitance and load resistor, the transit time of carriers across the depletion region and the avalanche build-up time.

The RC time constant is minimized by reducing the device area and increasing the depletion width. The pulse width (full-width at half-maximum, FWHM) for a device of capacitance $\approx 1$ pF measured with a 50 $\Omega$ sampling head is

$$\tau_{RC} \approx 2.2 \ \text{RC} \approx 100 \ \text{ps}. \quad (2.29)$$

In the absence of multiplication the intrinsic response time is limited by the transit time ($\tau_{tr}$) of carriers across the depletion region (width $L$). For a saturated drift velocity $v_{sat}$, the transit time is...
Electron and hole saturated drift velocities in GaAs and InP have been measured experimentally (Windhorn et al., 1982b; Holway et al., 1979) and calculated from Monte Carlo simulations (Brennan and Hess, 1984), and are \( \approx 7 \times 10^6 \, \text{cm-s}^{-1} \). The measured saturation velocity for electrons in Ga_{0.47}In_{0.53}As is \( \approx 6 \times 10^6 \, \text{cm-s}^{-1} \) (Windhorn et al., 1982a). Experimental (Hill et al., 1987) and theoretical (Brennan, 1987) values for the hole velocities in Ga_{0.47}In_{0.53}As are \( \approx 5 \times 10^6 \, \text{cm-s}^{-1} \) and \( \approx 6 \times 10^6 \, \text{cm-s}^{-1} \) respectively. Thus a typical transit time for a 1 \( \mu \)m depletion region is \( \approx 15 - 20 \, \text{ps} \).

Avalanche multiplication, in the case when both ionization rates are non-zero, introduces a gain-bandwidth product due to the increased effective carrier transit length in the presence of positive feedback. Emmons (1967) has calculated the frequency response for a P-I-N APD as \( \beta/\alpha \) varies from 0 to 1, by calculating the time dependent transport equations assuming that the carrier velocities are constant and equal for electrons and holes. For the case

\[
(\beta/\alpha)(M_e-1)\ln(M_e) \leq 1
\]  

(i.e. low multiplication and small \( \beta/\alpha \)) the multiplication is virtually constant with frequency up to the limits of the single carrier transit time and RC time constants. At a multiplication of 10, for example, this requires \( \beta/\alpha < 0.06 \); for \( M_e = 100, \beta/\alpha < 0.002 \). The multiplication at an angular frequency \( \omega \) is given by

\[
M(\omega) = \frac{M_e}{(1+\omega^2M_e^2\tau_1^2)^{1/2}}
\]  

for \( M_e > \alpha/\beta \). The gain-bandwidth limit at high frequencies is thus

\[
M(\omega) \cdot \omega = \frac{1}{\tau_1}
\]
The effective transit time $\tau_1$ is approximately $N(\beta/\alpha)\tau_tr$, where $\tau_tr$ is the transit time for a single carrier and $N$ is a numerical factor varying between 1/3 for $\beta/\alpha = 1$ and 2 for $\beta/\alpha = 10^{-3}$. At high multiplications and frequencies comparable to $1/\tau_1$, the gain-bandwidth product is thus directly proportional to $\alpha/\beta$.

2.3.5 Threshold Energies.

The energy required by a carrier for it to impact ionize an electron from the valence to the conduction band is the threshold energy, $E_{th}^{(e)}$ for electrons and $E_{th}^{(h)}$ for holes. In most theoretical treatments, a carrier is assumed to ionize as soon as it reaches the threshold energy, a so-called "hard threshold". Threshold energies are determined by the band structure of the semiconductor. The total energy of the final carriers is a minimum at threshold, subject to conservation of energy and momentum (Anderson and Crowell, 1972). For the case of parabolic conduction and valence bands with effective masses $m_e$ and $m_h$, the threshold energies for electrons and holes are

$$E_{th}^{(e)} = E_g \left[ 1 + \frac{m_e}{m_e+m_h} \right]$$

$$E_{th}^{(h)} = E_g \left[ 1 + \frac{m_h}{m_e+m_h} \right]$$

Thus

$$E_{th}^{(e)} = E_{th}^{(h)} = 3/2 \ E_g$$

for equal effective masses. Simple analytical formulae can also be obtained for the case of three parabolic bands (Pearsall et al., 1977; Pearsall, 1979) and for a non-parabolic conduction band (Ridley, 1977). The case of three parabolic bands illustrates the effects of the spin-orbit splitting. The electron- and hole-initiated ionization processes
are shown in figure 2.5. The calculated threshold energies are

\[
\begin{align*}
E_{\text{th}}^{(e)} &= E_g \left[ 1 + \frac{m_e}{m_{hh} + m_e} \right], \\
E_{\text{th}}^{(h)} &= E_g \left[ 1 + \frac{m_{s-o}(1 - A/E_g)}{2m_{hh} - m_{s-o} + m_e} \right],
\end{align*}
\] (2.36)

where \( A \) is the spin-orbit splitting. The electron ionization threshold is unchanged, whereas the hole threshold is reduced by the spin-orbit splitting. In the special case where \( A = E_g \) the hole threshold is simply equal to the band-gap, and zero momentum is transferred. The material system \( \text{Al}_x\text{Ga}_{1-x}\text{Sb} \) is well-described by three parabolic bands at energies up to the ionization threshold, and exhibits a strong resonant enhancement of the hole ionization rate at \( x = 0.065 \), when \( A = E_g \).

The band-structure of many III-V semiconductors shows non-parabolicity and anisotropy at energies below threshold. For more accurate calculations of the thresholds, realistic band-structures are required. Calculations have been reported (Anderson and Crowell, 1972; Pearsall et al., 1977, 1978; Pearsall, 1979) based on pseudopotential band-structures. Figure 2.2 shows the band structure of GaAs after Chelikowsky and Cohen (1976). The valence bands are fairly isotropic, so the hole thresholds are not strongly dependent on crystal orientation. The conduction bands, however, are strongly anisotropic. In the \( <100> \) direction, the electron threshold state is in the second conduction band. In the \( <110> \) direction, the threshold state is near the top of the first conduction band. However, in the \( <111> \) direction, the width of the first conduction band is less than the band-gap, and the second conduction band is separated in energy by more than 2 eV. Thus there is no threshold energy in the \( <111> \) direction. Despite the anisotropy in the threshold energy, experimental measurements have
Figure 2.5. Impact ionization in three parabolic bands;

(a) electron initiated,

(b) hole initiated.
generally shown that the ionization rates are independent of crystal orientation. This may be related to randomisation of the electron momentum by intervalley scattering, which dominates at high fields.

2.3.6 Theoretical calculation of impact ionization rates.

The rate at which carriers reach the threshold energy for ionization is determined by the rate of carrier heating by the field and the rate of energy relaxation by phonon scattering. For III-V semiconductors with band-gaps \( \gtrsim 0.5 \text{ eV} \) the dominant scattering mechanism at fields typical of impact ionization is intervalley scattering, due to collisions with zone-edge phonons of large wavevector and energies of a few tens of meV. The ionization rates have been calculated theoretically by finding the carrier distribution at high fields from solution of the Boltzmann transport equation. The ionization rate is given by

\[
\alpha = \left( \frac{eF}{E_{\text{th}}} \right) P(E_{\text{th}}),
\]

where \( P(E_{\text{th}}) \) is the probability of a carrier reaching the threshold energy. At high fields \( P \) tends to 1 and \( 1/\alpha \) tends to \( E_{\text{th}}/eF \), the minimum distance for a carrier to ionize. Parabolic energy bands, isotropic scattering, "hard" ionization thresholds and an energy-independent mean free path \( (\lambda_R) \) are generally assumed in these approaches.

The distribution function is determined by the ratio of the energy gained from the electric field between collisions \( (eF\lambda_R) \) to the energy lost per collision \( (E_R) \). Wolff (1954) considered the case when \( eF\lambda_R \gg E_R \). There are many collision events before a carrier reaches threshold, and thus a displaced Maxwellian for the high-field carrier distribution. This leads to an ionization rate
\[
\alpha(F) = \frac{eF}{E_{th}} \exp\left[ -\frac{3E_R E_{th}}{(eF\lambda)^2} \right].
\] (2.38)

Shockley (1961) considered the opposite extreme \( eF \lambda < E_R \), when a carrier will only reach threshold if it escapes phonon collisions ("lucky electron" theory). Thus the carrier travels ballistically over a distance \( E_{th}/eF \), giving an ionization rate

\[
\alpha(F) = \frac{eF}{E_{th}} \exp\left[ -\frac{E_{th}}{eF\lambda} \right].
\] (2.39)

Baraff (1962) noted that a typical breakdown fields neither of the above two extreme cases holds, and combined the two models by solving the Boltzmann equation using a distribution function containing a spherical component and a spike in the field direction. The Baraff curves are shown in figure 2.6 (dashed line) for different ratios of the phonon energy \( E_R \) to the ionization threshold \( E_{th} \). Experimental data have been widely fitted to these curves and used to determine \( \lambda_R \).

Ridley (1983a; 1983b) has proposed a model for impact ionization which gives an analytical expression for ionization rates in close agreement with the Baraff curves. At high energies, where intervalley scattering dominates, the relaxation rates for energy (\( \tau_E \)) and momentum (\( \tau_m \)) are different, with \( \tau_m \ll \tau_E \). This gives rise to a state ("lucky drift") in between those of Wolff and Shockley, where carriers can drift in an electric field without suffering energy relaxation, with a drift velocity determined by the momentum relaxation rate. The majority of carriers undergoing ionization start out at the average thermal energy and drift to threshold, suffering a few phonon collisions en route. Solution of the rate equations yields an analytical expression for the ionization rates in terms of the threshold energies, the optical phonon mean free path and the phonon energy (the same
Figure 2.6. Comparison of ionization rates calculated by Baraff (dashed lines) and lucky drift model (solid lines), for different values of the threshold energy and phonon energy (after Ridley, 1983a).
parameters as for the Baraff theory). Figure 2.6 shows a comparison of lucky drift theory (solid lines) with the Baraff curves. Inclusion of a soft threshold into the lucky drift model (Marsland, 1987; Ridley, 1987) has allowed accurate fits to the experimental data for a number of materials.

The effect of band-structure on carrier transport and impact ionization has been studied using Monte Carlo simulations, for GaAs (Shichijo et al., 1981; Brennan and Hess, 1984) and InP (Osaka et al., 1986(a)). Realistic band-structures calculated by pseudopotential techniques (for electrons) or k·p theory (for holes) methods were used. The simulations show that at fields of \(10^4\) V·cm\(^{-1}\) the electrons transfer into the X and L valleys due to the higher density of states. Impact ionization at higher fields is caused by electrons which avoid energy relaxation, although they suffer several phonon collisions en route to threshold. For holes, impact ionization is predominantly due to holes scattered into the light-mass spin-orbit band.

2.3.7 Survey of experimental results.

This section summarizes the most reliable experimental results on the III-V compound semiconductors GaAs, InP, Ga\(_{1-x}\)In\(_x\)As\(_y\)P\(_{1-y}\), Al\(_{0.48}\)In\(_{0.52}\)As and Al\(_x\)Ga\(_{1-x}\)Sb. Details of the measurement techniques can be found in chapter 5.

**GaAs:**

Results in the literature for the ionization rates in GaAs illustrate the difficulty in achieving reproducible results. Early workers found \(\beta > \alpha\), but observed strong field dependence (Stillman, 1977), orientation dependence (Pearsall et al., 1978) and temperature dependence (Capasso et al., 1977) of the ratio. Ando and Kanbe (1981) found \(\alpha / \beta \approx 2\) by measuring the excess noise associated with the
A careful consideration of the experimental conditions necessary to reliably achieve pure single carrier-type injection was carried out by Bulman et al. (1983). Mixed injection by Franz-Keldysh electroabsorption of recombination radiation was specifically avoided using the device structure discussed in chapter 5. A large number of P⁺-N junctions of varying doping levels were measured, and very good agreement obtained between wafers and with Monte Carlo calculations. The results were also verified by noise measurements. \( \alpha/\beta \) varies between 2.5 to 1.3 for fields of 2.2 to \( 6.25 \times 10^5 \text{ V.cm}^{-1} \).

**InP:**

Cook et al. (1982) have similarly made a detailed study of ionization rates in InP over a wide range of fields. \( \beta/\alpha \) ranges from 4.0 to 1.3 for fields of \( 2.4 - 7.7 \times 10^5 \text{ V.cm}^{-1} \). A number of workers have studied orientation dependence of the ionization rates (Armiento and Groves, 1983; Tabatabaie et al., 1983, Osaka and Mikawa, 1985) and deduced that the ionization is essentially isotropic.

**Ga\(_{1-x}\)In\(_x\)As\(_y\)P\(_{1-y}\):**

Pearsall (1980) measured \( \alpha/\beta \approx 2 \) in the (100)-oriented ternary Ga\(_{0.47}\)In\(_{0.53}\)As. More recent results (Osaka et al., 1985) find a similar value of \( \alpha/\beta \) but with \( \alpha \) and \( \beta \) about an order of magnitude lower than Pearsall's values.

Osaka et al. (1985) have measured ionization rates for the quaternary (lattice matched to InP) at several intermediate compositions between InP and Ga\(_{0.47}\)In\(_{0.53}\)As. The value of \( \alpha/\beta \) increased monotonically with \( y \). The composition dependence was attributed to alloy scattering and to the decrease in intervalley scattering of electrons as the intervalley separation increases.
Measurements of the ternary alloy $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$ showed $\alpha > \beta$, with $\alpha/\beta$ from 2.5 to 3 for fields of 2.2 to $4.3 \times 10^5$ V·cm$^{-1}$ (Capasso et al., 1984b).

This material system is of particular interest for APD's due to the resonant enhancement of the hole impact ionization rate which occurs at a composition when the spin-orbit splitting is equal to the band-gap ($x \approx 0.065$). The threshold energy for holes is then equal to the band-gap. Hildebrand et al. (1980; 1981) measured $\beta/\alpha \approx 1$ for GaSb increasing to 20 at $\Delta/E_g = 1.02$. Similar resonant enhancement is found in GaSb at 77K (Zhingarev et al., 1980). The high ionization ratio achievable in this material system makes it promising for 1.55 µm photodetectors. However there are a number of as-yet unsolved material problems, including high surface leakage currents (Law et al., 1981).
CHAPTER 3.

PROPERTIES OF QUANTUM WELL STRUCTURES.

This chapter describes some properties of quantum well structures. Section 3.1 considers the general case of band line-up at a heterojunction. In section 3.2, the process of carrier transport across a single heterojunction and its influence on the properties of heterojunction photodiodes is considered. The band structure of thin double heterostructures, or quantum wells, is briefly discussed in section 3.3. Optical absorption in quantum wells is discussed in section 3.4. Section 3.5 summarises previous work on resonant tunnelling in double barrier devices. Section 3.6 discusses electrical transport across quantum wells for an electric field applied perpendicular to the layers. Finally, in section 3.7 the application of quantum well structures to avalanche photodiodes is discussed. Experimental and theoretical evidence for the enhancement of ionization rates by the band discontinuities is reviewed.

3.1 Heterojunctions and band discontinuities.

Developments in crystal growth technology, which are described in chapter 4, have made possible the growth of high quality semiconductor heterostructures - single crystals with abrupt, controlled compositional changes between different regions. Due to the epitaxial nature of device-quality crystal growth, this involves the growth of thin layers of different composition, maintaining the lattice-spacing of the substrate.

In this section we consider the band line-up at a heterojunction between two general materials, A and B, with direct energy gaps $E_A$ and
EB (EA < EB). The band-gap discontinuity (ΔEg = EB - EA) is distributed between the conduction band discontinuity (ΔEc) and the valence band discontinuity (ΔEv). Present theories are not able to accurately predict how the band-gap is shared between conduction and valence bands, and the band offsets must be measured in each material system (see chapter 4).

Heterojunctions are classified into four types depending on the arrangement of the discontinuities (see figure 3.1). In type I heterojunctions (figure 3.1(a)), electrons and holes are confined in the same layer, and

$$\Delta E_g = \Delta E_C + \Delta E_V. \quad (3.1)$$

In type II heterojunctions, the electrons and holes are confined in different layers and the energy gap difference is given by

$$\Delta E_g = |\Delta E_C - \Delta E_V|. \quad (3.2)$$

Type II heterostructures may be staggered (figure 3.1(b)), or misaligned (figure 3.1(c)) if ΔEc > EB. Type III interfaces are formed in II-VI compounds where one constituent is semimetallic (figure 3.1(d)). The heterojunctions considered in this thesis (Ga0.47In0.53As/InP, Ga0.47In0.53As/Al0.48In0.52As and GaSb/AlSb) are type I. It is clear that the presence of the heterojunctions, forming potential barriers with heights much larger than typical low-field carrier energies, will strongly influence the electrical properties of the semiconductor. In this thesis we are concerned with transport in a direction perpendicular to the heterointerfaces ("perpendicular" or "vertical" transport).

3.2 Carrier transport across a heterojunction.

This section considers carrier transport across a single heterojunction, and the process of carrier collection in heterojunction
Figure 3.1. Band line-up at a heterojunction interface.

(a) Type I,
(b) Type II (staggered),
(c) Type II (misaligned),
(d) Type III.
photodiodes.

A single heterojunction in a majority carrier device can act as a rectifier since flow of carriers from the narrow band-gap region is impeded by the potential barrier, \( \phi \). Energetic carriers ("hot" carriers heated by drift in the electric field) incident at the heterojunction with energy greater than \( \phi \) will cross the barrier, while those with lower energy pile up at the interface, forming a reservoir of trapped or stored carriers. These carriers can either be emitted across the barrier or can recombine. In steady state, the rate of arrival of electrons in the reservoir is balanced by the sum of the carrier recombination rate and the emission rate. If the recombination rate is negligible then the "trap" has no effect on the steady state current flow, although transient currents will be affected if the emission time is slow.

Trapping at the heterointerface decreases with applied field since the average carrier energy and the proportion of carriers with energy greater than \( \phi \) increases. Emission of trapped carriers by thermionic emission or tunnelling also increases with field (see section 3.6).

In a heterojunction photodiode, carriers are commonly photogenerated in a narrow-gap region and collected by a wide-gap layer, as in the SAM-APD's described below.

### 3.2.1 Quantum efficiency of Separate Absorption and Multiplication region APD's.

As mentioned in chapter 2, the leakage current in reverse bias Ga\(_{0.47}\)In\(_{0.53}\)As P-I-N diodes is large at high fields due to band-to-band tunnelling. This precludes their use as low-noise APD's. To fabricate long-wavelength APD's with low dark-current, Nishida et al. (1979) proposed a heterostructure device with Separate Absorption and
Multiplication regions (SAM-APD). Carrier multiplication takes place within a P⁺-N junction grown in wide band-gap InP. Thus tunnelling currents at breakdown are negligible. Absorption of long-wavelength light occurs in an adjacent layer of Ga₀.₄₇In₀.₅₃As which is not subject to the high electric field of the P⁺-N junction. The N-type Ga₀.₄₇In₀.₅₃As is grown adjacent to the N-type side of the P⁺-N junction so that holes are injected into the high-field region, which is the required configuration for low-noise multiplication (β > α for InP). This device has been successfully fabricated by a number of groups (for example: Diadiuk et al., 1980; Susa et al., 1981; Kim et al., 1981).

The field at the heterointerface determines the collection efficiency of holes photogenerated in the absorption layer. Thermionic emission and tunnelling rates of heavy holes across the valence-band discontinuity are low due to the high effective mass (m^*ₕ ≈ 0.5) and large potential barrier (ΔEᵥ ≈ 0.4 eV, see chapter 4). At low bias before punch-through of the depletion layer into the absorption region, the rate of recombination with majority electrons is high, whilst emission rates are low. Thus the internal quantum efficiency is low (η_{int} ≈ 0). As the depletion edge punches through the heterointerface, the recombination rate decreases due to depletion of majority carriers, whilst the emission rate is increased by carrier heating in the applied field, and η_{int} tends to unity.

The correct operation of a SAM-APD imposes quite severe requirements on the device parameters. To avoid tunnelling currents the electric field at the heterointerface must be < 1.5 x 10⁵ V-cm⁻¹, but must be sufficiently high that the rate of thermal emission of heavy holes over the valence band discontinuity exceeds the carrier recombination rate. At the same time the field in the P⁺-N junction
must be \( \gtrsim 4.5 \times 10^5 \text{ Vcm}^{-1} \) for avalanche multiplication for InP.
These considerations have been treated in detail by Forrest et al. (1983).

### 3.2.2 Pulse response of SAM-APD's.

The pile-up of holes at the valence-band discontinuity can lead to slow pulse response in some Ga\(_{0.47}\)In\(_{0.53}\)As/InP SAM-APD's, as reported by Forrest et al. (1982). The pulse response showed both a fast component, with a pulse width less than 200 ps at all biases, and a slow component, varying from 1 \(\mu\)s at 300 K and low bias to 450 ps at bias close to breakdown and \( T \approx 470 \) K. The fast component was attributed to hot carriers which were not trapped at the discontinuity. The temperature dependence of the slow component showed that emission over the barrier was thermally activated. The activation energy was both bias and intensity dependent. The bias dependence was explained by the presence of compositional grading at the heterointerface. By calculating the potential profile from the measured free carrier concentration profiles for various grading lengths, Forrest et al. were able to deduce firstly that the activation energy was equal to the well depth and secondly that the grading length was of the order of 150 Å in their LPE-grown samples. The decay time of the slow pulse response decreased at high intensity. This was explained by barrier lowering due to charge pile-up. The decay of the pulse was no longer exponential since the barrier height changes as the well discharges. Barrier lowering becomes important for current densities in excess of 0.5 A cm\(^{-2}\) (a current of \( \approx 70 \mu\)A for devices of area 1.3 x \(10^{-4}\) cm\(^2\)), or an optical power of \( \approx 0.15 \) mW. A conclusion of this work was that measurement of intrinsic response times in heterostructure photodiodes must be made under conditions of low light.
intensity and long pulses.

A number of methods have been used to overcome the problem of carrier pile-up in SAM-APD's. Forrest et al. (1982) suggested the use of a graded band-gap region between the absorption and multiplication regions. Campbell et al. (1983) demonstrated a high performance SAM-APD with a GaInAsF layer of intermediate band-gap between the absorption and multiplication regions. To avoid the difficulty of growing compositionally graded layers, Capasso et al. (1984d) used a graded-gap pseudo-quaternary superlattice, where the grading is obtained by varying the ratio of the Ga₀.₄₇In₀.₅₃As to InP layer width within a constant period.

An alternative material system to Ga₀.₄₇In₀.₅₃As/InP for the fabrication of SAM-APD's is Al₀.₄₈In₀.₅₂As/Ga₀.₄₇In₀.₅₃As. In this material system, the valence-band offset is considerably smaller (see chapter 4) leading to reduced pile-up of holes and absence of long tails in the photoresponse (Capasso et al., 1984a) without the need for intentional interface grading. However, later measurements of the ionization rates showed that α > β (Capasso et al., 1984b). Thus for injection of the carrier with the higher ionization rate, the complementary structure with a P-type absorbing region should be used.

3.3 Energy levels in quantum wells.

A thin layer of narrow band-gap material between layers of wide-gap material acts to confine carriers within it. If the width of the narrow-gap layer is smaller than the DeBroglie wavelength of the electron and the path length for scattering, then quantum size effects determine the optical and transport properties. These properties are strongly dependent on the thickness of the "quantum well" and on the applied electric field.
The energy of carriers in a quantum well is quantized in the \( z \) direction (perpendicular to the plane of the layers) due to confinement in the one dimensional potential well. For a carrier of mass \( m^e \) confined in an infinite rectangular well of width \( L_z \), the eigenfunctions are given by

\[
\psi_n = A \sin \left( \frac{n \pi z}{L_z} \right) \quad (n \text{ even})
\]

and

\[
\psi_n = B \cos \left( \frac{n \pi z}{L_z} \right) \quad (n \text{ odd})
\]

and the eigenvalues are

\[
E_n = \frac{\hbar^2}{2m^e} \left( \frac{n \pi}{L_z} \right)^2 \quad \text{for } n = 1, 2, 3, \ldots
\]

Thus there is a set of discrete bound states in the well.

For a finite potential well of depth \( V_0 \), the wavefunctions extend into the barrier layers with exponential solutions corresponding to evanescent tails. The energy levels in the well are lowered compared to the infinite well (Schiff, 1968).

For a semiconductor quantum well there are bound states due to confinement in the conduction-band (well depth = \( \Delta E_C \)) and valence-band (\( V_0 = \Delta E_V \)) potential wells. In the valence band there are sets of states corresponding to the light hole (\( m^e_{lh} \)), heavy hole (\( m^e_{hh} \)) and spin-orbit (\( m^e_{so} \)) bands. In a real semiconductor there are a number of differences from the "text-book" solution for a finite well. This includes the effective mass change at the interface, band non-parabolicity, and interactions between bands which are close in energy. Real semiconductors have been treated using effective mass theories (Bastard, 1984; White and Sham, 1981; Altarelli, 1983) and tight-binding models (Chang and Schulman, 1982). In the envelope function approach (Bastard and Brum, 1986) the periodic parts of the Bloch
functions at the band-edges are assumed to be the same in the A and B layers (with energy gaps $E_A$ and $E_B$ and spin-orbit interactions $\Delta_A$ and $\Delta_B$), since the band-structures of the materials A and B are generally similar. Only the bands close to the gap ($\Gamma_6$, $\Gamma_7$ and $\Gamma_8$) are considered. The macroscopic boundary conditions satisfied by the envelope wave function are continuity of the wave function and the probability current, which requires that the factor $(1/m^*)d\psi/dz$ is continuous at the interface. For light-particle states, a three-band Kane model is generalised to the case of quantum wells. The solutions are

$$\frac{k_w}{m(A)(E)} \tan \left[ \frac{k_w}{2} \right] = \frac{k_B}{m(B)(E)} \hspace{1cm} \text{(even states)} \hspace{2cm} (3.5)$$

$$\frac{k_w}{m(A)(E)} \cotan \left[ \frac{k_w}{2} \right] = -\frac{k_B}{m(B)(E)} \hspace{1cm} \text{(odd states)}$$

where

$$K_B = \sqrt{\frac{2m^*(E-V_0)}{\hbar}} \quad \text{and} \quad K_w = \sqrt{\frac{2m^*E}{\hbar}}. \hspace{2cm} (3.6)$$

The energy-dependent effective mass $m(E)$ takes into account the non-parabolicity, and can be calculated from $k_p$ theory (Kane, 1957). When the non-parabolicity is small (i.e. in the limit of large $E_g$ and $\Delta$) and for the case of heavy holes, the mass reduces to the zone-centre band-edge effective mass.

The in-plane dispersion relation for the conduction sub-bands are found from solutions of the three dimensional Hamiltonian. The effective mass mismatch causes a weak dependence of $E_n$ on $k_{xy}$, which can usually be neglected. The total energy is thus
For each value of $n$ there is a two-dimensional energy band in the $k_{xy}$ plane, governed by a bulk-like dispersion relation. The carriers confined in the well are often referred to as a 2-dimensional electron gas. The density of states in two dimensions is independent of energy, giving rise to a step-like density of states for a series of bands.

For the $T_{8}$ valence sub-bands, which are degenerate at the zone centre, the off-diagonal terms in the valence band Hamiltonian introduce coupling between the light-hole and heavy-hole states away from the zone centre. If only diagonal terms are considered, the subbands show mass-reversal: the "heavy-hole" band shows a light in-plane mass and the "light-hole" band shows a heavy in-plane mass, with the light and heavy-hole bands crossing at some value of $k_{xy}$. The off-diagonal terms cause anticrossing and a mixing of the light and heavy natures of the bands. This effect is important in a number of situations, e.g. tunnelling of holes (see chapter 9). There is experimental evidence for band-mixing from luminescence studies in quantum wells (Miller et al., 1985b; Sooryakumar et al., 1984).

The energy levels in a quantum well under an applied electric field ($F$) have been calculated by a number of methods. For an infinitely-deep trapezoidal well (Miller et al., 1985a) the solutions of the Shroedinger equation are Airy functions, and the energy eigenfunctions can be found using series expansions. For the case of large fields, this reduces to the solution for an infinite triangular well

$$E_n = E_n + \frac{n^2}{2m^*} (k_{xy})^2. \quad (3.7)$$

The experimental energy levels in a quantum well under an applied electric field ($E_n$) have been calculated by a number of methods. For an infinitely-deep trapezoidal well (Miller et al., 1985a) the solutions of the Shroedinger equation are Airy functions, and the energy eigenfunctions can be found using series expansions. For the case of large fields, this reduces to the solution for an infinite triangular well

$$E_n = -\frac{|e|FL}{2} + \left[ \frac{3nFe\hbar(n-1/4)}{2\sqrt{2m^*}} \right]^{(2/3)} \quad (3.8)$$
Variational techniques have been used (Bastard et al., 1983) to find the energy eigenvalues and the wavefunctions in both infinite and finite quantum wells. The effect of the field is to push the electron and hole wavefunctions in opposite directions towards the edges of the well, resulting in a decrease in the spatial overlap. The heavy hole wavefunction is perturbed more due to its lower confinement energy.

A useful method for calculating the energy eigenvalues is the tunnelling resonance technique. This calculates the transmission resonances of a carrier tunnelling coherently through a double barrier structure. Effective mass discontinuities and non-parabolicities are easily included. This method is used to identify the resonant tunnelling transitions discussed in chapter 9.

3.4 Optical properties of quantum wells: excitons and electroabsorption

Optical absorption has been widely used to study quantum well structures. The step-like density of states is reproduced in the absorption spectra (Dingle, 1975), since allowed transitions only occur between conduction and valence sub-bands with $\Delta n = 0$. Photogenerated electron-hole pairs interact via a Coulomb force. Thus the electron and hole are described in terms of excitons (two-particle states consisting of a bound electron-hole pair) rather than single-particle states. Excitons can be pictured as hydrogenic states with a binding energy determined by the classical radius. As in the case of bulk semiconductors at low temperatures, excitons dominate the absorption close to sub-band edges. The absorption spectra thus consists of exciton peaks superimposed on the step-like structure. For small well widths, the peaks corresponding to each inter-subband transition split into two, corresponding to transitions from the light
and heavy-hole valence bands.

In a quantum well, unlike in bulk material, strong exciton absorption is observed at room temperature. The effect of confinement in a quantum well of width smaller than the exciton radius in the bulk material is to decrease the size of the exciton and thus increase its binding energy (Dingle et al., 1975). The electron-hole pair can complete a classical orbit before being ionized by an optical phonon, thus sharp resonances are observed even at room temperatures.

The effect of an electric field parallel to the plane of the layers is similar to the bulk case. The dominant electroabsorption process is Frank-Keldysh electroabsorption (Franz, 1958; Keldysh, 1958). Excitonic resonances shift to lower energies with field due to the hydrogenic Stark shift and broaden due to the decrease in the lifetime for field ionization. Ionization occurs at low fields, and the Stark shift is generally small.

The application of an electric field perpendicular to the layers has quite different effects (Chemla and Miller, 1987). As the field is increased, the intersubband absorption edges move to lower energies due to perturbation of the rectangular quantum wells by the applied field. The absorption strength of the lowest energy transitions decreases as the field increases, due to decreasing spatial overlap of the electron and hole wavefunctions. In an applied field, the electron and hole wavefunctions are no longer orthogonal, and these transitions become weakly allowed.

Large shifts are observed in the excitonic peaks with field perpendicular to the layers, and the peaks are present even at fields in excess of $10^5 \text{V-cm}^{-1}$. This effect was termed the quantum-confined Stark effect (QCSE) by Miller et al. (1984b). The excitons are observed at high fields because field ionization is inhibited by the barriers.
The tunnelling time is usually \( \gg 1 \) ps, while the exciton dissociation time from phonon collisions is \( \approx 300 \) fs at room temperature (Knox et al., 1985). Figure 3.2 (a) shows the room temperature absorption data for a Ga_{0.47}In_{0.53}As/InP multiple quantum well with 100 Å layers, after Bar-Joseph et al. (1987). A strong shift of the \( n=1 \) heavy-hole and light-hole peaks is observed as the field is increased to \( 1.5 \times 10^5 \) V-cm\(^{-1}\). The change in the electron and hole wavefunctions is indicated in figure 3.2 (b) and (c).

Electroabsorption is an important effect in quantum wells both for understanding the physical properties and for implementing new devices. Fast-modulated QW lasers (Yamanishi and Suemune, 1983) have been proposed, and electro-optic modulators (Wood et al., 1984) and nonlinear optical devices (Self-Electro-optic Effect Devices or SEEDS, Millar et al., 1984a) have been demonstrated.

3.5 Electrical properties of quantum wells: resonant tunnelling.

Using the techniques of epitaxial growth of semiconductor heterostructures, quantum mechanical tunnelling through potential barriers of controlled size and shape can be studied experimentally. Tsu and Esaki (1973) calculated the current-voltage characteristics of multiple barrier devices from the resonant transmission coefficient as a function of energy. Resonant tunnelling of electrons in double barriers was observed experimentally by Chang et al. (1974).

Further improvements in material quality led to the observation of room temperature negative differential resistance (Shewchuck et al., 1985), quantum well oscillators at frequencies up to 18 GHz (Sollner et al., 1984), resonant tunnelling at frequencies up to 2.5 THz (Sollner et al., 1983) and resonant tunnelling of holes (Mendez et al., 1985).

Resonant tunnelling through a double barrier is shown in figure
Figure 3.2. Quantum confined Stark effect in a Ga$_{0.47}$In$_{0.53}$As/InP 100 Å quantum well (after Bar-Joseph et al., 1987)

(a) absorption at zero field (solid line) and $10^5$ V·cm$^{-1}$ (dashed line)
(b) electron and hole n=1 energy levels and wavefunctions at zero field
(c) energy levels and wavefunctions at $10^5$ V·cm$^{-1}$ (not to scale).
3.3. N-type contacts are formed on either side of the undoped double barrier. Energy and lateral momentum (perpendicular to the field direction) are conserved in resonant tunnelling. As the field across the double barrier region increases, the Fermi sea in the emitter lines up with the quasi-bound level in the well, and resonant tunnelling occurs. The peak in the tunnelling current occurs when the emitter conduction-band edge lines up with the bound level. If the bias is increased further, the quasi-bound level is below the emitter conduction band-edge and resonant tunnelling is not possible. The current decreases, giving rise to a negative differential resistance (NDR). At higher bias, the current increases again due to Fowler-Nordheim tunnelling through the first barrier. The physics of resonant tunnelling in a double barrier has been reviewed by Ricco and Azbel (1984), Capasso et al. (1986c) and Mendez et al. (1987).

Ricco and Azbel (1984) considered Fabry-Perot-type resonance in double barriers, in which the phase coherence of the electrons is maintained throughout the structure. Under resonant conditions the wavefunction is peaked in the quantum well and interferes destructively with reflected waves and constructively with transmitted waves. The build-up of electrons in the well at resonance has two important consequences. Firstly, the charge accumulated will alter the potential profile of the barriers, which therefore must be calculated self-consistently. Secondly, the time taken for charge to accumulate in the well at the onset of resonance leads to an upper limit for the frequency of a resonant tunnelling oscillator.

Luryi (1985) has proposed an alternative mechanism to coherent resonant tunnelling to account for the observation of NDR at very high frequencies. For injection into a 2-dimensional state, NDR can arise solely through conservation of energy and lateral momentum, and does
Figure 3.3. Resonant tunnelling in a double barrier device.
not necessarily imply coherent tunnelling throughout the whole structure. Scattering in the double barrier will reduce phase coherence. If the scattering time is less than the time taken to build up the resonant mode in the well, then tunnelling proceeds sequentially through each of the barriers. Weil and Vinter (1987) have shown that resonant tunnelling and sequential tunnelling lead to the same predictions for the d.c. current in a double barrier device.

Results on resonant tunnelling in multiple quantum well structures are presented in chapter 9.

3.6 Carrier transport across a quantum well.

The performance of devices such as quantum well APD's is determined by carrier transport across the well. This is dependent on the recombination rate and the emission rate from the well, for carriers generated in the well. For carriers entering the well from the barrier layers, the capture probability is also important.

The collection efficiency for carriers generated in the well is given by

\[
\frac{\tau_e^{-1}}{\tau_e^{-1} + \tau_r^{-1}}
\]

where \(\tau_e^{-1}\) is the emission rate and \(\tau_r^{-1}\) is the recombination rate. Typical radiative recombination rates in bulk materials are \(\approx 1\) ns. The radiation rate is increased in quantum wells due to increased overlap of the electron and hole wavefunctions (Göbel et al., 1983). The field dependence of the luminescence lifetime has been studied by time-resolved photoluminescence, and depends strongly on the well width. For narrow wells (\(L_w \lesssim 100\) Å) the luminescence quenches at fields \(\lesssim 10^4\) V·cm\(^{-1}\) (Mendez et al., 1982) due to field-induced tunnelling
out of the well (Kash et al., 1985). For wider wells (Polland et al., 1985), the luminescence lifetime increases with field due to the reduced spatial overlap of the electron and hole wavefunctions (Bastard et al., 1983). Thus for wide wells the collection efficiency increases with field due to both the increasing emission rate from the well and the decreasing radiative recombination rate.

Emission from a quantum well can occur through the following mechanisms (see figure 3.4).

(a). At low temperatures and for thin barriers, tunnelling through the barriers dominates. The probability for tunnelling ($P_T$) through a barrier $V(z)$ is given in the Wentzel-Kramers-Brillouin (WKB) semiclassical approximation (Bohm, 1951) by

$$P_T \approx \exp \left[ -2 \int k(z) \, dz \right] \tag{3.10}$$

where

$$k(z) = \sqrt{\frac{2m^*}{\hbar^2}} \frac{V(z) - E}{E} .$$

This expression does not include resonance effects due to quantum mechanical reflection at the interfaces. For a square barrier of width $L_B$ and height $\phi$ (e.g. a quantum well at zero field), the tunnelling probability becomes

$$P_T \approx \exp \left[ \frac{-2}{\hbar^2} \sqrt{\frac{2m^*}{\hbar^2}} \left( \frac{\phi - E}{L_B} \right) \right] \tag{3.11}$$

Under a sufficiently high field the barrier is triangular and the tunnelling probability is

$$P_T \approx \exp \left[ \frac{-\phi}{3 \hbar \sqrt{(2m^*)^{3/2}}} \right] \tag{3.12}$$
Figure 3.4. Emission from a quantum well:

(a) tunneling,
(b) phonon-assisted tunnelling,
(c) thermionic emission,
(d) ionization by scattering with hot carrier.
The tunnelling probability is strongly dependent on the field. The tunnelling rate can be estimated (Masumoto et al., 1986) from the product of the tunnelling probability and an attempt frequency \( \frac{\pi \hbar}{2m^*L_z^2} \) for an infinite square well.

(b). As the lattice temperature increases, phonon-assisted tunnelling occurs. This can also be thought of as an effective reduction in the barrier height for thermoionic emission due to the high probability for tunnelling through the triangular barrier at high energies and high fields.

(c). At high temperature, or for wide barriers, thermoionic emission over the barriers dominates. At low fields, carriers in the well are in thermal equilibrium with the lattice. The probability for thermally-activated emission over the barrier is given by \( \exp \left( -\frac{\phi}{kT} \right) \) where \( T \) is the lattice temperature. In the barrier layers, carriers are heated by the field to temperatures much higher than the lattice temperature. Direct heating by the field of carriers in bound states in the well does not occur since the momentum in the field direction is quantised. Carriers in the well may be heated by Coulomb interaction with hot carriers from the barrier layers. Due to the quantisation of momentum in the field direction, only momentum exchange perpendicular to the field causes carrier heating. At high fields where carriers drifting in the barriers can have large perpendicular momentum, and at high carrier densities, this may lead to an increased emission rate.

(d). If the carriers in the barrier layers have sufficient energy they can impact ionize bound carriers across the band-edge discontinuity. This effect is somewhat similar to impact ionization of deep levels, and is discussed in more detail in chapter 8.

The capture rate into the well is also important for transport across the well. Shichijo et al. (1978) showed that the efficiency for
scattering into the well by LO phonon emission is low for narrow wells
($L_w \leq 100 \text{ Å}$). Quantum mechanical calculations in quantum wells
(Brum and Bastard, 1986) and superlattices (Babiker and Ridley, 1986)
indicate capture rates in the range $10^{10} - 10^{12} \text{s}^{-1}$, with oscillatory
structure related to resonant capture by virtual states. Under high
applied fields, the capture time might be expected to be slower than
the transit time across the well ($\approx 10^{-13} \text{s}$), leading to low capture
efficiencies.

The quantum efficiency and pulse response of the MQW APD's
reported in chapters 6 and 7 are interpreted on the basis of the above
discussion.

3.7 Enhancement of ionization rate ratios in multilayer APD's.

The first proposal for use of multiple quantum wells in APD's was
made by Chin et al. (1980). Two mechanisms were described which would
alter the ionization rates. The first concerns the effect of the
potential steps at the heterojunction interfaces on the threshold
energies for ionization. The second concerns the optical phonon
scattering rates for electrons and holes in quantum wells.

Carriers accelerated by the applied field in the barrier layers
and arriving at the heterojunction interface will experience a
discontinuous increase in their energy relative to the band edge (by an
amount equal to the band offset). These heated carriers will require
less field heating to reach the ionization threshold energy. Thus the
effective threshold energy for impact ionization in the wells, for
carriers whose drift starts out in the barriers, is $E_{th}^{(e)} - \Delta E_c$
for electrons and $E_{th}^{(e)} - \Delta E_V$ for holes (see figure 3.5). Since
the ionization rates depend exponentially on the threshold energies,
the ratio of ionization rates may be substantially enhanced if $\Delta E_c$
Figure 3.5. Ionization rates and threshold energies:

(a) in bulk material

(b) in a quantum well, for carriers starting out in the barriers.
and $\Delta E_V$ differ significantly.

The carrier ionization rate may be enhanced even though the net energy gain from the periodic field is zero over one period. This is because impact ionization is a non-linear process with an energy threshold. The rate is enhanced in the presence of a periodic field since it is an exponent of the field rather than the field itself that is averaged over a period.

It is clear that the enhancement will be greatest at low fields, where the heating by the potential step is most significant compared to heating by the field, and ionization in the barrier layers is negligible.

The ionization rates for the structure as a whole are expressed as a weighted average of the rates in the wells and barriers:

$$\alpha = \frac{\alpha_W L_W + \alpha_PL_B}{L_W + L_B}, \quad (3.13)$$

Note that $\alpha_W$ is not the same as the rate in the bulk material due to enhancement by the potential step. Thus $\alpha$ may be higher or lower than $\alpha_W$ depending on the degree of enhancement and layer thicknesses.

The first MQW APD's were demonstrated by Capasso et al. (1982). The Al$_{0.45}$Ga$_{0.55}$As/GaAs structure had 25 periods of 450 Å wells and 550 Å barriers, placed in the I region of a P$^+$-I-N$^+$ structure. $g$ and $\alpha$ were calculated from measured multiplication characteristics for pure electron and pure hole injection (see chapter 5). $g$ was virtually the same as in the bulk material whilst $\alpha$ was substantially enhanced, giving an ionization rate ratio of 8 compared to 2 in bulk GaAs (at a field of $2.5 \times 10^5$ V-cm$^{-1}$). Similar rates were measured by Juang et al. (1986) in an Al$_{0.4}$Ga$_{0.6}$As/GaAs MQW APD with 570 Å and 424 Å barrier and well widths.

The second mechanism discussed by Chin et al. applies mainly to
narrow well structures. Shichijo et al. (1978) observed that the carrier collection efficiency of a quantum well reduces for $L_W \leq 100 \text{ Å}$. As the well width approaches the LO phonon mean free path it becomes less likely that a carrier will be scattered whilst in the well layer. This can lead to an enhancement of the ionization rates since all ionization events will be initiated by carriers starting out in the barrier layers. Holes are scattered more often due to their higher effective mass, and thus may be more efficiently collected by the wells.

A study of superlattice APD’s with narrow well widths ($\leq 100 \text{ Å}$) was made by Juang et al. (1985). They found that the ionization rate for electrons remained fairly constant at close to its value in the bulk, whilst the hole ionization rate was strongly dependent on the well width.

Impact ionization in quantum wells has been studied theoretically by lucky drift theory and by Monte Carlo simulations. Ridley (1985) used the simplifying concept of an energy- and field-independent energy relaxation length ($l_E$). The finite energy relaxation length means that a carrier reaching the threshold energy in the barrier layer will remain there for a length $l_E$ before relaxing. Thus carriers may reach threshold, and impact ionize within a well, over a distance $L_W + l_E$. Thus the ionization processes are determined principally by the dimensions of the well and barrier layer thicknesses relative to the energy relaxation length.

The threshold energy for carriers starting lucky drift from a barrier layer and ionizing in the well was taken as $E_{th} - AE$, while that for a carrier starting out in the well layer was $E_{th}$ (ionization in the barrier layers was assumed to be negligible). Four modes of operation were evident, depending on whether $L_W$ and $L_B$ are less than or
greater than \( l_E \).

(1) Narrow well, narrow barrier. If \( L_B < l_E \), ionization can occur in the well wherever threshold is reached. If \( L_W < l_E \), no thermalization occurs in the well and all carriers reaching threshold originate in the barrier layers. This is therefore the optimum case for ionization rate enhancement by the band offsets.

(2) Narrow well, wide barrier. Again no thermalisation occurs in the well so only carriers drifting from the barrier layers are important. However, only carriers reaching threshold in the region \( L_W + l_E \) will ionize in the well, so the ionization rate is reduced by a factor \( (L_W + l_E)/L \) or \( \approx l_E/L_E \).

(3) Narrow barrier, wide well. In this case carriers can reach threshold starting from the barrier layers or from the well layers. The ionization rate ratio is reduced compared to the case for a narrow well due to the contribution from carriers starting from the well layers, which do not see the effect of the potential step.

(4) Wide barrier, wide well. In this case, the rates are reduced by geometrical factors (which are field dependent) governed by the spatial regions from which a carrier can undergo lucky drift and impact ionize in a well. This leads to field-dependent oscillations in the ionization rates at low fields.

An analysis of the structure of Capasso et al. using the above lucky drift analysis and including a soft threshold for ionisation gave good agreement with the experimental data for a conduction band offset of \( \Delta E_C = 0.60 \Delta E_g \) (Ridley, 1987).

Brennan et al. (1985) applied Monte Carlo techniques to a GaAs/AlGaAs structure with 500 Å layer widths. The pseudopotential band-structure for GaAs was used for both materials and the heterointerfaces were treated simply as potential steps. The
enhancement of the electron ionization rate was strongly dependent on the value of the conduction band offset. Close agreement with the values of Capasso et al. (1982) was found for a conduction band offset of $\Delta E_c = 0.75 \Delta E_g$ - rather higher than observed experimentally (see chapter 4). A number of significant features emerged from the simulation. Virtually all the ionization events occurred in the well layers, as expected from the threshold energies. The ionization rates depended exponentially on the size of the potential step, which supports the model for the effective reduction in the threshold energy. The electrons travelled a mean distance of the order of 100 Å in the well before ionizing. As the step size was increased, the distance travelled before ionization decreased, implying that the carrier gains energy from the step quasi-ballistically. The variation of the ionization rates with well width was also studied (Brennan, 1985). The rate for electrons in the well layers increased rapidly with decreasing well thickness, in agreement with lucky drift analysis. The variation in the hole rate was much smaller, since it was only slightly enhanced from the bulk value.

For layer widths of $\leq 100$ Å, quantum size effects can influence the ionization rates. The increasing ground state energy for narrow wells both increases the threshold energy and decreases the energy received from the potential step, thus reducing any enhancement due to the periodic structure. Quantum mechanical reflections at the interfaces may have significant effects on the mean free paths.

A number of other multilayer APD structures are reviewed by Capasso (1985). MQW APD's with compositionally graded interfaces (figure 3.6(a)) were proposed (Capasso et al., 1983) to reduce pile-up of electrons in the wells. In the the Staircase APD (Williams et al., 1982) (figure 3.6(b)) the entire threshold energy comes from the
Figure 3.6 (a). Graded-well MQW APD.

(b). Staircase APD.
heterojunction discontinuity. Since the hole ionization rate is negligible, this results in single carrier-type multiplication at low applied fields. The excess noise factor is low due to both the single carrier-type multiplication and the spatial localization of the ionization events. Each step simulates the behaviour of a dynode in a photomultiplier tube, hence the term "solid state photomultiplier". Note that the implementation of this device requires a material system in which $\Delta E_c \geq E_{\text{th}}(e)$, and $\Delta E_V$ is low. (Ga,Al)(As,Sb) and II-VI compounds such as (Hg,Cd)Te are possible candidates for staircase APD's.

Results on long-wavelength MQW APD's are reported in chapter 6, 7 and 8.
Semiconductors formed from combinations of the group III (Ga, Al, In) and group V (As, P, Sb) elements are important for optoelectronics applications. Section 4.1 of this chapter describes some of the properties of these materials. Section 4.2 briefly reviews properties of lattice-matched III-V heterostructures. The growth of such heterostructures by a number of different techniques is discussed in section 4.3.

### 4.1 Properties of III-V Semiconductors

This section describes properties of III-V binary, ternary and quaternary alloy semiconductors. Some relevant parameters (lattice constant, static dielectric constant, band-gap at the Γ-point, spin-orbit interaction energy and effective masses) are shown in Table 4.1.

#### 4.1.1 The binary III-Vs

The binary III-Vs form a class of semiconductors with band-gaps in the range 0.18 - 2.42 eV. The most well-characterized binaries, GaAs and InP, have band-gaps of 1.43 and 1.35 eV at room temperature respectively, and thus are not sensitive in the near-infrared. GaAs is important commercially for high speed applications due to its high electron drift velocity at low fields. The principle importance of InP is for high quality substrates for alloy semiconductors and heterojunction systems. GaSb has a room temperature band-gap of 0.73 eV and is thus of interest for photodiodes operating at 1.55 μm. However
Table I. Properties of some III-V semiconductors.

<table>
<thead>
<tr>
<th>Material</th>
<th>GaAs</th>
<th>AlAs</th>
<th>InAs</th>
<th>InP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$ (Å)</td>
<td>5.6533 (a)</td>
<td>5.6611 (b)</td>
<td>6.0584 (c)</td>
<td>5.8688 (c)</td>
</tr>
<tr>
<td>$\varepsilon_s$ ($\varepsilon_0$)</td>
<td>13.2 (d)</td>
<td>10.9 (e)</td>
<td>14.55 (f)</td>
<td>12.35 (g)</td>
</tr>
<tr>
<td>$E_g$ (eV)</td>
<td>1.43 (h)</td>
<td>2.16 (i)</td>
<td>0.36 (j)</td>
<td>1.35 (k)</td>
</tr>
<tr>
<td>$\Delta$ (eV)</td>
<td>0.341 (m)</td>
<td>0.29 (o)</td>
<td>0.38 (p)</td>
<td>0.11 (m)</td>
</tr>
<tr>
<td>$m_e$ ($m_0$)</td>
<td>0.0665 (q)</td>
<td>0.15 (b)</td>
<td>0.023 (j)</td>
<td>0.0803 (s)</td>
</tr>
<tr>
<td>$m_{lh}$ ($m_0$)</td>
<td>0.094 (q)</td>
<td>0.22 (o)</td>
<td>0.024 (j)</td>
<td>0.12 (m)</td>
</tr>
<tr>
<td>$m_{hh}$ ($m_0$)</td>
<td>0.34 (q)</td>
<td>0.49 (b)</td>
<td>0.40 (u)</td>
<td>0.45 (m)</td>
</tr>
<tr>
<td>$m_{s-o}$ ($m_0$)</td>
<td>0.154 (v)</td>
<td>0.15 (r)</td>
<td>0.14 (p)</td>
<td>0.21 (m)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Material</th>
<th>Ga$<em>{0.47}$In$</em>{0.53}$As</th>
<th>Al$<em>{0.48}$In$</em>{0.52}$As</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_s$ ($\varepsilon_0$)</td>
<td>13.9 (w)</td>
<td>12.4 (w)</td>
</tr>
<tr>
<td>$E_g$ (eV)</td>
<td>0.73 (x)</td>
<td>1.46 (x)</td>
</tr>
<tr>
<td>$\Delta$ (eV)</td>
<td>0.36 (w)</td>
<td>0.34 (w)</td>
</tr>
<tr>
<td>$m_e$ ($m_0$)</td>
<td>0.0427 (x)</td>
<td>0.075 (x)</td>
</tr>
<tr>
<td>$m_{lh}$ ($m_0$)</td>
<td>0.0503 (y)</td>
<td>0.086 (w)</td>
</tr>
<tr>
<td>$m_{hh}$ ($m_0$)</td>
<td>0.465 (y)</td>
<td>0.57 (w)</td>
</tr>
<tr>
<td>$m_{s-o}$ ($m_0$)</td>
<td>0.118 (w)</td>
<td>0.16 (w)</td>
</tr>
</tbody>
</table>
Table 4.1 continued: references.

(u) Assumed the same as for InSb.
(w) Linear interpolation between values for binaries.
4.1.2 The ternary alloys and lattice matching.

The range of available semiconductors is extended by making alloys of binary semiconductors. Alloy semiconductors are generally grown on binary substrates due to the availability of high quality bulk crystals. To grow single crystal epitaxial layers of particular alloy composition on a substrate, without generating large quantities of dislocations, the lattice constant of the alloy must be the same as that of the substrate. This is referred to as lattice-matching. For ternary alloys (solid solutions of two binary compounds with a common ion) there is generally a one-to-one relationship between the lattice constant and the band-gap. This allows only a few compositions to be grown on available substrates, and thus a limited range of band-gaps.

Figure 4.1 shows the lattice constant and energy gap for a number of III-V materials. The lines joining the binary compounds give the ternary energy gap and lattice constant. The lattice constant of grown layers is determined experimentally from X-ray rocking curves.

The band structure properties of a semiconductor alloy depend on the lattice constant (which is fixed in a lattice-matched system) and local fluctuations in alloy composition (alloy disorder). The alloy disorder is small on the scale of the free-electron radius (≈ 100 Å), thus the disorder does not cause "smearing" of the band-structure. Disorder on the scale of the lattice constant (≈ 5 Å) is observed in binary vibrational modes in Raman scattering (Pinczuk et al., 1978).

(Al,Ga)As:

The alloy of AlAs and GaAs is written as (Al,Ga)As or AlₙGa₁₋ₙAs where x is the mole fraction of AlAs. Since the atomic sizes of Al and Ga are very similar, the ternary compounds (Al,Ga)As, (Al,Ga)P and
Figure 4.1. Lattice constant and energy band-gap of III-V semiconductors.
(Al,Ga)Sb have almost constant lattice parameter with varying band-gap. (Al,Ga)As in particular has been thoroughly investigated due to the existence of high quality GaAs substrates. The room temperature band-gap varies from 1.43 eV (GaAs) to 2.16 eV (AlAs). Since GaAs is a direct band-gap material whilst AlAs is indirect, there is a transition in Al\textsubscript{x}Ga\textsubscript{1-x}As which occurs at a composition of x \approx 0.45. Graded band-gap devices with arbitrary band-edge profiles can be grown simply by changing the Al mole fraction.

(In,Ga)As:

(In,Ga)As has a band-gap of 0.73 eV at room temperature when lattice matched to InP. Precise control of the mole fraction is required during growth to produce lattice matching, due to the large difference in lattice constants of InAs and GaAs. The mismatch is expressed as

\[
\frac{a_0(\text{In}_{x}\text{Ga}_{1-x}\text{As}) - a_0(\text{InP})}{a_0(\text{InP})},
\]

where \(a_0\) is the lattice constant. Lattice matching to the InP substrate occurs at x = 0.53. A deviation of 1% from x = 0.53 causes a mismatch of 7 \times 10^{-4} (Cheng et al., 1981). For device-quality layers, the mismatch should be \(10^{-3}\) or less.

Electron (Nicholas et al., 1979) and hole (Alavi et al., 1980) effective masses are shown in Table 4.1. The temperature-dependence of the band-gap at the \(\Gamma\)-point has been given by Forrest et al., (1983) as

\[
E_g(T) = E_g(0) - \omega T,
\]

where \(E_g(0) = 0.822\) eV and \(\omega = 3.00 \times 10^{-4}\) eV.K\(^{-1}\), accurate to \(\pm 0.5\%\) for 125 < T < 450 K.
4.1.3 Quaternary alloys.

Adding a fourth component to ternaries removes the one-to-one correspondence of band-gap and lattice constant. Thus a continuous range of band-gap with a fixed lattice constant (matched to the substrate) can be achieved.

(Ga,In)(As,P):

The system (Ga,In)(As,P) can be matched to two available substrates: GaAs over a band-gap range of 1.42 - 1.94 eV which is similar to that for (Al,Ga)As, and InP over 0.73 - 1.35 eV which includes the 1.3 μm and 1.55 μm low loss windows for silica fibres. The latter system has been extensively studied for use in double-heterostructure and quantum well lasers and in photodetectors (for a review, see Pearsall, 1982).

Lattice matching to InP occurs for \( y \approx 2.2x \). The composition-dependence of the band-gap at the Γ-point is given by Pearsall (1982):

\[
E_g(\text{eV}) = 1.35 - 0.775y + 0.149y^2 \quad (T = 295 \text{ K})
\]

and

\[
E_g(\text{eV}) = 1.425 - 0.7668y + 0.149y^2 \quad (T = 4 \text{ K}).
\]

The electron effective mass measured by Nicholas et al. (1979) varies as

\[
m_e^*/m_0 = 0.060 - 0.039y.
\]

Other parameters are often estimated by linear interpolation between InP and Ga_{0.47}In_{0.53}As.

(Al,Ga,In)As:

The quaternary alloy Ga\(_x\)Al\(_y\)In\(_{1-x-y}\)As (lattice-matched to InP) offers some particular advantages for implementing optoelectronic devices. The band-gap covers a range 0.73 eV (Ga_{0.47}In_{0.53}As) to 1.47 eV (Al_{0.48}In_{0.52}As, at 300 K), or a wavelength range from 0.85 to
1.65 μm. Because of the similar atomic sizes for Ga and Al, growth of different compositions can be achieved simply by changing the Ga/Al ratio. Graded band-gap structures can be grown by compositional grading. Compositional grading in (Ga,In)(As,P) is much more difficult to achieve due to the necessity to change all fluxes to maintain lattice matching. The techniques for compositional grading are discussed under MBE growth of AlGaInAs.

The lattice mismatch $\Delta a/a_0$ for growth on InP substrates (Olego et al., 1982) is related to $x$ and $y$ by

$$x + y = 0.468 - 15(\Delta a/a_0) + 0.017y .$$  (4.6)

The lattice matching condition ($\Delta a/a_0 = 0$) for (Ga,In)As is thus $x = 0.468$ and $y = 0$, and for (Al,In)As is $x = 0$ and $y = 0.476$. If $x + y = 0.472$ throughout the grading, the lattice mismatch will be $\leq 3 \times 10^{-4}$, which is sufficient for device-quality layers. The band-gap at 295 K and the conduction-band effective mass were measured by Olego et al. (1982) with the results

$$E_g(eV) = (0.076\pm0.04) + (1.04\pm0.10)y + (0.87\pm0.13)y^2$$  (4.7)

$$m_e^* / m_0 = (0.0427\pm0.0015) + (0.0683\pm0.0007)y .$$  (4.8)

### 4.2 Heterojunction systems.

The physics and device applications of heterojunction systems were considered in chapter 3. The parameters that describe a heterojunction interface between two materials A and B are the band-gap discontinuity $\Delta E_g = E_g^A - E_g^B$, the fraction of the offset in the conduction band $Q_e = \Delta E_c/\Delta E_g$, the interface quality in terms of abruptness and uniformity and the carrier concentration at the interface which determines band-bending effects. This section reviews the measurements of the band-offsets in $Al_{x}Ga_{1-x}As/GaAs$, $Ga_{0.47}In_{0.53}As/InP$ and...
4.2.1 Measurement methods for band-offsets in GaAs/AlGaAs.

The experimental techniques for measurement of band offsets, and their limitations, are discussed using the example of the (100) GaAs/Al\(_x\)Ga\(_{1-x}\)As heterojunction. This is the most extensively characterized system, is intrinsically lattice matched, and can be grown with abrupt interfaces by techniques such as MBE.

Band offsets were measured by Dingle et al. (1974) via absorption spectra of single and multiple quantum wells of GaAs/Al\(_{0.2}\)Ga\(_{0.8}\)As. The position of the excitonic features for allowed transitions were fitted using a model with simple square wells, parabolic bands, and equal effective masses in the wells and barriers. A good fit to the data was obtained using \(Q_e = 0.85\). This method of calculating band offsets suffers a number of drawbacks. The excitonic absorption energies are strongly dependent on well width and effective masses, but are relatively insensitive to the value of the band-offsets. The binding energy of the excitons must be calculated. The energy levels may be affected by non-parabolicity for narrow well widths or by band-banding for wide wells. Despite this, the "85:15" rule was widely accepted and supported by early C-V (People et al., 1983) and photoluminescence measurements (Welch et al., 1984).

This value of the band-offsets was questioned by Miller et al. (1984). A fit to excitation luminescence data, for both allowed (\(\Delta n=0\)) and forbidden (\(\Delta n\neq0\)) transitions in both square and parabolic wells, led to a value of \(Q_e = 0.57\) using different values of the valence band effective masses from those used by Dingle et al. (see Table 4.1). Recent direct measurement (Wolford et al., 1986) of the valence band offset by pressure-dependent luminescence indicated
\[ \Delta E_V = (0.32 \pm 0.02) \Delta E_g. \]

Measurement of band offsets by C-V profiling through iso-type heterojunctions was described by Kroemer et al. (1980). Although Debye length smearing causes the measured carrier density to differ from the true free carrier density, conservation of the number of charge carriers and the moment of their distribution allows derivation of a value for the majority carrier heterojunction barrier height. This holds true independent of compositional grading (but not doping grading) at the interface (Kroemer, 1985). The conduction-band offset was found to be \( Q_e = 0.66 \). Other techniques which have been applied to this problem include inelastic light scattering in quantum wells (Menendez et al., 1986; \( Q_e = 0.69 \pm 0.03 \)), internal photoemission in heterojunction photodiodes (Haase et al., 1987; \( Q_e = 0.63 \pm 13 \% \)) and thermally activated transport measurements (Batey et al., 1985; \( Q_e \approx 0.60 \)).

It can be seen from the data presented above that care is required in deducing values of the band offsets from experimental data, in particular from measurements of the energy of excitonic features. Although there is general consensus that \( Q_e \) is in the range 0.6 to 0.7 further work is required to establish the effect of extrinsic factors such as interface quality and presence of defects.

4.2.2 Band offsets in Ga\(_{0.47}\)In\(_{0.53}\)As/InP.

The band-offsets in Ga\(_{0.47}\)In\(_{0.53}\)As/InP have received less attention. High quality quantum well structures for the measurement of band-offsets via optical methods have only recently become available. The transition energy of heavy holes in the quantum-well ground-state to the barrier layer conduction band-edge measured by absorption in MOCVD-grown samples (Skolnick et al., 1986) indicated \( Q_e \approx 0.38 \).
The CV-profiling technique has also been somewhat limited by the difficulty of achieving good Shottky barriers on InP. Forrest et al. (1984) have reported CV-profiling results using organic-on-inorganic diodes, which give higher barriers than metal-semiconductor Shottky junctions. Measurements over the entire range of (In,Ga)AsP/InP lattice-matched heterojunctions grown by LPE yielded $Q_e = (0.39 \pm 0.01)$.

Other results from C-V measurements are in disagreement. Ogura et al. (1983) measured $Q_e \approx 0.5 - 0.7$ in N-type heterojunctions, with $Q_e$ approaching zero as the temperature is reduced. Similar temperature-dependent results were obtained by Forrest and Kim (1982) and explained by them in terms of an interface trap. Measurements on P-type samples (Steiner et al. 1986), yielded $\Delta E_V = 0$.

The problem of capacitance measurements in heterojunctions has been addressed by Lang et al. (1987) following an analysis of admittance measurements on multiple quantum well samples (see chapter 9). The disparities mentioned above and the apparent temperature-dependence of the band-offsets were explained in terms of thermionic emission over a simple band-offset without defects. Data on complementary N- and P-type samples showed $Q_e = (0.42 \pm 0.02)$, in agreement with the results of Forrest et al.,(1984).

4.2.3 Band offsets in Al$_{0.48}$In$_{0.53}$As/Ga$_{0.47}$In$_{0.53}$As.

The difference in direct band-gap between Ga$_{0.47}$In$_{0.53}$As and Al$_{0.48}$In$_{0.52}$As is $\approx 0.7$ eV, leading to a large confining barrier for carriers in heterojunction devices. Measurements of the band offsets suggest that most of the discontinuity lies in the conduction band. This causes efficient confinement of electrons, and is significant for carrier pile-up effects and ionization rate ratios in heterojunction
People et al. (1983) used C-V profiling of isotype heterojunctions to find $\Delta E_C = (0.50 \pm 0.05)$ eV at 297K, corresponding to $Q_e = (0.71 \pm 0.07)$. However, in the same work the authors report a value of $Q_e = 0.88$ for GaAs/(Al,Ga)As, in contradiction with the current consensus (see above). Sugiyama et al. (1986) have measured the temperature-activated transport of electrons in rectangular barrier devices over the entire range of lattice-matched compositions. These results show that the offset is proportional to the Al content, $x$:

$$\Delta E_C = (0.53 \pm 0.05)x \text{ (eV)} \text{ for } 0 \leq x \leq 1.$$ 

This corresponds to $Q_e = 0.72 \pm 0.07$, in agreement with People et al.

Welch et al. (1984) studied the PL of single quantum wells, assuming the same effective masses in the wells and barriers but including electron non-parabolicity, and found $Q_e \approx 0.7$. Wagner et al. (1985) measured the luminescence of quantum wells down to 100 Å and fitted the data using $\Delta E_C = 0.5$ eV and the valence band effective masses for well and barrier layers shown in Table 4.1. Weiner et al. (1985) have fitted absorption data of high-quality Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW's exhibiting room temperature excitonic behaviour. Exciton binding energies were calculated using a variational approach, but only allowed transitions were observed. The results showed $\Delta E_C = 0.44$ eV and $\Delta E_V = 0.29$ eV, giving $Q_e = 0.60$ which is lower than the value obtained by other techniques.

Thus there appears to be some agreement that $\Delta E_C \approx 0.5$ eV, although a rigorous study of quantum well absorption (including forbidden transitions) along the lines of Miller et al. (1984) has yet to be made.
4.3 Growth techniques for heterostructures.

The performance of heterostructure devices is critically dependent on the growth of thin layers of uniform thickness and composition, with abrupt interfaces on the scale of an atomic monolayer. Until recently, molecular beam epitaxy (MBE) has been the predominant growth method for heterostructures due to its atomic layer resolution. Metal organic chemical vapour deposition (MOCVD) has been intensively investigated particularly for (GaIn)(AsP), and has greater prospects for scaling up for commercial production. Quantum well samples have also been produced by vapour phase epitaxy (VPE). The use of gas sources in MBE systems has blurred the distinction between vapour deposition and molecular beam techniques and offers some of the advantages of both.

The photoluminescence linewidth of single quantum wells has been widely used as an indicator of interface quality (Weisbuch et al., 1981). Emission energies and linewidths are plotted against well width in Figure 4.2 for Ga\textsubscript{0.47}In\textsubscript{0.53}As/InP quantum wells, showing recent data for various growth techniques (after Miller et al., 1986). The solid line in figure 4.2(b) shows the calculated linewidth assuming an interface width of one monolayer. An interface width of less than one monolayer, which corresponds to large domains with well thickness \( L_w \) with a small fraction of islands of thickness \( L_w + a_0/2 \), is achievable.

In this section, these growth techniques are described, concentrating on MBE which has been used to grow most of the samples studied in this thesis. The suitability of each technique for growing the different material systems described in the previous section is considered.
Figure 4.2. Comparison of optical properties of Ga$_{0.47}$In$_{0.53}$As/InP quantum wells grown by different growth techniques (after Miller et al., 1986).
4.3.1 Liquid Phase Epitaxy.

Liquid phase epitaxy (LPE) is a common epitaxial growth technique for III-V's due to its capability of producing high-quality material using a simple and inexpensive apparatus consisting of a horizontal furnace and sliding graphite boat. Since the growth occurs under near-equilibrium conditions there are few crystal defects or impurities. However, for the growth of multi-layer structures, there is a lack of controllability for thin layers, and wide interfaces due to melt-back. For these reasons, LPE has not been widely used for multilayer devices.

4.3.2 Vapour Phase Epitaxy.

Vapour phase epitaxy (VPE) can produce very high purity (Ga,In)(As,P) from a simple, low-cost apparatus (Olsen, 1982; Beuchet, 1985). Transport of the group V elements is via the hydrides (hydride VPE) or the trichlorides (chloride VPE). VPE has some important advantages over LPE: high uniformity in thickness and composition, flexible control of alloy composition and high throughput with the option to scale up for mass production. High purity material is achieved using chloride transport. However there are difficulties with the growth of compounds containing Al and Sb.

Hydride Transport:

HCl is used to transport group III metals, and arsine and phosphine provide the group V source. The growth process involves passing HCl gas over hot In or Ga metal to form metal monochlorides which combine with cracked AsH₃ or PH₃ at the heated wafer.

Control of the flux ratios for growth of (Ga,In)(As,P) is easy in hydride transport since the flow rates are all independently variable. However the purity of starting products is not as high as for chloride
transport, thus Ga\(_{0.47}\)In\(_{0.53}\)As layers grown by hydride transport typically have background dopings of low \(10^{15}\) cm\(^{-3}\).

**Chloride Transport:**

VPE using trichloride transport is safer than MOCVD and hydride-transport VPE since no arsine, phosphine or metal alkyls are used. Very low levels of background doping in InP and GaAs are achievable due to the high purity of the starting products.

For the growth of binary compounds, saturated two-phase or binary sources are used. For GaAs the process is as follows. Pure hydrogen flows through an AsCl\(_3\) bubbler; the mixture enters the source zone where AsCl\(_3\) decomposes to HCl and As vapours. The Ga metal source is saturated with As and a solid crust forms on the surface. HCl reacts with this crust forming GaCl and As vapours; this reaction is reversed at the substrate leading to deposition.

**Growth system:**

A typical single chamber VPE system consists of a horizontal quartz reactor tube, with a loading zone, source zone and a growth zone heated by resistance furnaces. Typical temperatures are 750-850°C for the source zone and 650-725°C for the growth zone. The trichlorides or HCl for transport of the group III elements enter the reactor chamber at one end and pass over the sources. They are mixed with a further stream of trichlorides or hydrides. P-type doping is achieved by passing H\(_2\) over a Zinc bucket, N-type by adding H\(_2\)S gas to the group V line. The gas mixture passes over the substrate and exhausts at the other end of the chamber. Flow rates are regulated with electronic mass flow controllers.

**Growth of multilayer systems:**

Growth of high purity multilayers requires special procedures to avoid contamination of one layer by the reactants of preceding layers.
This can occur due to deposition on the reactor walls or during the transition period while flows are equilibrating.

In multichamber reactors, the substrate is switched between different chambers for growth of the different layers. The flow of gas in each chamber remains constant allowing good compositional control with minimum cross-contamination of the layers. The manual switching time of the substrate between chambers is of the order of 1 s, allowing abrupt interfaces to be achieved if the growth rates are sufficiently slow. Komeno et al. (1983) have reported a dual growth-chamber reactor for trichloride VPE growth of Ga$_{0.47}$In$_{0.53}$As/InP. Hydride VPE was used by DiGiuseppe et al. (1983) in a similar dual reactor to produce a 10 period Ga$_{0.47}$In$_{0.53}$As/InP MQW with 100 Å layer thickness.

Ga$_{0.47}$In$_{0.53}$As/InP MQW structures were grown by Mattera et al. (1986) using a single chamber trichloride VPE process in a reactor previously developed by Cox et al. (1983, 1985). A H$_2$ counterflow system was used to prevent build-up of deposits on the reactor walls. Very high compositional and thickness uniformity were achieved using a horizontal rotating substrate holder. The sources were a large area In source (over which high purity HCl flowed) in the main tube and high purity undoped GaAs and InP sources in two auxiliary tubes (over which AsCl$_3$/H$_2$ and PCl$_3$/H$_2$ mixtures, respectively, were passed). Growth of alternate InP and Ga$_{0.47}$In$_{0.53}$As multilayers was achieved by controlling the flow rates of the trichlorides through the source tubes. Uniform high purity layers with background concentrations of $(1-3) \times 10^{14}$ cm$^{-3}$ in Ga$_{0.47}$In$_{0.53}$As and mid $10^{14}$ cm$^{-3}$ in InP were obtained. However the interface width is likely to be higher than for growth in multichamber reactors. Ga$_{0.47}$In$_{0.53}$As/InP MQW APD's grown by this technique are described in chapter 6.
4.3.3 Metal Organic Chemical Vapour Deposition (MOCVD).

Chemical vapour deposition using metal alkyls as vapour phase sources offers a number of advantages over conventional VPE. These include growth of Al and Sb alloys, and convenient growth of lattice-matched quaternary alloys and heterostructures (Hirz et al., 1982; Razeghi, 1985).

Sources.

Metal alkyls (in particular trimethyls, TM and triethyls, TE) are useful sources (Stringfellow, 1985) since they have high vapour pressures at storage temperatures (-20 to +20°C) and decompose to the III-V elements at convenient temperatures for growth. Metal alkyl sources for P and As do not readily decompose, and the hydrides are often used as group V sources. The electrical characteristics of the layers are dependent on the source purity: Carbon contamination can be a problem, particularly from the TM-Al source. The standard p-type dopant is Zn from DEZn. S from H₂S is used as n-type dopant for InP and (Ga,In)(As,P).

Apparatus:

MOCVD growth occurs in a vertical or horizontal quartz reaction tube. The substrate is typically mounted on a graphite susceptor heated by an RF coil.

The organometallic sources are kept in stainless steel bubblers. N₂ or H₂ is passed through the bubbler and the transported gas diluted by H₂. The flow rates are equilibrated prior to growth and regulated using electronic mass flow controllers.

Growth Process:

The hot susceptor catalyses decomposition of metal alkyls and hydrides into the elemental species. The growth rate is proportional to the flow of group III species, and not strongly dependent on
temperature and group V partial pressure. This shows that the rate
limiting step is diffusion of the group III alkyls across a "boundary
layer" adjacent to the growing surface.

Low pressure MOCVD (LP-MOCVD):

A particular problem for growth of InP and (Ga,In)(As,P) by MOCVD
at atmospheric pressure (AP-MOCVD) is low reaction efficiencies due to
the formation of parasitic polymers. This is eliminated in low pressure
MOCVD (Duchemin et al., 1978). For this technique, the reaction chamber
is evacuated to 10-300 Torr. LP-MOCVD also offers other advantages. The
gas speeds are higher, thus new gas compositions for heterostructures
are more quickly established leading to abrupt heterointerfaces.

Growth of Ga$_{0.47}$In$_{0.53}$As/InP quantum wells and superlattices:

Ga$_{0.47}$In$_{0.53}$As/InP quantum wells and superlattices grown with
metalorganic sources were first reported by Razeghi and co-workers
(Razeghi and Duchemin, 1984), using low pressure techniques. The PL
linewidth of the emission from a 8 Å quantum well was 9 meV, indicating
interface abruptness better than one monolayer (see figure 4.2).

Miller et al. (1986) reported the growth of high quality
Ga$_{0.47}$In$_{0.53}$As/InP quantum wells of well width from 135 Å down to 10 Å
using AP-MOCVD. Arsine and Phosphine were used as group V sources, and
TM-Ga and TM-In for group III. The growth temperature was 625°C with
the substrate held on a Si-C coated graphite susceptor and heated with
i.r. lamps. The growth time of the 10 Å layer was $\approx$1 s. Pressure-
balanced vent-run operation was used with a fast-switching gas manifold
with low dead volume. Transient pressures during gas switching, which
can lead to transients in the III-V ratios and thus lattice mismatch,
were avoided by automatically keeping the pressure across the manifold
constant to within ± 0.2 mTorr. The carrier gas flow rate was also
increased to 5-10 l/min, which gives gas velocities close to those of
LP-MOCVD. A pause of \( \approx 1 \) s before growth of the InP layers allowed complete removal of AsH\(_3\) to avoid incorporation of As if both gases are present. The background carrier concentration of undoped Ga\(_{0.47}\)In\(_{0.53}\)As was \( \approx 2 \times 10^{15} \) cm\(^{-3}\). PL linewidths show that the average interface abruptness was \( \approx 1 \) monolayer (see figure 4.2). The Ga\(_{0.47}\)In\(_{0.53}\)As/InP MQW APD's characterized in section 6.2 were grown by this technique. Low-loss waveguides (Koren et al., 1986) and phase modulators (Koren et al., 1987) have also been fabricated.

4.3.4 Molecular Beam Epitaxy.

Molecular beam epitaxy (MBE) is an epitaxial growth technique for semiconductor, metallic and insulator thin films. Growth involves the reaction of thermal beams of atoms or molecules with a crystalline substrate surface held at a suitable temperature under ultra-high vacuum (UHV) conditions. The UHV provides a very clean growth environment. Chopping of the molecular beam with mechanical shutters together with the slow growth rate (variable from 0.1 to 10 \( \mu \)m/hour) allows rapid variation of beam composition or doping, giving precise control down to the level of a single atomic layer. Coupled with the very uniform growth (thickness and composition variations less than 1\% over a 3" wafer) achieved using rotating substrate holders, this has made MBE a leading contender for the growth of high quality heterojunction devices.

Monocrystaline GaAs was first grown by Arthur (1968); further developments in the growth of high quality material for device applications were made by Cho (1983). Kinetic studies (Foxon and Joyce, 1977) showed that stoichiometric GaAs can be grown under excess As since the As sticking coefficient is very low above 500°C unless combined with Ga. Ternaries and quaternaries with As as the only group
Element \((A,\text{Ga})\text{As}, (A\text{l},\text{Ga})(\text{In,As})\) are also As stabilized. Composition and lattice matching therefore depend only on the flux ratios of the group III elements. However, alloys with two group V constituents such as \((\text{Ga,In})(\text{As,P})\), require control of the As/P ratio via control of beam fluxes and deposition conditions.

**Growth apparatus:**

Figure 4.3 shows a schematic diagram of a typical MBE reactor. The growth chamber is maintained under UHV conditions by a suitable pump (ion, cryo- or turbomolecular pump), and the sample is placed in the chamber via a load lock to avoid breaking the vacuum. Liquid nitrogen shrouds enclose the growth area to minimize water vapour and carbon containing gases. Substrate rotation (at 0.1 - 5 rev/min) is used to overcome the spatial variation of the intensity profiles of the beams over the substrate surface. The sources are Knudsen effusion cells made of pyrolytic boron nitride, with their temperature controlled to ± 1°C.

**Substrate preparation:**

Substrate preparation is critical for the growth of high quality material by MBE, unlike for LPE or VPE which use in-situ meltback or etching. It is essential to obtain a clean substrate, atomically flat and damage-free without C or O contamination. The standard method is to use oxide protection and in situ desorption. This prevents adsorption of Carbon which is not evaporated by later heat treatment. The full preparation procedure is described by Cheng et al, (1981).

**Dopants for III-V's:**

The commonly used n-type dopants for MBE are Sn and Si. Sn exhibits surface segregation at high growth temperature leading to non-abrupt doping profiles (Cho, 1975). For p-type doping, Be is the most often used.
Figure 4.3. Schematic diagram of typical MBE system (after Cho, 1983).
MBE growth of (Ga,Al)As:

The first high quality material to be grown by MBE was GaAs. The ternary (Al,Ga)As, and GaAs/(Al,Ga)As heterostructures are also easy to grow under As stabilization since the lattice matching is not critically dependent on the Al/Ga ratio. Good dimensional control in multilayer devices was first demonstrated in GaAs/(Al,Ga)As periodic structures (Cho, 1971). Graded band-gap devices obtained by compositional grading via control of the Al/Ga flux ratios have included graded base bipolar transistors (Malik et al., 1985) and graded-gap avalanche photodiodes (Capasso et al., 1981).

MBE growth of $\text{In}_x\text{Ga}_{1-x}\text{As}$:

To avoid lattice mismatch, the mole fractions must be controlled to better than 1% and uniform across the molecular beam. Cheng et al. (1981) used a coaxial In-Ga oven to ensure compositional uniformity across the beam. The In and Ga fluxes were controlled by aperture sizes and fine-tuned by the oven temperature (since Ga and In have different temperature coefficients of vapour pressure). They achieved uniform growth with a lattice mismatch of less than $5 \times 10^{-4}$, or a composition variation of less than 0.75%, over 3 cm square substrates.

MBE growth of InP and (Ga,In)(As,P):

InP and other P containing alloys are hard to grow by MBE due to the high partial pressure of P required. Special handling systems are required using either cooled shutters or a separate P removal chamber adjacent to the growth chamber (Tsang et al., 1982). High quality Ga$_{0.47}$In$_{0.53}$As/InP quantum wells have been grown by MBE using a two pump arrangement to control the high P vapor pressure (Marsh et al., 1985). The linewidth of a 10 Å well was 15 meV, indicating high quality interfaces.

MBE growth of the quaternary (Ga,In)(As,P) is hampered by the
complex nature of sticking coefficients of group V on group III elements.

These problems have been overcome by the use of gas sources for MBE growth (see below).

**MBE growth of (Al,Ga,In)As:**

(Al,Ga,In)As is suitable for MBE growth for the following reasons. Only one group V element is involved, so the As/P flux ratio problem does not occur. Ga, Al and In all have near-unity sticking coefficients at epitaxial temperatures, thus it is possible to control mole fractions precisely for lattice matching. Since no P is involved in the growth, no special pumping arrangements are required, and the same MBE kit as for growth of (Al,Ga)As (with the addition of an In source) is used.

The optical properties of Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As quantum wells were investigated by Welch et al. (1983). The smallest well width was 15 Å, giving a 0.474 eV shift in emission. Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW structures grown at 580°C without growth interruption showed excitonic behaviour in the room temperature absorption characteristic (Weiner et al., 1985).

As mentioned previously, (Al,Ga,In)As is particularly suited for MBE growth of compositionally-graded material for long-wavelength optoelectronic devices. Multiple graded-well APD's were grown by Alavi and Cho (Alavi et al., 1987) and are studied in chapter 8. The MBE growth kit (described earlier by Cheng et al., 1981) was modified to incorporate computer control of the cell temperatures and shutters. The In flux was kept constant (constant cell temperature) while changing the Ga and Al cell temperatures simultaneously, to keep the total Ga plus Al flux constant whilst linearly grading the composition. The growth rates were determined experimentally from thickness measurements for different values of the measured cell temperatures. The graded
layer was grown at a constant rate and the required cell temperatures for the desired composition profile were read from a look-up table. Multiple graded layers grown by this technique showed good surface morphology with no cross-hatching.

4.3.5 Gas Source Molecular Beam Epitaxy (GSMBE).

As mentioned above, the control of the As/P ratio in the MBE growth of (Ga,In)(As,P) alloys is difficult when conventional solid sources and effusion ovens are used. Panish reported the use of gas sources for the growth of GaAs and InP (Panish, 1980) and (Ga,In)(As,P) (Panish and Sumski, 1984). The As\(_2\) and P\(_2\) beam fluxes could be accurately controlled allowing growth of layers of precisely determined composition.

The GSMBE apparatus was a conventional MBE system with the following modifications. The effusion ovens for group V sources were replaced by a gas cracker. A high capacity pump was used to remove H\(_2\). A high pressure gas source was used due to its high conversion efficiency. The cracking occurred in resistively-heated alumina tubes with a "leaky seal" into the low pressure growth region. A gas handling system with low dead-volume provided fast switching of the beam fluxes.

The high quality of the heterointerfaces in Ga\(_{0.47}\)In\(_{0.53}\)As/InP heterostructures grown by this technique has been demonstrated by the growth of thin quantum wells (Panish et al., 1986). The photoluminescence linewidths are consistent with interface roughness significantly less than one monolayer (figure 4.2b). Ga\(_{0.47}\)In\(_{0.53}\)As/InP MQW photodiodes grown by GSMBE are discussed in section 9.3.
4.3.6 Chemical Beam Epitaxy (CBE).

Chemical beam epitaxy (or metalorganic MBE) combines the growth techniques of MBE and MOCVD (Tsang, 1984). Growth occurs in an MBE reactor using metal alkyl gas phase sources. The group V alkyls have a lower purity than the hydrides used in GSMBE but are less hazardous in use. The group V alkyls are thermally decomposed at low pressure with a Ta or Mo catalyst. The group III alkyls are heated in a separate effusion oven and pyrolysed at the heated substrate surface. Uniform composition of Ga_{0.47}In_{0.53}As across the wafer is obtained through use of the same effusion oven for In and Ga (Tsang, 1985).

High quality Ga_{0.47}In_{0.53}As/InP quantum wells of width 10 - 150 Å were grown using metal alkyls for group III and hydrides for group V sources (Tsang and Schubert, 1986). PL linewidths comparable to those obtained in MBE-grown GaAs/(Al,Ga)As quantum wells were obtained, with a linewidth of \( \approx 9 \) meV for a 10 Å well. This indicates an effective interface roughness of 0.12 \( \text{Å} \).

The above results represent some of the best heterostructures achieved to date in the Ga_{0.47}In_{0.53}As/InP material system, and open the way for the realization of new heterostructure device concepts and investigation of the physics of very high quality quantum wells.
EXPERIMENTAL TECHNIQUES

5.1 Sample design for measurement of ionization rates.

As noted in chapter 3, the extraction of the ionization rates from the multiplication versus field curves is non-trivial in the general case. However, careful control of the device structure and experimental conditions yields simple analytical formulae. The multiplication is measured separately under conditions of pure electron or pure hole injection into the high field region in the same junction, to avoid device-to-device variation. This is usually achieved by minority carrier injection in a reverse-biased P-N junction. Short wavelength light, which is completely absorbed in the neutral region, is shone on the P or N side of the junction, creating electron-hole pairs. The majority carriers are collected by the contacts, while the minority carrier diffuse into the depletion region. This is illustrated in figure 5.1.

Simple analytical expressions for the ionization rates can be obtained for a P⁺-I-N⁺ diode, a single-sided abrupt P⁺-N junction, a P⁺-N-N⁺ junction at punch-through and a linearly-graded P-N junction (Stillman and Wolfe, 1977). The P⁺-I-N⁺ diode is the favoured configuration due to the constant collection efficiency with field, and the position independence of the ionization rates. The relation between multiplication and ionization rates for a P⁺-I-N⁺ diode was given in chapter 2. Thus

\[ F = \frac{(V+V_{d1})}{W} \]  \hspace{2cm} (5.1)

\[ \alpha(F) = \frac{1}{W} \left[ \frac{M_e(V)-1}{M_e(V)-M_n(V)} \right] \ln \left[ \frac{M_e(V)}{M_n(V)} \right] \]  \hspace{2cm} (5.2)
Figure 5.1. Minority-carrier injection of electrons and holes by selective illumination of P-N junction.
\[ \beta(F) = \frac{1}{W} \left[ \frac{M_h(V)-1}{M_h(V)-M_e(V)} \right] \ln \left[ \frac{M_h(V)}{M_e(V)} \right] \quad (5.3) \]

\[ k(F) = \frac{\alpha(F)}{\beta(F)} = \frac{M_e(V)-1}{M_h(V)-1} . \quad (5.4) \]

Note that the P and N regions must be heavily doped (> \(10^{16} \text{ cm}^{-3}\)) to reduce change in the collection efficiency due to depletion of these layers.

Pure single carrier-type injection is required for accurate measurement of the ionization rate, but mixed injection can occur in a number of ways:

1. Incomplete absorption of the light in the neutral regions will result in absorption in the depletion region. The P⁺ and N⁺ regions are thus made sufficiently thick to absorb virtually all the incident light. For example, 99.8% of light from a He-Ne laser (\(\lambda = 0.63 \mu\text{m}\)) will be absorbed in 1 \(\mu\text{m}\) of InP (absorption coefficient \(\approx 6 \times 10^4 \text{ cm}^{-1}\) at 0.63 \(\mu\text{m}\) (Burkhard et al., 1982)).

2. Mixed injection can occur through stray light falling on the wrong side of the P-N junction (e.g., for the case of electron injection, on the N side) or by direct absorption of stray light at the exposed surface of the P-N junction (for a mesa device). Careful focussing of the light is essential to avoid these problems.

3. If the neutral region width is greater than several minority carrier diffusion lengths (\(L_p\)), carriers generated near the surface by strongly-absorbed radiation will recombine with majority carriers before they reach the depletion region. In a homostructure device, this will lead to low quantum efficiency. The recombination gives rise to sub-bandgap radiation. This radiation will not be absorbed at low fields, but Franz-Keldysh electroabsorption (Stillman and Wolfe, 1977) in the depletion region is significant at high fields (figure 5.2(a)).
Figure 5.2(a). Mixed injection by Franz-Keldysh electro-absorption of recombination radiation.

(b). Device structure to avoid recombination in the substrate.
This gives rise to both mixed injection, and a unity baseline photocurrent which is very strongly field dependent. This can be an important source of error in the experimental determination of ionization rates (Bulman et al., 1982). Typical minority diffusion lengths have been quoted from \( \approx 1 \, \mu m \) to \( \approx 15 \, \mu m \) (Bulman et al., 1982; Diadiuk et al., 1983; Osaka et al., 1985) depending on the doping. The effect is therefore particularly likely for illumination of the thick substrate side of the P-N junction. The effect can be avoided by thinning the substrate to within a few \( \mu m \) of the depletion region. Bulman et al. (1982) used an etch-stop layer of different composition to achieve uniform thinning by selective chemical etching (figure 5.2(b)).

Other problems which can occur are non-uniform electric fields across the active area of the device (due to doping fluctuations, microplasmas or edge effects), high leakage currents (leading to gain saturation) and variation in the unity gain base-line (due to increasing collection efficiency in P\(^+\)-N junctions or carrier recombination at low fields in heterojunction APD's).

The above problems demand careful sample design and characterisation for the accurate determination of ionization rates. The typical samples are P\(^+\)-I-N\(^+\) diodes with P, I and N regions of 1-2 \( \mu m \) thickness, fabricated as mesa devices.

5.2 Device fabrication procedures.

The grown wafers were processed by making electrical contacts to electrically isolated devices of areas of the order of \( 10^{-4} \, cm^2 \). This reduces the probability of containing a defect within the active area of the device and low leakage diodes can thus be formed. For ease of processing, mesa-etched devices were used throughout.
The procedure for fabrication was as follows. The wafer was first thinned, by mechanical polishing with emery powder on a quartz pad, to a thickness of about 300 μm. Samples for experiments requiring back illumination were thinned to 150 μm to reduce free carrier absorption in the substrate. The final polishing stage was a chemo-mechanical polish with bromine-methanol (2% 1%) to produce a mirror-like finish.

P-type contacts to the top layer were formed by evaporation in an electron beam evaporator. A shadow mask in contact with the epitaxial layers was used to define circular pads of diameter 50 μm and spacing 500 μm. The contacts were formed by evaporating 100 Å of Au-Be alloy (1% Be by weight) followed by 1500 - 2000 Å of Au to facilitate bonding and probing. Ohmic contacts were produced by alloying at 420°C for 1-2 seconds on a graphite strip heater. The contact resistance could be measured at this stage by measuring between adjacent contact pads on the surface.

N-type contacts were formed on the back by evaporation (or electroplating) of Au (200 Å)/Sn (100 Å)/Au (1500 Å) and subsequent alloying. Windows (50 μm diameter) were produced below the active area of the devices using lift-off techniques or etching of the Au with I₂/KI etch. The top and back photolithography patterns were aligned using a mask aligner with infra-red transmission optics.

Mesa patterns were defined in photoresist using standard photolithography techniques with either positive or negative resist. The standard mesas were ovals of area 1.3 x 10⁻⁴ cm², of dimensions roughly 100 μm x 150 μm. The contact is at one end of the mesa, giving a clear region of the active area for illumination. The mesas were etched down to about 1 μm inside the lower N⁺ contact layer to ensure electrical isolation. The best results for InP-based multilayer structures were obtained with Bromine-Methanol etch (0.1%) and
Waycoate negative resist. This etch is not material-selective, although the etch rate is fast and negative resist is required. Note that the shape of the etched mesa is important in determining the breakdown behavior of the device. The condition of the exposed P-N junction is also important. After resist stripping and final cleaning, absorbed water vapour was removed by baking at 120°C in a hydrogen atmosphere for 30 minutes. The finished devices were then stored in a dessicator.

Some of the devices were passivated by applying an insulating epoxy (Sylvax 600) to the mesa edges. This appears to increase the lifetime of the devices and reduce the likelihood of damage during mounting and bonding. The epoxy was applied by hand to individual devices.

Later in the course of the work it was found desirable both to find a more convenient method of protecting and passivating the mesas, and to design a configuration less sensitive to stray light absorption (e.g. for use in low temperature cryostats). The method used was a modification of a technique used by Cinguino et al. (1985) for minimising the contact area of a top-illuminated photodiode. Mesas were formed by etching a ditch (≈ 25 µm wide), and planarisation and passivation was carried out by applying Merck HTR-3 polyimide. A large area, thick Au/Ti contact is evaporated on top of the device and windows opened for illumination. Illumination of the mesa edges or the substrate by stray light is then prevented by the metallisation. Bonding is also greatly facilitated as it can be carried out away from the mesa. Figure 5.3(a) shows the photoresponse of a device showing strongly enhanced photocurrent at the mesa edges. The metallisation shown in figure 5.3(b) restricts the illumination to the central uniform region.

Contacts to the wafers were made by electrical micro-probes. For
Figure 5.3(a). Spatial photoresponse of a mesa device, showing edge effects.

(b). Edge illumination prevented by metallisation.
some of the low-temperature work, the samples were mounted on sapphire pads (which have high transparency for back-illumination and good thermal conductivity), using silver paste. The devices were bonded using an Indium ball bonding technique (Mil'Shtein et al., 1985). This requires no heating of the substrate or pressure during bonding, and thus avoids the problems of device degradation sometimes associated with thermocompression ball-bonding of minority carrier devices.

5.3 Electrical and optical characterisation.

Current-voltage (I-V) and capacitance-voltage (C-V) measurements were made to determine the properties of the P-N junctions. Optical measurements (spectral response and quantum efficiency) were used to give information on the internal efficiency of photodiodes and on the quality of quantum well samples.

The current characteristics were measured in the bias range -100 V to +100 V using an HP 4145A parameter analyser. Currents as low as $10^{-13}$ A (0.1 pA) could generally be measured, limited by stray leakages in the experimental set-up. Rucker and Koll probes, with coaxial shielding to within a few mm of the probe tip, were used to minimize pick-up. Reverse-bias currents in excess of $\approx 10^{-5}$ A generally led to irreversible breakdown.

C-V curves are used to find the electric field profile within the device. For a P⁺-I-N⁺ device the capacitance is constant with applied bias, with a typical value of $\approx 1.5$ pF ($C = \varepsilon_s A/d$, where $\varepsilon_s \approx 13 \varepsilon_0$, $A = 1.3 \times 10^{-4}$ cm², and $d = 1 \mu$m). The C-V curves were measured at 1 MHz on an EG&G 410 meter, which includes a variable-rate voltage ramp. Variable-frequency capacitance measurements were made with an EG&G 5204 lock-in amplifier. Depletion profiles were also obtained for some MQW structures using a Miller feedback profiler (see chapter 9).
The quantum efficiency was measured by imaging the light from a pin-hole at the exit of a monochromator onto the active area of the photodiode, using a 25X microscope objective (spot size ≈ 40 μm for a ≈ 1 mm pinhole). The device was biased, and the photocurrent measured (with computer-controlled data aquisition), as described in the following section. The photocurrent of a calibrated Si or Ge detector was also measured on each occasion. The device quantum efficiency is then calculated from the known quantum efficiency of the calibrated detectors.

5.4 Apparatus for measurement of ionization rates.

The ionization rates were measured using the method of minority carrier injection described in section 5.1. Light from a He-Ne laser (λ = 0.63 μm) was focussed onto the top and back of the mesa. An image formed by the imaging lens was demagnified by a 25x microscope objective to give a laser spot of diameter < 10 μm. The long working distance objective allowed room for convenient operation of the prober tips. The light was incident via a beam-splitter, allowing simultaneous observation of the sample for probing purposes and verification of the position and focusing of the light spot (figure 5.4). The beam-splitter had a semi-reflecting coating for long wavelengths, with non-parallel faces to avoid multiple light spots. Absence of scattered radiation was verified by focussing the light spot some way from the device, and observing negligible photocurrent. The sample and prober were mounted on a platform movable in three dimension by means of micropositioners. The sample was illuminated from behind using a 20x ω-corrected objective. Positioning and focussing of this light spot were verified by removing the device from the optical path and viewing the light spot from the upper microscope. The intensity of the incident illumination
Figure 5.4. Experimental arrangement for illumination of mesa device.
could be varied over several orders of magnitude by the insertion of appropriate neutral density filters.

The incident illumination was chopped using a variable-frequency mechanical chopper, and the synchronous photocurrent detected on a lockin amplifier. The photocurrent was preamplified by an EG&G 181 current amplifier. The device was biased via a simple RC bias tee (see figure 5.5). A low-pass filter in the DC arm reduced noise from the bias supply. The DC output of the lockin amplifier, which is proportional to the photocurrent, was fed to a chart recorder.

The spatial uniformity of the photoresponse of the devices were measured by raster scanning the laser spot over the active area ("flying spot" microscope). This was achieved using two perpendicular oscillating mirrors (General Scanning Inc.). Signals from the driver amplifiers, proportional to the position of the light spot, were fed to the X and Y inputs of an oscilloscope, and the photocurrent (from the output of the current preamplifier) was added to the Y signal. This produced images of the type shown in figure 5.6. The dip in the photoresponse at one end of the mesa is due to the contact pad and prober. Devices were selected for detailed study on the basis of low dark current and uniform gain.

5.5 Measurement of the pulse response of photodiodes.

The photodiodes were characterised at high frequency by measuring the response to a fast light pulse from a semiconductor laser. 1.3 µm and 1.5 µm lasers were used. The laser, mounted in a temperature-controlled package, was biased close to threshold, and driven by a pulse generator. This was either a comb generator, producing pulses of ~100 ps duration at a 100 MHz repetition rate, or an HP214B pulse generator. The laser pulses were monitored from one end of the laser
Figure 5.5. Experimental arrangement for photocurrent measurements.
Figure 5.6. Raster scan of a mesa photodiode with uniform response.
with an ultrafast Ge detector (response time \( \approx 90 \) ps FWHM). The radiation from the other laser facet was focussed onto the photodiode. The photodiode was mounted with silver epoxy onto one arm of a 50 Ω stripline fabricated on printed circuit board, and contacted to the other arm with a gold wire bonded to the device. One side of the stripline was connected via high frequency coaxial cable to the S4 sampling head of a high speed oscilloscope. The cable was impedance matched at the end close to the stripline by a parallel 50 Ω termination (see figure 5.7). The DC side of the stripline was connected to the bias supply via a long length of standard coaxial cable. This was found to damp out reflections in this line due to the finite resistance of the reverse-biased photodiode in the avalanche region. A Tektronix 541 curve tracer was used as the bias supply, since it allowed simultaneous observation of the photocurrent for optimising alignment. For measurement of the pulse response of low intensity pulses, the signal from the photodiode was preamplified by two B&H 7020AH amplifiers in series, each with 20 dB gain, and the output of the oscilloscope was connected to an EG&G 4001 signal averager.

5.6 Low temperature techniques.

Low temperatures were used both to study the temperature dependence of various transport mechanisms and to "beat the leakage". For routine measurements, a miniature Joule-Thompson cooler (MMR Inc.) was adapted to allow probing of the devices at temperatures between \( \approx 90 - 100 \) K and 400 K. This assembly could be placed under the flying-spot microscope in place of the standard prober, acting as a cold stage and allowing the full range of low-frequency measurements to be carried out at low temperature, without the necessity for bonding the devices. Careful calibration showed that the temperature at the device could be
Figure 5.7. Experimental arrangement for measuring pulse response of photodiodes.
up to 20 K higher than that displayed, at the lowest temperatures, due to the limited capacity of the cooler. For lower temperature work, a continuous flow cold-finger Helium cryostat (≈ 8 K - 300 K) was used. Light was focussed onto the sample in this case by an Ealing long-working-distance reflecting objective. This had the additional advantage of being achromatic. For temperatures below 8 K and when the temperature needed to be accurately known, a Helium immersion cryostat was used. Light focussing was more difficult in this case.
**CHAPTER 6**

Ga$_{0.47}$In$_{0.53}$As/InP MULTIPLE QUANTUM WELL APD'S GROWN BY VPE AND MOCVD.

In this chapter, APD's with Ga$_{0.47}$In$_{0.53}$As/InP MQW multiplication regions, grown by VPE and by MOCVD, are described. In particular, room temperature avalanche multiplication characteristics and carrier pile-up effects were investigated.

In section 6.1, the characterization of Ga$_{0.47}$In$_{0.53}$As/InP MQW APD's, grown by continuous trichloride vapor phase epitaxy, is discussed. Devices exhibited multiplication factors of $\approx 6$ at room temperature. The response to a fast light pulse from a semiconductor laser showed no long tails due to carrier pile-up.

In section 6.2, Ga$_{0.47}$In$_{0.53}$As/InP MQW photodiodes grown by atmospheric pressure MOCVD are described. These exhibited DC gains of more than 20 at room temperature. Carrier pile-up was observed only at low fields. High resolution TEM showed that the interfaces are considerably sharper than for the VPE samples considered in section 6.1.

In section 6.3, the ionisation rates in Ga$_{0.47}$In$_{0.53}$As/InP MQW APD's are discussed.

**6.1 Ga$_{0.47}$In$_{0.53}$As/InP MQW APD's grown by continuous trichloride vapour phase epitaxy.**

**6.1.1 Material growth.**

The Ga$_{0.47}$In$_{0.53}$As/InP MQW APD's studied in this section were grown by Mattera et al. (1986) using the continuous single-chamber trichloride VPE process described in section 4.3.2.

The Sn-doped substrates used were oriented 2° off the 100
direction. The growth temperature was 675 °C. A 1 µm-thick undoped InP spacer layer was grown. Then, alternating layers of Ga₀.₄₇In₀.₅₃As and InP were grown, followed by an InP cap layer (2.2 µm thick). A Ga₀.₄₇In₀.₅₃As cap (0.2 µm thick) was then grown prior to withdrawal of the substrate from the furnace.

6.1.2 Characterization of layers.

The growth rates, background carrier concentrations (obtained by capacitance-voltage profiling) and layer thicknesses (obtained by transmission electron microscopy (TEM)) of the two wafers grown are summarized in Table 6.1.

Table 6.1. Characteristics of MQW's grown by VPE.

<table>
<thead>
<tr>
<th>#</th>
<th>Layer</th>
<th>Growth rate (A/s)</th>
<th>Carrier concentration (10⁵ cm⁻³)</th>
<th>Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>InP (MQW, x9)</td>
<td>5.8</td>
<td>1.0</td>
<td>350 A</td>
</tr>
<tr>
<td></td>
<td>Ga₀.₄₇In₀.₅₃As (MQW, x10)</td>
<td>13.3</td>
<td>1.0</td>
<td>950 A</td>
</tr>
<tr>
<td></td>
<td>Total MQW thickness</td>
<td></td>
<td></td>
<td>1.27 µm</td>
</tr>
<tr>
<td></td>
<td>InP (cap)</td>
<td>5.8</td>
<td>0.8</td>
<td>2.2 µm</td>
</tr>
<tr>
<td>B</td>
<td>InP (MQW, x19)</td>
<td>6.7</td>
<td>1.9</td>
<td>650 A</td>
</tr>
<tr>
<td></td>
<td>Ga₀.₄₇In₀.₅₃As (MQW, x20)</td>
<td>11.6</td>
<td>1.0</td>
<td>950 A</td>
</tr>
<tr>
<td></td>
<td>Total MQW thickness</td>
<td></td>
<td></td>
<td>3.14 µm</td>
</tr>
<tr>
<td></td>
<td>InP (cap)</td>
<td>6.7</td>
<td>0.1</td>
<td>2.2 µm</td>
</tr>
</tbody>
</table>

Note the large difference in the purity of the InP cap layer (Table I) between Sample A and Sample B. The carrier density in sample B is about an order of magnitude lower than material grown under the same conditions without a multilayer buffer. A possible explanation for this reduction in the background doping is that the multilayer region
acts to getter impurities. Previous studies have shown that the photo-

luminescence properties of GaAs layers grown by MBE are improved by the

insertion of quantum well and superlattice buffer layers (Gossard et al., 1982; Masselink et al., 1984).

Data from secondary-ion mass spectroscopy (SIMS), and cross-

sectional transmission electron microscopy (TEM) were provided by co-

workers. The SIMS analysis was performed on a Cameca IMS-3F instrument

using Cs⁺ ions at 10 keV for sputtering and electronegative secondary

ions at -4.5 keV for detection. This configuration optimized the

sensitivity to the principal material constituents (In, Ga, As and P).

A primary beam focused to a 50 μm diameter spot was rastered over a

250 x 250 μm square area. An aperture was used to restrict the signal

detected by the spectrometer to a 60 μm diameter region at the centre

of the crater. This enhances depth resolution by minimizing

contributions from material sputtered from the crater walls. Figure 6.1

shows the SIMS data for sample B. The results for sample A were very

similar. The periodic behavior of the profiles of the four elements

traced shows the formation of compositionally well-defined layers. The

slope in the SIMS trace at each interface is due to both the limited

depth resolution of the technique and to compositional grading.

Figure 6.2 shows a TEM micrograph of sample B. The results

indicate that there are no large-scale dislocations present within the

MQW region.

6.1.3 Device fabrication and characteristics.

The P⁺-region was defined by Zn diffusion into the N⁻-InP top

layer. The junction depth was measured by lapping and staining to be

1.5 μm in Sample A and 1.84 μm in Sample B. Note therefore that the

junction is within the 2.2 μm-wide InP cap layer. Mesa devices (100 μm
Figure 6.1. SIMS data for sample B (20 periods).
Figure 6.2. TEM micrograph of sample B (20 periods).
in diameter) were etched using Br-Methanol and Ohmic contacts were formed as described in section 5.2.

The buffer layer is doped due to out-diffusion of Sn from the substrate during growth. Capacitance-voltage measurements showed that the capacitance at punch-through is 0.35 pF (sample A) and 0.30 pF (sample B). This corresponds to a depletion region thickness of \( \approx 2.7 \) \( \mu \)m for sample A and \( \approx 3.1 \) \( \mu \)m for sample B. Punch-through occurs at \( \approx -35 \) V in both structures.

The inset of Figure 6.3 shows the reverse-bias current-voltage characteristic of sample B, measured using an HP 4145 parameter analyzer which ramps up to 100 V bias. Similar results were obtained for sample A but with a slightly lower breakdown voltage. The dark current at half the breakdown voltage is \( \approx 20 \) nA for a device area of \( 7.8 \times 10^{-5} \) cm\(^2\). The current at \(-120\) V was found to be \( \approx 1 \) \( \mu \)A using a curve tracer. Some devices on the wafer showed a large dark current component increasing from \( \approx 200 \) nA at \(-75\) V to \( \approx 100 \) \( \mu \)A at \(-100\) V. The exponentially-increasing behaviour suggests that this may originate in surface current components (Moss and Ritchie, 1983), or possibly in tunnelling via defect states. Figure 6.3 also shows the reverse-bias photocurrent as a function of voltage for radiation of wavelength \( \lambda = 6328 \) \( \text{\AA} \) (He-Ne laser) incident on the top of the detector. This corresponds to electron injection since at this wavelength all the light is absorbed in the top layer. Between zero bias and about \(-10\) V, the photocurrent increases slightly. The photocurrent is almost constant between \(-10\) V and \(-100\) V, corresponding to unity multiplication and unity internal quantum efficiency. At higher bias, the photocurrent increases due to avalanche multiplication. The reduced quantum efficiency at low fields is probably due to carrier recombination within the wells. The emission rate from the wells
Figure 6.3. Reverse-bias current-voltage (insert) and photocurrent-voltage (main figure) curves for sample B.
increases with field as discussed in section 3.6, and the photocurrent saturates when the recombination rate is small compared to the emission rate. The maximum multiplication factor is ≈ 6 and is uniform (to better than 10%) across the active area of the device. Because the P-N junction lies within the InP layer, the peak electric field occurs in the n^-InP. However, because the ionisation threshold for InP is much larger than for Ga_{0.47}In_{0.53}As, the multiplication is expected to occur only in the Ga_{0.47}In_{0.53}As layers. Avalanche breakdown would occur in the InP at several hundred Volts at these carrier concentrations.

6.1.4 High speed response.

For the study of carrier pile-up effects on the pulse response, the diodes were mounted on a 50 Ω stripline coupled to a sampling scope, as described in chapter 5. The light source was a semiconductor laser emitting at 1.5 μm. The laser was driven either by a comb generator (which produces pulses with a FWHM of ≈ 75 ps) or with a HP214B pulse generator (giving a square pulse of width 50 ns). The output of the S4 sampling head was signal averaged to allow a precise measurement of the pulse width of the detector at low incident optical powers. Devices fabricated from both wafers exhibited similar high speed characteristics. Figure 6.4(a) shows the pulse response of sample A at -50 V bias. The FWHM is ≈ (100 ± 10) ps. Figure 6.4(b) shows the pulse height as a function of reverse-bias. Below punch-through at ≈ 35 V, the pulse height is reduced because the thermal emission rate of photo-generated carriers out of the undepleted wells is slow compared to the laser pulse. A small amount of multiplication is observed at 90 V.

Forrest et al. (1982) showed the importance of measuring the pulse response for low intensity and longer duration pulses. Figure 6.5 shows the response of sample A to laser pulses with a peak power of ≈ 50 μW
Figure 6.4. Pulse response of sample A.

(a) Response at -50 V to a 75 ps pulse from a 1.5 μm laser.

(b) Pulse height as a function of reverse bias.
Figure 6.5. Response to a 1.5 μm 75 ps FWHM laser pulse;
(a) at 50 μW incident peak power,
(b) at 5 mW incident peak power.
(a) and \( \approx 5 \) mW (b). The response to a longer laser pulse is shown in Figure 6.6; (a) is the response of a fast Ge photodiode, (b) is the MQW photodiode response at incident power of 500 \( \mu \)W and (c) is the response at \( \approx 10 \) \( \mu \)W. Forrest et al. (1983) estimated a critical incident optical power for SAM-APD's of \( \approx 150 \) \( \mu \)W, above which barrier lowering due to charge accumulation occurs. The absence of long tails in the trailing edge of the pulse over this range of optical intensities indicates that pile-up effects of carriers in the wells are small.

An intrinsic response time of between 30 and 70 ps can be estimated using a sum of the squares approximation (valid for a Gaussian pulse-shape) and the known laser FWHM (\( \approx 75 \) ps) and the RC time constant of the diode (\( \approx 2.2 \) RC \( \approx 39 \) ps). The intrinsic response time in the absence of gain is limited by the transit time of the device plus the emission time of carriers from the wells. The saturated drift velocities of electrons and holes in InP and Ga\(_{0.47}\)In\(_{0.53}\)As (see section 2.3.4) indicate a transit time for a device of depletion width 3 \( \mu \)m of about 50 ps. Thus the high frequency response of these devices appears to be limited by transit times, and not by carrier pile-up effects at the heterointerfaces.

As discussed in chapter 3, various techniques for grading the heterointerface region have been shown to improve considerably the speed of response of InP/Ga\(_{0.47}\)In\(_{0.53}\)As SAM APD's. The high speed of response of these VPE MQW APD's can be partially explained by interface grading. A grading length of 100 \( \AA \) is sufficient to reduce the barrier height for holes and electrons to zero at fields close to breakdown.

In the next section, Ga\(_{0.47}\)In\(_{0.53}\)As/InP MQW APD's grown by MOCVD with more abrupt interfaces are studied.
Figure 6.6. Response to a 50 ns laser pulse:

(a) fast Ge diode,

(b) MQW photodiode at $\approx 500 \, \mu W$ peak power,

(c) MQW photodiode at $\approx 10 \, \mu W$ peak power.
6.2 High speed $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ MQW APD's grown by MOCVD.

In this section the first MQW APD's grown by MOCVD are reported. The layers were grown by U. Koren and B.I. Miller using the techniques described in section 4.3.3.

6.2.1 Material growth.

The P-I-N structure was grown on a Zn-doped $\text{P}^+$ (3 x $10^{18}$ cm$^{-3}$) <100> InP substrate, starting with a 0.5 μm buffer layer (P-type, $\approx 2 \times 10^{17}$ cm$^{-3}$). The undoped $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$/InP MQW region is 2 μm thick comprising 100 periods of 100 Å thick individual layers. The 1 μm thick S-doped InP top layer (N+, $\approx 10^{18}$ cm$^{-3}$) is separated from the MQW by a 0.2 μm InP spacer layer (N-type, $\approx 2 \times 10^{17}$ cm$^{-3}$). Background doping was $\leq 10^{16}$ cm$^{-3}$ for InP and $\leq 2 \times 10^{15}$ cm$^{-3}$ for $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$. TEM studies indicated abrupt interfaces for $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$-on-InP; the inverted interface showed some roughness.

6.2.2 Device characteristics.

Devices of area $3.0 \times 10^{-4}$ cm$^2$ were fabricated using standard wet-etching techniques. Ohmic contacts were applied with Au-Zn/Au (p-type) and Sn/Ni/Au (N-type). The MQW region of the P-I-N devices was completely depleted at a few volts reverse bias at room temperature (as expected from the doping levels), with a punch-through capacitance of 1.6 pF.

The room temperature reverse bias I-V characteristic is shown in the inset to Figure 6.7. The dark current is dominated by thermal generation at low bias. The dark current at $\frac{1}{2} V_B$ is $\approx 50$ nA. At higher bias ($\approx 55$ V), the current increases strongly with bias due to band-to-band tunnelling. The slope of the characteristic increases
Figure 6.7. Photocurrent multiplication of MOCVD APD at 1.3 μm.

Inset: Reverse-bias current-voltage at room temperature.
further at \( \approx 65 - 70 \) V due to avalanche multiplication. Figure 6.7 also shows the photocurrent as a function of reverse bias for top illumination with light of wavelength 1.3 \( \mu \text{m} \). The photocurrent is constant with bias between about 10 and 60 V, corresponding to unity internal quantum efficiency. The gradual increase in photocurrent between 0 and 10 V can be explained by carrier recombination in the wells and may be associated with incomplete depletion at low bias. This is studied in more detail in section 9.3. At bias in excess of \( \approx 60 \) V (a field \( \approx 3 \times 10^5 \) V-cm\(^{-1}\)), the photocurrent increases due to avalanche multiplication, with a maximum multiplication \( \approx 20 \). The sharp increase in the multiplication between 60 and 70 V indicates that the ionisation rates ratio is close to unity (see section 6.5). The spatial uniformity of the photoresponse was examined using a focussed (\( \approx 5 \mu \text{m} \)) He-Ne laser spot raster-scanned over the active area of the device. The photocurrent in the active area is uniform to within \( \approx 10\% \), but is strongly enhanced at the mesa edge. The existence of uniform multiplication in the active area was verified by observing a line scan on the CRO as the bias was increased.

The quantum yield, which is the product of the external quantum efficiency and the multiplication, is shown in Figure 6.8. The APD's were illuminated with light from a monochromator, and the measured photocurrent normalized to the response of a calibrated Ge (\( \lambda > 0.95 \) \( \mu \text{m} \)) or Si (\( \lambda < 0.95 \) \( \mu \text{m} \)) detector. The response at 40 V reverse-bias corresponds to unity multiplication. The maximum quantum yield is 75\%, corresponding to approximately unity internal quantum efficiency for a photodiode without antireflection coatings. The quantum yield at 0 V is reduced by carrier recombination in the quantum wells, and is increased at 65 V due to multiplication. The maximum quantum yield shown is \( \approx 1.5 \) at 65 V, corresponding to a
Figure 6.8. Quantum yield as a function of photon energy at three different reverse biases.
multiplication of $\approx 2$. The response at photon energies $> 1.35$ eV is reduced due to absorption in the InP contact layer.

The features in the spectral response are associated with inter-subband transitions. A study of the room temperature absorption spectra of these structures (Bar-Joseph et al., 1987) showed excitons associated with the $n=1$ ($E_1 \approx 0.8$ eV) and $n=2$ ($E_2 \approx 0.9$ eV) transitions, which shifted in energy with applied field due to the quantum-confined Stark effect. The features at higher energies (e.g. at $\approx 1.05$ eV) in the spectral response of figure 6.9 may be transitions associated with the $n=3$ quasi-bound level, or weakly-forbidden transitions which become allowed as the field increases. The low leakage current and the observation of excitonic behaviour at room temperature shows the high electrical and optical quality of these structures.

6.2.3 High speed behaviour.

The pulse response of the MOCVD devices was measured as described previously. The response to 100 ns square pulses from a laser emitting at 1.3 $\mu$m at an optical power of $\approx 1$ $\mu$W is shown in Figure 6.9(a). The response is averaged over 100 pulses. The response at zero bias shows an exponential rise and fall with a tail of length $\approx 30$ ns. This is similar to effects seen in SAM-APD's (see section 3.2.2) due to pile-up of holes at the valence-band heterojunction interface. At low fields the response is limited by the thermionic emission rate of heavy holes from undepleted quantum wells. An emission rate of $\approx 30$ ns is estimated for holes of mass $0.45 m_0$ over a 0.3 eV barrier at a lattice temperature of 300 K, using the expressions of section 3.6. The pulse height at zero bias is reduced due to recombination in the wells, thus the recombination rate is of the order of $10^8$ s$^{-1}$. At a reverse bias of
Figure 6.9. Response of MOCVD APD to 1.3 μm laser pulses.

(a) 100 ns pulse at optical power of $\approx 1 \mu W$, at zero bias (lower trace) and $-10^3$V (upper trace).

(b) 90 ps pulse at optical power of $\approx 1$ mW and $-60$ V bias.
a few volts the quantum well region is fully depleted. The recombination rate in the well decreases with field due to the decreased electron-hole overlap, and the emission rate from the well increases due to tunnelling and hot carrier effects. Thus the long tail in the pulse disappears, and the pulse height increases as recombination becomes negligible (figure 6.9(a)). The FWHM of the response to a 90 ps pulse is about 260 ps at 10 V reverse-bias; this reduces to 200 ps at 60 V (Figure 6.9(b)). The peak height of the pulse also increases at high bias, showing the presence of multiplication at frequencies of several GHz. No evidence for gain-bandwidth product limitation was found at the small multiplications measured. The intrinsic response time is estimated by deconvoluting the RC time constant (≈ 2.2 RC ≈ 170 ps) and the laser pulse width (90 ps). The response time decreases from ≈ 170 ps at 10 V to ≈ 50 ps at 60 V. The reduction in response time suggests that there is still some pile-up of holes at fields of the order of 10^5 V·cm^{-1}.

The transit time for a 2 μm depletion region is ≈ 30 - 40 ps (see section 2.3.4). The measured response time of 50 ps, which includes an estimated error of about ± 20 ps, is thus close to the transit time. This suggests that at fields close to avalanche breakdown the high frequency response is not limited by carrier pile-up effects. The tunnelling time for heavy holes at the lattice temperature at a field of 3 x 10^5 V·cm^{-1} is estimated to be greater than 10^{-7} s. Thus there appears to be significant heating of the carriers in the well at high fields, in order to account for the observed lack of pile-up. High speed operation with a pulse response of 200 ps has also recently been reported in multiple quantum well (Al,Ga)As/GaAs P-I-N photodiodes grown by MBE (Larson et al., 1985). Ga_{0.47}In_{0.53}As/InP P-I-N photodiodes grown by GSMBE (Temkin et al., 1985), which have very
abrupt interfaces, have also shown high speed of response without evidence for carrier pile-up.

6.3 Ionisation rates in Ga$_{0.47}$In$_{0.53}$As/InP MQW APDs.

The above structures are not optimized for pure single carrier-type injection into the depletion region and thus for accurate ionisation rate measurements. Further, devices grown by both VPE and MOCVD showed strong enhancement of the photocurrent at the mesa edge, although the photoresponse of the active area was spatially uniform. This complicates measurements due to the difficulty of completely ruling out absorption of stray light at the mesa edge. Finally, the devices tend to have a short lifetime at high applied fields at room temperature, making measurements of the multiplication factors difficult.

The photocurrent was measured in the MOCVD-grown sample for top illumination with 1.3 µm, corresponding to absorption within the depletion region and thus mixed carrier initiation of multiplication, and 0.63 µm radiation, which is strongly absorbed in the top 1 µm N$^+$ InP layer and therefore injects holes. The multiplication curve does not change appreciably, indicating that $\alpha/\beta \approx 1$ in this structure.

The ionisation rate ratio in Ga$_{0.47}$In$_{0.53}$As is approximately $\alpha/\beta \approx 2$ (Pearsall, 1980). In the Ga$_{0.47}$In$_{0.53}$As/InP heterojunction, the valence-band discontinuity is greater than the conduction-band discontinuity. Thus any enhancement of the ionisation rates would be expected to preferentially enhance the hole ionisation rate ($\beta$). With increasing enhancement, the ratio $\beta/\alpha$ ($\approx 0.5$ in the bulk) would therefore increase, passing through unity. Note that for the 100 Å quantum wells in the MOCVD-grown samples, quantum-size effects may have a significant effect on the ionisation rates.
Recently, more detailed measurements of the ionisation rates in Ga$_{0.47}$In$_{0.53}$As/InP MQW's were made by Osaka et al. (1986). The 25 period structure with 350 Å Ga$_{0.47}$In$_{0.53}$As wells and 450 Å InP barriers was grown by chloride VPE. The undoped 2 µm MQW region, with 0.5 µm of undoped InP on either side, was grown between heavily doped P and N regions. The photocurrent multiplication was measured with 0.83 µm illumination from the top (pure injection of electrons) and 1.3 µm illumination through the substrate (mixed injection via absorption in the Ga$_{0.47}$In$_{0.53}$As quantum wells). An ionisation rate ratio $\beta/\alpha \approx 2$ was derived from these measurements. The electron ionization rate was similar to that in bulk Ga$_{0.47}$In$_{0.53}$As. The hole ionisation rate was enhanced over the bulk rate due to the large valence band discontinuity ($\Delta E_V = 0.38$ eV). The crossover in ionisation rates compared to the bulk material is strong evidence for enhancement of the ionisation rates by the band-edge discontinuities as discussed in section 3.2.3.
In this and the following chapter, long-wavelength Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW APD's grown by MBE are studied. This heterojunction system is attractive for use in MQW APD's due to the large conduction-band discontinuity, thus the ionisation rate for electrons is expected to be enhanced considerably more than the hole ionisation rate. Furthermore, the small valence-band discontinuity reduces pile-up of holes at the heterojunction interfaces. Chapter 8 investigates current multiplication effects occurring in samples exhibiting large reverse-bias leakage currents. This chapter deals with band-to-band impact ionisation in a high-quality quantum well structure.

7.1 Device structure and growth.

The samples were grown by A.Y. Cho using solid-source MBE (see chapter 4) at a growth temperature of 600°C. The structure for wafer D418 is shown in figure 7.1. The layers were grown on an N$^+$ S-doped InP (100) substrate. The 0.57 µm Si-doped (N$^+$, $10^{17}$ cm$^{-3}$) Al$_{0.48}$In$_{0.52}$As buffer layer is followed by the MQW region, which consists of 35 periods of Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As with 139 Å well and barrier layer thickness, for a total thickness of 0.97 µm. The MQW's are not intentionally doped and have a background doping (N-type) of the order of $10^{14}$ cm$^{-3}$. The 1 µm Be-doped (P$^+$, $2 \times 10^{18}$ cm$^{-3}$) window layer is capped with 150 Å of Ga$_{0.47}$In$_{0.53}$As Be-doped (P$^+$, $2 \times 10^{18}$ cm$^{-3}$) for contact purposes. Structures with 50 periods of 103 Å layers were also
Figure 7.1. Structure of $\text{Al}_{0.48}\text{In}_{0.52}\text{As}/\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ MQW APD (D416).
7.2 Device fabrication and characteristics.

Mesa devices of area $\approx 1.3 \times 10^{-4}$ cm$^2$ were fabricated by the procedures described in chapter 5. The room temperature current-voltage characteristic of the device with 139 Å wide wells and barriers is illustrated in Figure 7.2. The breakdown voltage, due to avalanche multiplication, varies across the wafer between $\approx 36$ V and $\approx 37$ V, corresponding to a field of $\approx 3.7 \times 10^5$ V·cm$^{-1}$. This is to be compared with breakdown fields of over $4 \times 10^5$ V·cm$^{-1}$ at room temperature in Ga$_{0.47}$In$_{0.53}$As P-I-N's grown by MBE (Cinguino et al., 1985), suggesting a lowering of the ionisation threshold in the present structure. The dark current at $\frac{1}{2} V_B$ is $\approx 20$ nA (corresponding to a current density $J \approx 1.5$ A·cm$^{-2}$), and at 0.9 $V_B$ is $\approx 1.5$ μA ($J \approx 10^2$ cm$^{-2}$). The slope of the characteristic increases at $\approx 25$ V, suggesting that the leakage current is dominated by thermal generation at low bias and by band-to-band tunnelling at high bias. This was confirmed by studying the temperature dependence of the leakage current (see section 9.2). The slope of the characteristic increases again above $\approx 30$ V due to avalanche multiplication.

The effect of the heterojunctions on the leakage current can be examined by comparing with values in MBE-grown Ga$_{0.47}$In$_{0.53}$As P-I-N diodes. Cinguino et al. (1985) reported dark current densities of $\approx 1$ A·cm$^{-2}$ at $\frac{1}{2} V_B$ and $\approx 3 \times 10^2$ A·cm$^{-2}$ at 0.9 $V_B$ in MBE-grown Ga$_{0.47}$In$_{0.53}$As P-I-N's. In a MQW structure the current generation rates are reduced by the presence of the wide band-gap layers where the carrier generation rate will be smaller, and are also reduced at low fields by carrier recombination in the wells. Deep levels associated with the heterointerfaces or with the Al$_{0.48}$In$_{0.52}$As layers can lead to
Figure 7.2. Current-voltage characteristic of D418 in the dark (lower curves) and under white light illumination (upper curves).
increased leakage currents. The similar values for the Ga_{0.47}In_{0.53}As P-I-N's suggests that the heterointerfaces do not have a strong effect on the leakage current.

Figure 7.2 show the characteristic in the dark (lower curve) and under unfocussed white-light illumination (upper curve). At low bias the photocurrent is reduced due to carrier recombination in the wells. The photocurrent is approximately constant between about -10 and -25 V, and increases for reverse bias > 25 V due to avalanche multiplication.

Figure 7.3 shows the reverse-bias C-V characteristic at room temperature. The capacitance increases at low bias (< 2 V) due to incomplete depletion of the I-region. The capacitance decreases slightly with bias, from ~1.9 pF at 0 V to ~1.7 pF at 30 V, due to depletion of the heavily-doped contact layers. There is some structure in the capacitance between 0 and 20 V. This may be due to dark-current carriers which are confined in the wells at low fields but emitted as the field increases, or possibly to depletion of quantum wells adjacent to the heavily-doped contact layers, which have become doped due to dopant diffusion. At higher fields the capacitance curve is flat, indicating that the quantum wells are fully depleted and there is no carrier accumulation in the wells. The sharp decrease in the measured capacitance at ~34 V is due to the dark current.

Figure 7.4 shows the photocurrent due to illumination of the top of the mesa with radiation from a He-Ne laser. This corresponds to injection of electrons into the high field region. The maximum multiplication observed is ~65. The inset shows the response as a focussed (~5 μm) laser spot is scanned across the active area, at a reverse bias of 20 V (unity gain) and 35 V (M ~ 10). The photocurrent is zero under the gold contact and is elsewhere constant to within
Figure 7.3. Capacitance-voltage characteristic of D418 at 300 K.
Figure 7.4. Photocurrent as function of reverse-bias for top illumination at 0.63 µm.

Inset: line scan of device at 20 V (Mx1) and 35 V (Mx10).
10 %, showing uniform multiplication without edge effects and microplasmas.

Figure 7.5 shows the room temperature quantum yield at different bias voltages (from Mohammed et al., 1985). As the reverse bias is increased from 0 to 10 V the absorption edge shifts to lower energy accompanied by an order of magnitude increase in the quantum efficiency. The shift is due to the quantum-confined Stark effect (see section 3.4). The low quantum efficiency at low bias is due to the large recombination rate of electrons and holes in the quantum wells compared to the emission rate of carriers out of the wells. The quantum efficiency is approximately constant between 10 and 25 V, since at fields above \( \approx 10^5 \text{ V} \cdot \text{cm}^{-1} \) most of the carriers are emitted from the wells and collected. The external quantum efficiency of \( \approx 50 \% \) at 1.0 \( \mu \text{m} \) and -20 V corresponds to approximately unity collection efficiency in this structure. Above 25 Volts the quantum yield increases due to impact ionisation.

The photocurrent at low bias shows features due to the step-like two-dimensional density of states in the quantum wells. The photocurrent at zero bias and \( T \approx 150 \text{ K} \) for the sample with 139 Å layers is shown on a linear scale in figure 7.6. Intersub-band transitions in the \( n = 1, 2 \) and 3 sub-bands are observed. As the field increase to \( 2 \times 10^5 \text{ V} \cdot \text{cm}^{-1} \), the photocurrent due to the \( n=1 \) transition decreases due to decreasing spatial overlap of the electron and hole wavefunctions (see figure 7.4). The excitonic peaks are not well resolved, presumably due to the presence of ionised impurities or other imperfections. The sample with 103 Å wide layers (figure 7.4; see also Capasso et al., 1986) shows plateaus corresponding to the \( n=1, 2 \) and 3 transitions and exciton peaks for the light hole and heavy hole \( n=1 \) excitons, and the heavy hole \( n=2 \) exciton. The features are in agreement
Figure 7.5. Quantum yield (external quantum efficiency times multiplication) against photon energy for devices with 139 Å and 103 Å layer widths (from Mohammed et al., 1985).
Figure 7.6. Photocurrent against photon energy for device with 139 Å layers at T \approx 150 \text{ K} and zero bias.
with the absorption data of Weiner et al. (1985) for Al$_{0.48}$In$_{0.52}$As /Ga$_{0.47}$In$_{0.53}$As MQW's with 110 Å wells.

7.3 Impact ionisation rates.

The multiplication at low frequency (≈ 1 kHz) was measured under illumination of the front and back of the device with 0.63 μm radiation. The former corresponds to injection of electrons into the depletion region; the latter to mixed injection of electrons and holes as discussed below. Low primary photocurrents (≈ 0.1 nA) were used to avoid perturbation of the applied field by photogenerated carriers in the wells. No intensity dependence was observed up to $I_{p0} ≈ 1 \mu A$. The photocurrent at room temperature under these illumination conditions is shown in figure 7.7. The photocurrent below 10 V is reduced by recombination in the wells as discussed previously. The photocurrent continues to increase between 10 and 20 V, particularly for electrons. This may be due to increased injection efficiency into the high-field region due to depletion into the P$^+$ contact layer, or to the presence of some trapping and recombination even at fields > $10^5$ V·cm$^{-1}$ (as observed in the capacitance data). The primary (unmultiplied) photocurrent $I_{p0}$ at higher bias was estimated by fitting a straight line to the photocurrent between 10 and 20 V (a reasonable fit could be obtained in this region). The multiplication is given by $M(V) = I_p(V)/I_{p0}(V)$. Ionisation rates were calculated using the equation for P-I-N diodes, assuming that back illumination gives rise to pure injection of holes. Depletion of the contact layers was not taken into account in calculating the field. The results are shown in figure 7.8 (dashed lines). The electron ionisation rate is about $10^4$ at $3.5 \times 10^5$ V·cm$^{-1}$. The ratio varies from $α/β = 2.5$ at $3 \times 10^5$ V·cm$^{-1}$ to 1.9 at $3.5 \times 10^5$ V·cm$^{-1}$. Note that the calculation of the rates by
Figure 7.7. Photocurrent for front and back illumination with 0.63 μm radiation at room temperature.
Figure 7.8. Ionization rates at 300 K,
(a) assuming pure injection of holes (dashed lines),
(b) mixed injection for back illumination, with
\[ \alpha_0 = 3.7 \times 10^4 \text{ cm}^{-1} \] (solid lines).
Allam et al. (1986) contained an algebraic error, and the conclusion $\alpha/\beta \approx 3 - 3.5$ is incorrect. The ratio is similar to that obtained for bulk Ga$_{0.47}$In$_{0.53}$As by Pearsall (1980).

As the temperature is decreased, the ionisation rates increase due to the decreasing phonon scattering. Figure 7.9 shows the photocurrent for back illumination at temperatures of 300, 210 and $\approx 90$ K, with the intensity adjusted to give a photocurrent of 0.1 nA at 20 V. The breakdown voltage decreases from $\approx 36$ to $\approx 33$ V. As the temperature decreases the photocurrent at low fields decreases. Extrapolation of figure 7.9 suggests that unity efficiency at zero bias will occur at a temperature of $\approx 500$ K, giving an estimated thermionic emission rate for holes of $2 \times 10^{10}$ s$^{-1}$. The low collection efficiency for holes at $10^5$ V-cm$^{-1}$ at 90K suggests that carriers in the wells are not strongly heated by the field. Approximately unity efficiency occurs at $\approx 2 \times 10^5$ V-cm$^{-1}$ at $T \approx 90$ K. The low efficiency at low fields shows that the capture rate into the wells is higher than expected. This might be due to quantum-mechanical reflection at the interfaces, or to the presence of undepleted wells which capture carriers injected from the contact layers more efficiently. It becomes more difficult to estimate a baseline at low temperatures since the onset of multiplication occurs at lower bias closer to the onset of unity efficiency. Note that the small increase in photocurrent for back illumination between -10 and -20 V at 300 K, and the small photocurrent between 0 and -10 V at 90 K, may be associated with electrons from mixed injection. Figure 7.10 shows the photocurrent for back and front illumination at $T \approx 100$ K. The electron multiplication shows a similar temperature dependence to the hole multiplication. The barrier for thermionic emission of electrons ($\approx 0.5$ eV) is higher than that for holes ($\approx 0.2$ eV). For low fields ($\lesssim 10^5$ V-cm$^{-1}$) and low
Figure 7.9. Temperature dependence of the photocurrent for back illumination.
Figure 7.10. Photocurrent against reverse-bias at $T \approx 100$ K.
temperatures (\(\lesssim 100 \text{ K}\)) the electron current is dominated by tunnelling. Features at \(\approx 2.5\) and \(\approx 7.5 \text{ V}\) are associated with sequential resonant tunnelling of electrons between the wells (see chapter 9). At high fields (\(\approx 3 \times 10^5 \text{ V-cm}^{-1}\)) the electron current may be dominated by Fowler-Norheim tunnelling with estimated emission rates of \(\approx 10^{11} \text{ s}^{-1}\).

The above calculation of the ionisation rates assumed that back illumination of the substrate with strongly absorbed radiation caused pure injection of holes. This is not correct. The substrate is several hundred microns thick and heavily doped, thus the diffusion length of holes is of the order of a few microns and essentially all the photogenerated holes will recombine with majority electrons before reaching the depletion edge. This is confirmed by the low quantum efficiency observed for back illumination. The recombination gives rise to radiation close to the band-gap of the substrate, which can be absorbed in the narrow-gap layers of the MQW region. This effect will be essentially independent of the bias, thus absorption of recombination radiation can give rise to mixed injection without the strong bias-dependence seen in the bulk due to Franz-Keldysh electroabsorption.

The ionization rates in the case of mixed injection due to absorption in the depletion region can be calculated numerically from the multiplication curves, if the absorption coefficient is known (Osaka et al., 1985). The absorption coefficient in the quantum wells is estimated as \(\alpha_0 \approx 3.7 \times 10^4 \text{ cm}^{-1}\), assuming that back illumination is equivalent to illumination with radiation of energy just below the InP band-edge. The results are shown in figure 7.8 (solid lines). The electron ionisation rate is little changed, but the hole rate is decreased. The ratio varies from \(\approx 8\) at \(3 \times 10^5 \text{ V-cm}^{-1}\) to \(\approx 3\) at
3.5 x 10^3. Thus it is clear that α/β in this structure, and that α/β ≥ 2. More precise conclusions cannot be drawn due to the mixed injection and the field-dependence of the primary photocurrent. Pure injection of holes could be achieved using selectively-etched wells in the substrate as described in chapter 5 or using the device structure suggested in chapter 10.

Capasso et al. (1981) measured α/β ≈ 8 in GaAs/Al_{0.45}Ga_{0.55}As MQW's with ≈ 500 Å layers. Due to the larger fraction of the band-gap discontinuity in the conduction band in Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As, the enhancement of α/β in this material system is expected to be larger. Furthermore, theoretical studies (see chapter 3) suggest that the ratio increases as the layer thickness is decreased to ≈ 100 Å. Note however that these studies do not include quantum confinement effects. Brennan (1987) has calculated the ionization rates in an Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As MQW APD with 250 Å layers. The results are shown in figure 7.11 compared with data for bulk Ga_{0.47}In_{0.53}As (Osaka et al., 1985; Pearsall, 1980) and Al_{0.48}In_{0.52}As (Capasso et al., 1984b). The experimental results for 139 Å layers presented in this chapter are not in good agreement with the theoretical results for 250 Å layers of Brennan.

7.4 High frequency response.

For the study of the pulse response the diodes were mounted on a 50 Ω microwave stripline coupled to a sampling scope as described in chapter 5. The source was a 1.3 μm semiconductor laser. The response at 36 V to a 100 ps laser pulse of intensity ≈ 300 μW is shown in the inset to figure 7.12. The FWHM is ≈ (230 ± 10) ps. The response time was unchanged for intensities down to ≈ 3 μW and pulse widths of ≈ 100 ns. Tails in the pulse response due to carrier pile-up in the
Figure 7.11. Ionization rates for Al$_{0.48}$In$_{0.52}$As (Capasso et al., 1984b), Ga$_{0.47}$In$_{0.53}$As (Osaka et al., 1985; Pearsall, 1980), Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW with 139 Å layers (experimental, this chapter) and Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW with 250 Å layers (theoretical, Brennan, 1987).
wells were not observed. The intrinsic response time at -36 V can be estimated by deconvoluting the laser pulse width ($\tau_L \approx 100 \text{ ps}$) and the contribution of RC time constants ($2.2 \text{ RC} = 187 \text{ ps}$ for $C \approx 1.7 \text{ pF}$) using a sum of the squares approximation. This leads to an effective transit time, which includes the avalanche build-up time, of the order of 90 ps.

The microwave avalanche gain was examined by plotting the pulse height as a function of reverse bias (figure 7.12). The microwave gain was measured at laser intensities between 1.5 mW and 12 $\mu$W and was unchanged. The efficiency is low up to $\approx 15 \text{ V}$. The reduced efficiency at $10^5 \text{ V cm}^{-1}$ compared to the d.c. photocurrent is due to carriers which are not emitted from the well within the duration of the pulse. From the ratio of the peak height at -36 V to the value at -20 V (approximately unity gain) a maximum microwave gain of $\approx 12$ is obtained. This implies a gain-bandwidth product in excess of 40 GHz. The d.c. gain at the same bias and the same illumination conditions was $\approx 30$. The gain-bandwidth product is given by equation 2.33

$$M(\omega)\omega = \frac{1}{\tau_1} \approx \frac{1}{N\gamma(\beta/\alpha)}$$

at high frequencies. $\tau$ is the transit time ($\approx L/v_{\text{Sat}} \approx 15 \text{ ps}$) and $N$ is approximated by (Kuvas and Lee, 1970)

$$(1/3)(1+\frac{1}{\ln(\alpha/\beta)})$$

Thus for $\alpha/\beta \approx 2$, $\tau_1$ is about 3 ps and the gain-bandwidth product is 50 GHz. The avalanche build-up time ($\tau_2$) is given by

$$\tau_2 = M(0)\cdot \tau_1$$

or about 90 ps, in agreement with the measured value.

These results show that high speed operation is possible at high fields. Carrier pile-up effects are not observed despite the 0.5 eV barriers for electrons. Thus MQW APD's can be used in microwave
Figure 7.12. Pulse height as a function of reverse-bias.
Inset: response of detector biased at -36 V to a 100 ps laser pulse.
receivers without the need for compositional grading of the heterojunction interfaces.
CHAPTER 8

IMPACT IONIZATION ACROSS THE BAND-EDGE DISCONTINUITY IN MULTIPLE QUANTUM WELL STRUCTURES.

8.1 Introduction.

In chapters 6 & 7, photocurrent multiplication in Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As and Ga$_{0.47}$In$_{0.53}$As/InP MQW's was discussed. The samples were grown with low background doping and exhibited low reverse-bias dark-current at room temperature. Avalanche multiplication in these devices showed the frequency- and temperature-dependence expected for band-to-band impact ionization.

Structures grown in the same material systems and with the same growth techniques, but under slightly non-optimum growth conditions, can exhibit much higher reverse-bias dark currents through defect-assisted thermal generation and tunnelling, whilst still being free of microplasmas and non-uniform breakdown caused by active large-scale defects. Photocurrent multiplication in such devices shows characteristics quite different from band-to-band impact ionization. The characteristics appear to be consistent with impact ionization across the band-edge discontinuity of dark-current carriers dynamically stored in the wells, a process called here "quantum-well impact ionization". This process is similar to ionization of deep levels, and the quantum wells can be regarded as artificial "traps".

This mechanism was mentioned by Smith et al. (1983), who studied the response to sub-bandgap radiation of a GaAs/Al$_{0.3}$Ga$_{0.7}$As MQW structure with heavily-doped wells (N \approx 3 \times 10^{19} \text{ cm}^{-3}). The photoresponse was due to free carriers in the wells being directly photoionized over the conduction-band discontinuity. The high
responsivity of 200 A/W resulted from photoconductive gain due to the large ratio of the lifetime of the photoexcited electron to the transit time. Quantum well impact ionization was not directly observed in this structure.

A structure with N-type doped wells and undoped barriers is shown in figure 6.1(a). If the barriers are sufficiently thick to prevent tunnelling at high fields, electrons from the donors will be confined in the wells, providing a reservoir of stored electrons. Under a high electric field, electrons will be heated in the barrier layers with enough energy on entering the well to impact ionize a stored electron over the conduction-band discontinuity, resulting in current multiplication. The threshold energy for this process is at least $(2\Delta E_C - E_F)$, where $E_F$ is the Fermi level in the well, since both carriers must escape the well for multiplication to occur. To maintain multiplication, the reservoirs in the wells would have to be replenished by supplying charge from electrical contacts to the wells. Since the ionization mechanism involves only one type of carrier, feedback of holes is avoided, resulting in low-noise multiplication. This makes the mechanism attractive for optical detectors, if the technological problems of applying selective contacts to quantum wells can be overcome. The ionization rates in such a structure have been calculated by Chuang & Hess (1986), and are discussed in section 6.2.

Similar ionization effects may also occur in undoped MQW structures which exhibit large leakage currents. The charge reservoirs arise from dark-current carriers such as thermally-generated electron-hole pairs (figure 6.1(b)), which are dynamically stored in the wells. Since both electrons and holes are stored, both carrier types are multiplied. The ionization rates are dependent on the threshold energies and on the density of stored carriers. The threshold energies
Figure 8.1. Impact ionization across the band-edge discontinuity
(a) in a sample with N-type doped wells,
(b) in an undoped sample with electrons and holes created by thermal generation.
are determined by the band-offsets, thus the rates for electrons and holes can be very different if the band offsets differ significantly. The density of stored carriers depends on the dark current, and also on the emission rate from the wells and therefore on the barrier widths, the band-offsets and the temperature. Photocurrent multiplication data for MQW structures with large leakage currents provides evidence for this ionization effect, as reported in section 8.3.

In section 8.4, single carrier-type multiplication is reported in multiple graded-well Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As avalanche photodiodes. The grading of the Ga$_{0.47}$In$_{0.53}$As-on-Al$_{0.48}$In$_{0.52}$As heterojunction interface means that, at high reverse bias, electrons are not confined in this structure, thus there is no multiplication of electrons by ionization across the discontinuity. Holes, on the other hand, are multiplied by the ionization of holes stored at the abrupt valence band hetero-interface, resulting in pure hole multiplication. This is the first observation of single carrier-type multiplication in a III-V semiconductor material.

8.2 MQW's with doped wells: theoretical calculation of ionization rates.

Chuang and Hess (1986; 1987) have calculated the ionization rates for impact ionization across the band-edge discontinuity in GaAs/AlGaAs MQW photodiodes with N-type doped wells. The wavefunctions of the electrons participating in the impact ionization process were represented by simple propagating or bound solutions. The quantum-mechanical transition rate was calculated from Fermi's Golden rule using a screened Coulomb potential for the electron-electron interaction. From the energy distribution of the incident hot electrons, which was assumed to be Maxwellian, the average ionization rate was calculated numerically. Values of the electron temperature in
the barriers as a function of field were found from Monte Carlo simulations. The ionization rate was strongly dependent on the well widths, the doping concentration in the wells and the value of the band-offset. The effect of tunnel-assisted ionization was included (Chuang and Hess, 1987) using an Airy function expression for the barrier transmission. This significantly increases the ionization rate for electrons with low energies. Ionization rates at $77 \text{ K}$ were calculated for a device with 50 periods of 200 $\text{Å}$ wells doped to $7 \times 10^{18} \text{ cm}^{-3}$ and 400 $\text{Å}$ $\text{Al}_{0.25}\text{Ga}_{0.75}\text{As}$ barriers. The ionization rates were $10^3 - 10^4 \text{ cm}^{-1}$ at fields of $(0.25 - 1) \times 10^5 \text{ V-cm}^{-1}$, and the multiplication was $\approx 10$ at $F \approx 1.5 \times 10^5 \text{ V-cm}^{-1}$.

### 8.3 MQW's with undoped wells: experimental evidence.

Photocurrent multiplication was studied in $\text{Al}_{0.48}\text{In}_{0.52}\text{As} / \text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ MQW structures of varying layer widths, as well as in $\text{GaSb/AlSb}$ and $\text{Ga}_{0.47}\text{In}_{0.53}\text{As/InP}$ MQW structures.

The $\text{Al}_{0.48}\text{In}_{0.52}\text{As}/\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ samples were grown by Cho and Sivco using MBE. The devices consisted of a nominally undoped multiple quantum well region sandwiched between 2 $\mu$m thick $\text{P}^+$ and $\text{N}^+$ $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ layers (except for sample D384 where the $\text{P}^+$ and $\text{N}^+$ regions were $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$). The layers were grown lattice-matched on an $\text{N}^+ <100>$ InP substrate. The MQW dimensions are shown in Table 8.1. The area of the mesa-etched devices was $1.3 \times 10^{-4} \text{ cm}^2$.

The capacitance was constant ($\approx 1.3 \text{ pF}$) as a function of reverse-bias when measured at $1 \text{ MHz}$ and $T \approx 90 \text{ K}$. Measurements at lower frequencies showed that the MQW region became fully depleted at a reverse-bias of $\approx 10 \text{ V}$, indicating a background carrier concentration of $\approx 10^{15} \text{ cm}^{-1}$. The constant capacitance at high frequencies and low temperatures is due to carrier "freeze-out" in the undepleted wells, as
Table 8.1 Dimensions of multiple quantum well samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Composition</th>
<th>No. of periods</th>
<th>Barrier width</th>
<th>Well width</th>
</tr>
</thead>
<tbody>
<tr>
<td>D418</td>
<td>Al$<em>{0.48}$In$</em>{0.52}$As/Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>35</td>
<td>139 Å</td>
<td>139 Å</td>
</tr>
<tr>
<td>D384</td>
<td>Al$<em>{0.48}$In$</em>{0.52}$As/Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>48</td>
<td>104 Å</td>
<td>104 Å</td>
</tr>
<tr>
<td>D620</td>
<td>Al$<em>{0.48}$In$</em>{0.52}$As/Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>20</td>
<td>200 Å</td>
<td>200 Å</td>
</tr>
<tr>
<td>D622</td>
<td>Al$<em>{0.48}$In$</em>{0.52}$As/Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>20</td>
<td>465 Å</td>
<td>230 Å</td>
</tr>
<tr>
<td>D676</td>
<td>Al$<em>{0.48}$In$</em>{0.52}$As/Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>20</td>
<td>450 Å</td>
<td>450 Å</td>
</tr>
<tr>
<td>A064</td>
<td>AlSb/GaSb</td>
<td>25</td>
<td>200 Å</td>
<td>200 Å</td>
</tr>
<tr>
<td>25FebG6</td>
<td>InP/Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>10</td>
<td>500 Å</td>
<td>300 Å</td>
</tr>
</tbody>
</table>

discussed in section 9.3.

Figure 8.2 shows the reverse-bias current-voltage characteristic for sample D622 at T $\approx$ 85 K, with white light illumination (upper curve) and the in dark (lower curve). Up to $\approx 5$ V, the dark current is approximately constant at $\approx 10^{-11}$ A. The dark current increases almost exponentially from $\approx 10^{-11}$ A at 5 V to $10^{-6}$ A at 20 V, and is $\approx 10^{-8}$ A at 12 V. The photocurrent (the difference between the current with and without illumination) is almost constant at low bias, but starts to increase above $\approx 12$ V.

It should be noted that there was some variation of the dark current from device to device on the same wafer; in particular, it is difficult to ascertain the contribution from surface leakage currents. There is also variation between devices of the photocurrent at low bias: although most devices show fairly flat baselines, some others show structure such as peaks or steps. These may be associated with tunnelling through defects in the barriers as reported by Capasso et al. (1986d). However, there is greater uniformity across the wafer for
Figure 8.2. Reverse-bias current-voltage characteristic for sample D622 at $T \approx 85$ K.
the multiplication characteristics than for the baselines.

In figure 8.3 the quantum yield is plotted as a function of photon energy for top illumination of samples D384 and D622. Sample D620 showed near-identical behavior to D622. For sample D622 the efficiency increases between zero and -10 V due to increased collection efficiency from the wells. The response at -10 V is in agreement with estimates of the maximum external quantum efficiency for this structure, which is limited by absorption and recombination in the top layer and by the thickness of the intrinsic layer. Thus the increase in the quantum yield as the reverse-bias is increased to 20 V is due to a multiplication process. As the bias is increased the quantum yield increases more rapidly at longer wave-lengths, showing that holes are multiplied significantly more than electrons. The largest gain was observed in sample D384. Two representative curves are shown in figure 8.3. The quantum yield exceeds 10 for \( \lambda < 1.35 \mu m \) at -18.6 V.

Multiplication at high fields is characteristic of an impact ionization process determined by a threshold energy. The onset of multiplication occurs at an electric field of \( \approx 8 \times 10^4 \) V-cm\(^{-1}\). This is about a factor of two smaller than the field at which impact ionization across the band-gap occurs in bulk Ga\(_{0.47}\)In\(_{0.53}\)As at the same temperature, indicating a smaller threshold energy. The multiplication factors of electrons and holes were found by measuring the photocurrent-voltage curve for illumination of the P\(^{+}\) and N\(^{+}\) sides of the mesa with 0.63 \( \mu m \) radiation. The results are shown in figure 8.4 for a diode of sample D622 at \( T \approx 80 \) K. By attenuating the laser beam, the primary electron and hole photocurrents before the onset of multiplication were made equal (\( I_{po} \approx 150 \) pA). The increase of the photocurrent as the reverse-bias is increased from zero to 2.5 V is due to the increased collection efficiency from the wells. The multipli-
Figure 8.3. Quantum yield for samples D384 and D622.
Figure 8.4. Photocurrent as a function of reverse bias for sample D622 for electron and hole injection.
cation factor for holes is greater than that for electrons. This is consistent with the smaller threshold energy for quantum well impact ionization of holes \((2\Delta E_V \approx 0.4 \text{ eV})\) compared to electrons \((2\Delta E_C \approx 1.0 \text{ eV})\) in \(\text{Al}_{0.48}\text{In}_{0.52}\text{As}/\text{Ga}_{0.47}\text{In}_{0.53}\text{As}\).

In the general case of carrier multiplication in the presence of trapping it is not possible to derive ionization coefficients for electrons and holes from the measured electron- and hole-initiated multiplication factors (Landsberg and Robbins, 1975; Robbins, 1980). This is because the ionization rates depend on the electron and hole carrier densities and therefore on the injection conditions. However, calculation of the ionization rates using the standard P-I-N formula serves as an indicator of the order of magnitude of the ionization rates ratio for comparison with other impact ionization devices. These "effective" ionization rates are shown in figure 8.5. The ratio \(\beta/\alpha\) is \(\approx 20\) at a field of \(1.5 \times 10^5 \text{ V/cm}^{-1}\), corresponding to a hole multiplication \(M_h \approx 4.5\). This value is roughly twice the ionization rate ratio enhancement previously reported in MQW APD's employing band-to-band ionization (see chapter 3) and is in the opposite sense \((\beta > \alpha)\).

The multiplication was strongly temperature-dependent. As the temperature was increased, the dark current and the multiplication at a given field increased (see figure 8.17, below). The onset of multiplication also shifted to lower voltages. This shows the opposite temperature dependence to that observed for band-to-band impact ionization. The temperature dependence can be explained by the increasing generation rate of carriers in the well at high temperatures.

The frequency response was measured by plotting the photocurrent recorded by the lock-in as the chopping frequency was slowly swept from
Figure 8.5. Effective ionization rates for electrons and holes in D622.
10Hz to 5kHz at a given voltage. The results are shown in figure 8.6 for 0 V, -10 V (unity gain) and for -21 V. The frequency response at zero bias is related to carrier transport in the undepleted region of the MQW, as discussed in chapter 9. At -10 V, the MQWs are fully depleted and the photocurrent is virtually independent of frequency (up to 5 kHz). The photocurrent in the presence of gain (at -21 V) is strongly frequency dependent with a time constant of $\approx 10^{-3}$ s.

The frequency-dependence for quantum well ionization in samples with undoped wells is due to the finite density of carriers in the reservoirs. Ionization causes depletion of the wells, which must be refilled by the dark current generation mechanism. The multiplication will be determined by the rate of depletion of the reservoirs by ionization and the refilling rate. In the experiment chopped square wave illumination was used. At low frequencies the charge reservoir can recover if the time for which the light is off is long compared to the refilling time. Increased multiplication is expected if the mark/space ratio of the chopped light is reduced, since depletion of the wells by ionization is reduced while the recovery time is increased. Further experimental work is required to investigate this. For thermal generation of carriers, the generation rate is strongly dependent on the trap density and energy (equation 2.12), and can be slow at low temperatures. However, this time is expected to be strongly temperature dependent and to vary from wafer to wafer, depending on the deep level concentration. The similar time constant measured for different material systems and different temperatures (see below) suggests that the time constant may be determined by some structural property such as an RC time constant. The equivalent circuit of a doped well structure is a series of parallel capacitors and resistors, where the capacitance is given by the thickness of the barrier layers and the resistor
Figure 8.6. Frequency dependence of the photocurrent in sample D622.
corresponds to the effective resistance of the dark current generation mechanism. The RC product for a 1 pF capacitance and a $10^9$ Ω resistance (leading to a dark current of 10 nA at 10 V bias) is $\approx 10^{-3}$ s, which is close to the time constant measured in all the structures studied.

The multiplication was strongly dependent on the intensity of the illumination, as shown in figure 6.7. The intensity was varied to give primary photocurrents between 10 pA and 100 nA. At the highest intensity the photocurrent at low bias is reduced since a higher applied field is required to deplete the photogenerated carriers. However the multiplication is strongly intensity dependent even at very low primary photocurrents ($\approx 10$ pA). The intensity dependence can also be explained by the finite carrier reservoir in the wells. Ionization across the heterojunction will deplete the reservoirs, and thus reduce the number of carriers available for scattering with hot carriers from the barrier layers. The multiplication is thus reduced at high primary photocurrents.

Another sample (D676) showed similar multiplication effects at room temperature. The photocurrent at 300 K for illumination of the top $P^+ \text{Ga}_0.47\text{In}_0.53\text{As}$ layer is shown in figure 8.8. Illumination at $\lambda = 1.5 \, \mu m$ is not completely absorbed in the top layer and thus corresponds to mixed injection, showing that holes are multiplied more than electrons. The onset of gain occurred at a field of $\approx 10^5$ V-cm$^{-1}$ at room temperature and at $\approx 3 \times 10^5$ V-cm$^{-1}$ at a temperature of $\approx 100$ K. The intensity and frequency dependence were similar to those for sample D620.

The observed multiplication characteristics suggest that the current multiplication is due to impact ionization of deep levels or of quantum wells. Homojunction $\text{Al}_0.48\text{In}_0.52\text{As}$ and $\text{Ga}_0.47\text{In}_0.53\text{As}$ as well as single heterojunction $\text{Al}_0.48\text{In}_0.52\text{As}/\text{Ga}_0.47\text{In}_0.53\text{As}$ detectors, grown
Figure 8.7. Intensity dependence of the photocurrent for sample D622.
Figure 8.8. Photocurrent of sample D676 at 300 K.
under the same conditions as the MQW samples, were also studied, but similar multiplication effects were not observed. The density of deep levels required to observed significant multiplication can be estimated as follows. The ionization rate for holes ($\beta$) can be expressed as the product of the probability for interaction between hot holes and holes in the reservoir and the probability of a hot hole having an energy in excess of threshold. For a Maxwellian distribution with a hot hole temperature $T_{h,b}$ in the barrier layers:

$$\beta = \sigma p \exp \left[ \frac{-E_{th}}{k_B T_{h,b}} \right]$$

(9.4)

where $p$ is the hole density in the wells and $\sigma$ is an effective cross section. For a screened Coulomb interaction, the interaction distance is of the order of $1/q$, where $q$ is the dielectric screening parameter which is equal to the reciprocal of the Debye length for a Maxwell-Boltzmann distribution (Chuang and Hess, 1986). Thus for hole concentrations of $10^{15} - 10^{16} \text{ cm}^{-3}$ the interaction length is of the order of a few hundred Angstroms. This gives ionization rates of the order of $10^4 \text{ cm}^{-1}$ at high fields where the exponential term approaches unity. For deep level ionization the range of the central cell potential is approximately $10 - 20 \text{ A}$ (Hess, 1986). Thus the effective cross section for ionization across the band discontinuity is expected to be about two orders of magnitude greater than that of deep level ionization. In order to observe ionization rates of $10^4 \text{ cm}^{-1}$ due to deep level ionization, an unrealistically high concentration of deep levels ($\sim 10^{17} \text{ cm}^{-3}$) would be required.

Similar multiplication effects were also observed in structurally similar samples of different material systems: AlSb/GaSb grown by Malik using MBE and Ga$_{0.47}$In$_{0.53}$As/InP grown by Panish using gas source
MBE. The AlSb/GaSb P+-I-N+ structures consisted of 25 periods with 200 Å layer widths, sandwiched between a P+ Al$_{0.40}$Ga$_{0.60}$Sb window layer and an N+ GaSb buffer. The photocurrent at T $\approx$ 90 K for front and back illumination with a He/Ne laser is shown in figure 8.9. The onset of gain occurred at similar electric fields as that for sample D622 for devices with similar dark current characteristics, with holes multiplied more than the electrons. The intensity and frequency dependence are also similar to those for D622. The band-gaps in GaSb/AlSb are close to those in Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As and $\Delta E_C > \Delta E_V$. Thus the similar multiplication characteristics are strong evidence that the multiplication is due to a structural effect rather than ionization of deep levels. The two peaks observed in the photocurrent for top illumination are not understood, although they are reminiscent of the resonant tunnelling peaks observed in D418 (see section 9.1).

Multiplication at low fields was also observed in the Ga$_{0.47}$In$_{0.53}$As/InP MQW structures. The unintentionally-doped MQW region consisted of 10 periods of 300 Å wells and 500 Å barriers. The structure is described in more detail in section 9.3. The top P+ layer was InP, so that top illumination with 0.63 µm radiation injected electrons whereas 0.95 µm radiation was absorbed in the quantum wells leading to mixed injection. The photocurrent at 100 K under these conditions is shown in figure 8.10. The increasing quantum efficiency for illumination at 0.95 µm is due to collection of photogenerated carriers as the wells, which are doped to about 5 x 10$^{16}$ cm$^{-3}$ by background impurities, are successively depleted (see section 9.3). The figure shows that electrons are multiplied significantly, although given the uncertainties in the injection it is not clear whether electrons or holes are multiplied more. In this material system the
Figure 8.9. Photocurrent as a function of reverse bias for sample A064 (GaSb/AlSb MQW) for
(a) back illumination,
(b) front illumination.
Figure 8.10. Photocurrent as a function of reverse bias for 
InP/Ga\textsubscript{0.47}In\textsubscript{0.53}As MQW structure.
conduction-band discontinuity is less than that for holes, thus electrons might be expected to be multiplied more by impact ionization across the discontinuity. The frequency-dependence of the electron photocurrent is shown in figure 8.11. The decreased response at low fields and higher frequencies is related to the emission time of electrons from the undepleted wells (see section 9.3). The time constant for the multiplication is again of the order of $10^{-3}$ s. The density of deep levels in this sample was measured by Lang and Sergent (1986), and was found to be less than $5 \times 10^{14}$ cm$^{-3}$. Thus the multiplication can be explained by quantum well rather than deep level ionization.

The ionization rates for quantum well ionization are strongly dependent on the density of carriers in the wells, as shown in the calculations of Chuang and Hess (1986). This is shown in the experimental results by the correlation between the multiplication and the dark current. The onset of multiplication occurs at similar values of the dark current in different wafers. For wafers (such as D418) in which the dark current in the same voltage range is small (of the order of nA at room temperature), the above multiplication effects are not observed and band-to-band impact ionization occurs at significantly higher fields (see chapter 7).

The density of carriers dynamically stored in the wells can be estimated as follows. Only multiplication of holes is considered, i.e. feedback due to electron multiplication is neglected. Consider at first the current in the absence of illumination and at a field lower than that required for the onset of quantum well impact ionization. In steady state, the sum of the carrier generation rate and the capture rate into the $i$th well is equal to the sum of the recombination rate and the thermal emission rate out of the well. The dark-current density
Figure 8.11. Frequency response of photocurrent for InP/Ga_{0.47}In_{0.53}As MQW photodiode.
$J_{d,i}$ emitted at the $i^{th}$ well is written as the product of a charge density ($L_W n_i^* e$), an attempt frequency ($v_{th}/L_W$) and the probability of thermal emission ($\exp\left(-\Delta E/k_B T\right)$). Thus the carrier density dynamically stored in the well is given by

$$n_i = \frac{J_{d, i}}{e \cdot v_{th}} \exp\left[\frac{\Delta E}{k_B T}\right]$$

(9.1)

where

$$v_{th} = \sqrt{\frac{k_B T}{2\pi m}}$$

is the average thermal velocity perpendicular to the layers.

It can be seen from equation 9.1 that large carrier densities in the wells can be obtained if the band offsets are sufficiently larger than the carrier thermal energies. The carrier temperature in the wells is considerably smaller than in the absence of the barrier for the same value of the electric field, as long as the field does not exceed appreciably $10^5 \text{ V-cm}^{-1}$. Evidence for this was found in the quantum efficiency and high speed response of MQW APD's in the previous chapters. Carrier densities of the order of $10^{16} \text{ cm}^{-3}$ can be attained at temperatures of $\approx 100 \text{ K}$ for dark currents of 10 nA.

The photocurrent multiplication at each well is given by $(1+\xi_1)$, where $\xi_1$ is the yield, which depends on the carrier density in the well, the well width, the threshold energy and the distribution of hot carriers in the barrier layers. The total multiplication is thus given by

$$M = \prod_{i=1}^{N} (1 + \xi_i).$$

(9.3)

Neglecting carrier capture, depletion of the reservoirs by ionization
and feedback due to multiplication of electrons, the carrier concentration and thus the yield is the same in all the wells. For a 20 period device with a multiplication of 6, the yield is \( \approx 0.1 \).

Equation 9.1 does not include the effect of quantum well impact ionization initiated by dark current carriers. This will tend to deplete the carrier reservoir. Assuming that the photocurrent is negligible compared to the dark current, the rate of carriers ionized out of the reservoir in the \( i \)th well is given by \( J_{d,i} \cdot e_i/e \), where \( J_{d,i} \) is the dark current density in the \( i \)th barrier. Thus the number ionized in a time window of \( 10^{-3} \) s for a dark current of 1 nA is of the order of \( 10^{15} \) cm\(^{-3} \).

The effect of carrier capture into the wells will be to increase the carrier concentration in the reservoirs. However, the constant photocurrent as a function of reverse bias at fields below multiplication suggests that either the capture rate or the recombination rate in the wells is low. In the absence of recombination in the wells, all the dark current carriers will eventually be released from the well and there can be no d.c. multiplication by the impact ionization process discussed here.

Another current multiplication mechanism which should be considered was described by McIntyre (1986). This is related to perturbation of the field by the photogenerated carriers. This increases the emission rate from the wells, leading to an increased current. Due to the near-exponential nature of the I-V characteristics observed in these samples a small increase in the field can cause a large increase in the total current. The perturbation of the field is approximately \( \Delta F = Q/\varepsilon sA \), where \( Q \approx I_p \tau \) is the photogenerated space-charge in a time window \( \tau \) (the thermal release time). For a photocurrent of \( \approx 1 \) nA and an emission rate of \( 10^4 \) s\(^{-1} \) (estimated for
emission of holes over a 0.2 eV barrier with a carrier temperature of 120 K) the field perturbation is of the order of 10^3 V·cm⁻¹, or about 1% of the field at which multiplication is observed. It is not clear if this effect is important in our samples since the carrier temperatures in the well and the emission rates are not well known. The multiplication for this effect is expected to increase with intensity, since the perturbation of the wells increases. This is not in agreement with the experimental results presented above, where the largest multiplication is observed at primary photocurrents as low as 10 pA.

6.4 Single Carrier-Type Multiplication in Multiple Graded-Well Structures.

This section discusses current multiplication in multilayer APD's with compositionally-graded regions at the interfaces between the low band-gap and high band-gap materials. The Al₀.₄₈In₀.₅₂As/Ga₀.₄₇In₀.₅₃As structures exhibited single carrier-type multiplication of holes over certain bias ranges. The multiplication characteristics are similar to those described in the previous section and can be explained by impact ionization across the band-gap discontinuity. These structures cannot act as staircase photomultipliers as described by Capasso et al. (1983) since the conduction-band discontinuity is significantly less than the threshold energy for ionization of electrons in the lower band-gap material.

The band-structure of a multiple graded-well APD under operating bias is shown in figure 8.12. At zero bias, electrons are confined in the conduction-band wells by the abrupt heterojunction discontinuity on one side and by the quasi-field in the graded region on the other side. Under sufficient reverse bias, the applied field becomes greater than the quasi-field and the electrons are no longer confined. Holes, on the other hand, will be confined at the abrupt valence-band discontinuity.
Figure 8.12. Impact ionization of holes across the band-edge discontinuity in a multiple graded-well structure.
Thus impact ionization of holes across the valence-band discontinuity may occur, leading to single carrier-type multiplication of holes.

The multiple graded-wells were positioned in the nominally undoped region of a P⁺-I-N⁺ photodiode structure. The structures were grown by Alavi and Cho using computer-controlled MBE as described in section 4.3.4. Four wafers with varying multilayer dimensions and either 3 or 5 periods were grown (R397, R407, R408 & R409). The dimensions are given in Table 8.2. The graded regions consisted of Ga₀.₄₇₋ₓAlₓIn₀.₅₃As linearly graded from x=0 to x=0.47. A final graded region was capped with a 2 µm P⁺ contact, doped to \( \approx 7 \times 10^{18} \, \text{cm}^{-3} \). Sample R397 had a wide band-gap Al₀.₄₈In₀.₅₂As P⁺ top window layer; the other samples all had Ga₀.₄₇In₀.₅₃As P⁺ layers. A Sn-doped N⁺ Al₀.₄₈In₀.₅₂As buffer layer (10¹⁸ cm⁻³) was grown on the substrate, followed by a 0.4 µm nominally undoped Al₀.₄₈In₀.₅₂As layer, to prevent diffusion of Sn from the substrate into the quantum wells during subsequent growth. Surface segregation of Sn in GaAs, leading to non-abrupt doping profiles, has been reported for growth temperatures > 550°C (Cho, 1975). The structures were analysed by S. M. Abys using SIMS to investigate the extent of Sn penetration. The SIMS data for sample R409 is shown in figure 8.13. The 5 periods are clearly observed in the Ga trace. The Sn trace shows there is negligible penetration into the multilayers, as expected at the growth temperature of 540°C.

Mesa devices of area 1.3 x 10⁻⁴ cm² were fabricated and ohmic contacts were formed. Windows were opened in the back metallization to allow illumination of the substrate. No photocurrent multiplication was observed at room temperature due to excessive leakage current (\( \approx 10 \, \mu\text{A} \) at a reverse bias of 2.5 V). To enable further characterization the samples were cooled to \( T \approx 90 \, \text{K} \). Figure 8.14 shows the reverse-bias current-voltage characteristics of R407 in the dark and with broad area
Figure 8.13. SIMS data and structure for sample R409.
Figure 6.14. Reverse bias current-voltage characteristic of R407 at 90 K, in dark (lower curves) and with white light illumination (upper curves).
Table 8.2. Dimensions of multiple graded-well samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Top layer</th>
<th>No. of periods</th>
<th>Barrier width</th>
<th>Well width</th>
<th>Graded width</th>
</tr>
</thead>
<tbody>
<tr>
<td>R397</td>
<td>Al$<em>{0.48}$In$</em>{0.52}$As</td>
<td>5</td>
<td>700 Å</td>
<td>300 Å</td>
<td>700 Å</td>
</tr>
<tr>
<td>R407</td>
<td>Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>3</td>
<td>501 Å</td>
<td>292 Å</td>
<td>1022 Å</td>
</tr>
<tr>
<td>R408</td>
<td>Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>3</td>
<td>720 Å</td>
<td>292 Å</td>
<td>720 Å</td>
</tr>
<tr>
<td>R409</td>
<td>Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>5</td>
<td>292 Å</td>
<td>292 Å</td>
<td>720 Å</td>
</tr>
</tbody>
</table>

illumination from a tungsten lamp. The photocurrent is low at zero bias but increases with reverse bias up to \( \approx 2 \) V. The photocurrent is approximately constant from 2 to 6 V. At higher fields the photocurrent increases sharply, and the quantum yield (external quantum efficiency \( \times \) multiplication) exceeds 70\% close to breakdown. This indicates significant photocurrent multiplication. The dark current in the multiplication region increases from \( \approx 1 \) nA to \( \approx 1 \) \( \mu \)A.

The electron and hole multiplication factors were estimated by measuring the photocurrent as a function of bias under different illumination conditions. For electron injection, a He-Ne laser was focussed on to the top of the mesa. To obtain hole injection, the He-Ne laser was focussed on to the top of the substrate close to the edge of the mesa or on to the back window of the mesa; these methods gave essentially the same results. Mixed injection could also be achieved by focussing 1.0 \( \mu \)m light from a monochromator on to the mesa top. In all cases, the primary photocurrent (I$_{po}$) was kept small (\( \approx 0.1 \) nA) to avoid perturbation of the band-structure by the photogenerated carriers. The photoresponse of the devices was spatially uniform to within about 15\% at reverse bias close to breakdown, showing the absence of significant edge response and microplasmas.
Figure 8.15 shows the hole- and electron-initiated photocurrent at a chopping frequency of 200 Hz and a temperature of 90 K, for sample R407. For hole injection, multiplication occurs at a reverse bias of 7 V and reaches $M_h \approx 20$ at 12 V. For electron injection, the multiplication ($M_e$) is negligible at reverse bias up to 10 V ($M_e < 1.02$ as verified by careful inspection of the electron multiplication curve on an expanded scale). Thus virtually single carrier-type multiplication of holes occurs between 7 and 10 V ($B/\alpha$ is at least several hundred at 10 V). At higher reverse bias, the electron multiplication increases to a value of $\approx 1.4$ at 12 V, resulting in an ionization rate ratio, estimated by $(M_h^{-1})/(M_e^{-1})$, in excess of 50.

The gain is strongly dependent on the chopping frequency of the illumination, with a time constant of the order of $10^{-3}$ s. As the temperature is increased, the onset of multiplication occurs at lower bias. The low field at the onset of multiplication, and the observed dependence on the temperature, intensity and frequency, show that band-to-band ionization is not responsible for the multiplication, and the characteristics can be explained by ionization across the band-edge discontinuity as for the abrupt-interface samples described in the previous section. The maximum gain for single carrier-type multiplication in sample R407 would be 8 ($\approx 2^3$, corresponding to the 3 periods). This corresponds to our observation of single carrier-type multiplication of holes in the reverse-bias range 7 to 10 V, with a multiplication at 10 V of $\approx 3.7$. The yield (fraction of carriers ionizing per stage, $\epsilon$) can be calculated from the expression for the gain of a photomultiplier

$$M_h = (1+\epsilon)^N$$

where $N$ is the number of stages. At -10 V, prior to the onset of significant electron multiplication, $\epsilon \approx 0.55$. The yield strongly
Figure 8.15. Photocurrent as a function of reverse bias for R407 for injection of electrons and holes.
depends on the density of holes stored at the heterointerface, but the high value may also be due to the large width of the hole accelerating region (barrier plus graded region) and the quasi-field in the graded region.

At reverse-bias greater than 10 V, the small electron multiplication ($<1.4$ at 12 V) can cause significant feedback effects. This allows the hole multiplication to exceed the limit of 8 (maximum measured gain $\approx 50$ at 10 Hz and 12 V). The onset of electron multiplication occurs at a reverse field of $\approx 2 \times 10^5$ V-cm$^{-1}$, which is similar to the field at which band-to-band ionization occurs in bulk Ga$_{0.47}$In$_{0.53}$As. Note also that the band-to-band ionization of electrons may be enhanced by the conductio-band discontinuity. Thus it is suggested that single carrier-type multiplication of holes occurs via impact ionization over the band-edge discontinuity, and the feedback is accounted for by band-to-band impact ionization of electrons.

The other wafers show similar behavior. Fig 8.16 shows the photocurrent for R409 at $T \approx 100$ K. The lower curve is for short-wavelength light shone on the mesa top, and corresponds to injection of electrons. There appears to be slight peak in the photocurrent at $\approx -3$ V. This may correspond to a small amount of electron multiplication due to ionization of carriers stored at the heterointerface before the field is high enough to invert the structure. This occurs at a field of $\approx 7 \times 10^4$ Vcm$^{-1}$ (a bias of a few Volts). There is also a small increase in the photocurrent for reverse bias above $\approx 8$ V, possibly due to band-to-band ionization as discussed above. The upper part of figure 8.16 shows the multiplication with top illumination with radiation of 1.0 $\mu$m wavelength, which causes mixed injection of electrons and holes. Again, holes are multiplied much more than electrons.
Figure 8.16. Photocurrent as a function of reverse bias for R409

(a) mixed injection of electrons and holes,

(b) electron injection.
The temperature-dependence of the reverse-bias current-voltage characteristics for R408 is shown in figure 8.17, at temperatures of 100, 150, 200 and 300 K. The lower curve shows the dark current, and the upper curve is for broad area white light illumination. At 100 K, the onset of multiplication occurs at a reverse-bias of \( \approx 5 \) V, corresponding to a dark current of \( \approx 10^{-10} \) A; as the temperature increases, the onset voltage decreases to \( \approx 4 \) V (\( \approx 5 \times 10^{-9} \) A) at 150 K, and \( \approx 2 \) V (\( \approx 10^{-7} \) A) at 200 K. The dark current at the onset of gain increases with temperature due to the increasing rate of thermal emission from the well and thus the reduction in carrier pile-up at high temperatures. The photocurrent for sample R397 is shown in figure 8.18, for top illumination with 1.4 \( \mu \)m radiation. This sample has an Al\(_{0.48}\)In\(_{0.52}\)As top layer, thus absorption occurs in the quantum wells. The steps in the photocurrent are due to successive depletion of the first three wells, as discussed in chapter 9. The background doping in the multiple graded-wells is estimated to be \( \approx 8 \times 10^{15} \) cm\(^{-3}\) for this sample. The multiplication increases with temperature. For band-to-band ionization, the multiplication decreases with temperature as shown in figure 7.9.

8.4 Conclusions and recent further work.

Experimental evidence for impact ionization across the band-edge discontinuity was presented in this chapter. The performance of the present devices are limited by high dark-current, and low speed. The noise is expected to be dominated by the dark-current noise. For a low-noise device, the carrier reservoirs would have to be replenished by means of electrical contacts to the individual wells, rather than via the dark current. A schematic diagram of a "three terminal" APD is shown in figure 8.19. The shaded regions are the wells, which
Figure 8.17. Temperature dependence of reverse-bias current-voltage characteristic for R408.
Figure 8.18. Temperature dependence of photocurrent for sample R397.
Figure 8.19. Structure of a three terminal solid-state photomultiplier.
correspond to the dynodes of a photomultiplier. The potential drop across the barriers is maintained by a resistive network. A possible method for implementing this is shown in figure 8.19(c). A narrow band-gap side-contact is regrown on the mesa edge. The contact is lightly-doped N-type with a similar resistivity to the barrier layers. Under suitable bias injection of electrons from the contact into the wells will cause replenishing of the reservoirs, while injection into the wide-gap barriers will have a lower efficiency.

These results also have implications for the realization of MQW and staircase photomultipliers employing band-to-band ionization, where it will be necessary to suppress ionization over the discontinuity which occurs at lower fields. This could be achieved in devices with low dark current, or in the case of staircases by using a staggered (type II) structure where no pile-up of carriers at the interface can occur in reverse bias (Capasso et al., 1984c).

Further evidence for impact ionization across the band-edge has recently been found by other workers. Levine et al. (1987) observed multiplications of $\approx 3$ in GaAs/Al$_{0.36}$Ga$_{0.64}$As MQW's with 70 Å N$^+$ wells. The photocurrent was initiated by intersub-band absorption, with responsivities of $\approx 7$ A W$^{-1}$ at a wavelength of 10.3 μm. Yu et al. (1987) have studied Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW APD's with characteristics similar to those described above. Analysis of noise measurements showed that $\beta/\alpha \approx 6$. At fields $\gtrsim 2 \times 10^5$ V cm$^{-1}$ band-to-band ionization occurred and $\beta/\alpha$ decreased towards unity.
CHAPTER 9

PERPENDICULAR TRANSPORT MECHANISMS IN MULTIPLE QUANTUM WELL STRUCTURES.

In this chapter, some aspects of perpendicular transport in MQWs are discussed. Resonant tunnelling was investigated in Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW's with high quality interfaces and low background doping. Sequential resonant tunnelling between conduction-band quantum wells, as first observed by Capasso et al. (1986b), is discussed further in section 9.1. At higher fields, tunnelling between bound states in valence- and conduction-band quantum wells (resonant Zener tunnelling) was observed, as reported in section 9.2.

The reverse-bias current-voltage and capacitance-voltage characteristics of Ga$_{0.47}$In$_{0.53}$As/InP MQW's are discussed in section 9.3. The carrier concentration in the wells was $\approx 5 \times 10^{16}$ cm$^{-3}$ due to background doping. Perpendicular transport through the quantum wells was dominated by the behaviour of carriers localised in the wells in the undepleted region.

9.1 Sequential resonant tunnelling in multiple quantum wells.

Resonant tunnelling in double barrier structures was briefly described in section 3.5. Multilayer structures which have been studied include superlattices (with narrow barriers and coupled wells) and multiple quantum well structures (tight-binding superlattices with uncoupled wells). Negative differential conductance (NDC) in superlattices due to electron transfer into the negative mass regions of the miniband was predicted by Esaki and Tsu (1970). This effect, and the associated Bloch oscillations, have not been observed experimentally. NDC was observed (Esaki et al., 1972) in a GaAs/Al$_{0.5}$Ga$_{0.5}$As superlattice with 70 Å periods, associated with the transition from
miniband conduction to hopping which occurs when the potential drop across one period exceeds the miniband width. A different kind of multiple NDC was observed in GaAs/AlAs 85 Å period superlattices with 10^{17} \text{cm}^{-3} doping (Esaki and Chang, 1974). The current oscillations were related to resonant tunnelling between adjacent wells which occurred in a high field domain which propagated through the superlattice as the field increased.

In a multiple quantum well with weak coupling between the wells and low background doping, an applied field will drop uniformly across the MQW region, rather than forming high field domains. NDC in such a structure was predicted by Kazarinov and Suris (1971; 1972) due to resonant tunnelling between the ground state of a well and an excited state of an adjacent well. Resonance occurs when the potential drop across a period (eF_{A}) is equal to the difference between the ground state energy (E_{g}) and the excited state energy (E_{n}). The electron in the excited state relaxes to the ground state and the resonant tunnelling proceeds sequentially though all the periods of the MQW if the field is uniform. This sequential resonant tunnelling (SRT) was observed experimentally by Capasso et al. (1986b) in an Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As MQW P^{+}-I-N^{+} diode with 35 periods of 139 Å layer widths. The structure was described in detail in chapter 7 (sample D418). The photocurrent for top illumination with He-Ne radiation (i.e. injection of electrons into the depletion region) was measured as a function of reverse-bias at low temperature. Their results are shown in figure 9.1, together with the associated transitions. A tunnelling resonance calculation showed that the two peaks correspond to tunnelling into E_{2} and E_{3}.

In this section, sequential resonant tunnelling is studied further. Sample D418 was processed into mesa devices of area
Figure 9.1. Sequential resonant tunnelling into the first two excited states in an Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW diode (from Capasso, Mohammed and Cho, 1985).
1.3 \times 10^{-4} \text{ cm}^2 as described in chapter 7. The mesas were passivated with Merck HTR3 polyimide for physical protection; this had a negligible effect on the dark currents. The current-voltage characteristic is shown in figure 9.2 in the dark (lower curve) and with white light illumination (upper curve), measured on an HP4145 parameter analyser. The features in the reverse-bias photocurrent at \( \approx 2 \text{ V} \) and \( \approx 6.5 \text{ V} \) correspond to the SRT transitions observed by Capasso et al. A further feature (shown below to be due to SRT into \( E_4 \)) is apparent at \( \approx 11 \text{ V} \). These features are not observed in the dark current since under reverse bias the dark current, due to thermal generation in the depletion region, is lower than the detection capabilities of the experimental set-up (\( \approx 1 \text{ pA} \)). SRT was observed in the photocurrent at temperatures up to about 100 K (see figure 7.10). At higher temperatures, phonon-assisted tunnelling and thermal emission over the barrier dominate. The barrier height is given by \( \Delta E_C - E_4 \), or about 0.4 eV at \( 10^5 \text{ V} \cdot \text{cm}^{-1} \). Features in the reverse-bias dark current at \( \approx 24 \text{ V} \) and \( \approx 29 \text{ V} \) are associated with resonant-Zener tunnelling as discussed in the section 9.2.

The forward-bias dark current shows a large amount of structure associated with different tunnelling transitions. Under forward-bias majority carriers are injected from the contacts into the depletion region when the bias exceeds the built-in potential \( V_{bi} \) (flat-band condition). In a homostructure P-N junction, the forward-bias current increases very rapidly with bias. In this MQW structure, however, the dark current is \( \approx 10^{-4} \text{ A} \) at a forward bias of 15 V at \( T = 8 \text{ K} \). This is due to the high capture and recombination rate in the wells and the low thermal emission rate at low fields, as observed in the low quantum efficiency for the reverse-bias photocurrent below 15 V.

The forward-bias characteristics were studied by mounting the
Figure 9.2. Current-voltage characteristic of device D418 at 8 K.
devices onto sapphire pads, bonding gold wire to the contacts using the In ball bonding technique described in chapter 5, and cooling the devices to 4 K in a Janis He cryostat. The DC characteristics were measured on a Tektronix curve tracer. The results for one device are shown in figure 9.3. The same features were observed in all 4 devices studied in detail. At forward bias greater than about 10 V, hysteresis effects were observed and the characteristics were measured for both increasing and decreasing bias. The features observed are as follows.

(i). A forward current flows when the bias is equal to the built-in potential (figure 9.3(a)). The built-in potential for electrons in thus ≈0.9 V. This is closer to the band-gap of Ga0.47In0.53As than of the Al0.49In0.52As which forms the P+ and N+ contact layers.

(ii). Two peaks are observed at forward biases of (3.8±0.2) V and (7.5±0.5) V (the errors represent the spread of data for the 4 devices measured). These correspond to fields of ≈3 x 10^4 V·cm^-1 and ≈7 x 10^4 V·cm^-1. The field was estimated from F = (V-V_b)/L, where L is the width of the MQW region.

(iii). A further peak (feature X) was observed at (12.1±0.4) V for increasing bias and (12.6±0.3) V for decreasing bias. A step in the current (feature Y) was observed at (13.4±0.3) V for increasing bias and (14.1±0.2) V for decreasing bias (see figure 9.3 (d, e, g and h)).

(iv). A peak (feature Z) was observed at a bias of (15.0±0.3) V, for both increasing and decreasing bias. The peak is more pronounced for increasing bias (figure 9.3(h)).

(v). Two regions of hysteresis are observed, between ≈12 and ≈13.5 V and between ≈14 and ≈15 V.

The two peaks at 3 and 7 x 10^4 V·cm^-1 are in reasonable agreement with the peaks observed in the reverse-bias photocurrent at 2.2 V (F ≈ 3 x 10^4 V·cm^-1) and 6.6 V (F ≈ 8 x 10^4 V·cm^-1) and are associated
Figure 9.3. Forward-bias characteristic of D418 at 4 K.
Figure 9.3. (Continued).
Figure 9.3. (Continued).
with SRT from \( E_1 \) into \( E_2 \) and \( E_3 \). It is tempting to associate the peaks X and Z with SRT into higher energy levels in the wells (\( E_4 \) and \( E_5 \)). The step Y may be associated with tunnelling into the continuum or into a barrier resonance, since this will result in an increase in the current due to reduced capture into the well.

In order to identify the X, Y, and Z features, the energy levels in the wells were calculated from a numerical solution of the one-dimensional Schroedinger equation for the envelope function

\[
-\frac{\hbar^2}{2m^*} \frac{d^2\psi(z)}{dz^2} = (E-V(z)) \psi(z) \tag{9.1}
\]

The continuity conditions at the interfaces are that \( \psi(z) \) and \( (1/m^*(E))(d\psi(z)/dz) \) are continuous, in accordance with the envelope function approximation (Bastard, 1981). The coupling between the wells is weak (due to the relatively thick barriers) and the levels in the MQW are assumed to have the same energy and width as in the case of a single well. Perturbation of the field by charge accumulation in the wells was not considered.

Solutions for the energies were obtained using a tunnelling resonance technique. A three-layer test structure was used as shown in figure 9.4. The width, band-edge energy, effective mass and non-parabolicity of each layer and of the end layers were specified. The transmission for a carrier tunnelling coherently through the structure at an energy \( E \) was calculated by applying the continuity conditions at each interface. Both double barrier and single barrier simulations were used (figure 9.4(b)). The applied field was dropped across the three central layers. The linear potential due to the field was approximated by dividing each layer into constant-potential slices (figure 9.4(c)). Numerical tests showed that 10 slices per layer were sufficient. The zero of energy for the double barrier is taken as the band-edge at the
Figure 9.4. Double and single barrier structures used in tunnelling resonance calculation.
centre of the well. The transmission of the double barrier shows peaks due to the bound states in the well and the resonances in the barriers. The single barrier simulation is used to identify the barrier resonances. The zero of energy as shown in figure 9.4(b) gives the same energy scale for single and double barrier simulations.

Layer widths of 139 Å and effective masses of 0.041\textit{m}_0 for Ga_{0.47}In_{0.53}As and 0.076\textit{m}_0 for Al_{0.48}In_{0.52}As were used, together with a conduction-band offset of 0.5 eV. The non-parabolicity of the bands was taken into account using an energy dependent effective mass

\begin{equation}
m^*(E) = m^*(0)(1 + a_1E)
\end{equation}

in equation 9.1, where \(a_1\) is the non-parabolicity factor. Three band \(k\cdot\pi\) theory gives a non-parabolicity of

\begin{equation}
a_1 = \left(1 - \frac{m^*}{m_0}\right)\frac{2}{E_g}
\end{equation}

which is 1.13 eV\textsuperscript{-1} for Ga_{0.47}In_{0.53}As and 0.55 eV\textsuperscript{-1} for Al_{0.48}In_{0.52}As at 4 K. Experimental measurements (Shantharama, 1986) showed that the non-parabolicity in Ga_{0.47}In_{0.53}As is rather larger, given by 1.81 eV\textsuperscript{-1} for Ga_{0.47}In_{0.53}As compared to 0.64 for InP. These values were used in the simulation.

Figure 9.5 shows the transmission of a double (solid line) and single (dashed line) barrier structure as a function of energy at zero field. The five peaks in the transmission correspond to the bound levels in the wells. The peak transmission is unity since the left and right barriers are symmetrical. This can be verified using a higher energy resolution. The resonances at energies greater than 0.5 eV are virtual states associated either with the well (e.g. \(E_6\)) or barrier layers (e.g. \(E'_1\) or \(E'_2\)).
Figure 9.5. Transmission of double barrier (solid line) and single barrier (dashed line) at zero field.
The transmission at a field of $10^5$ V cm$^{-1}$ is shown in figure 9.6. At this field $E_5$ is no longer bound and is thus considerably broadened. There is also strong interaction between bound levels in the well and the barrier resonances $E'_1$ and $E'_2$, as shown in figure 9.6 by the coupling of $E_5$ and $E'_2$.

The energy levels as a function of field are shown in figure 9.7. Of the bound levels, $E_4$ is the most strongly perturbed by the field and shows a quadratic Stark shift. The dashed lines show the barrier resonances. The shift of the barrier resonances with field is mainly due to the shift of the conduction band-edge at the centre of the barrier with respect to that at the centre of the well. Clear anticrossing effects are observed between $E_5$ and $E'_1$ and $E'_2$, indicated by the dotted lines. Interaction of the barrier resonances with lower energy levels is smaller due to the triangular barrier, which is of width $\approx 60$ Å at the field for interaction of $E'_1$ and $E_4$ and $\approx 15$ Å for interaction of $E'_2$ and $E_4$. The solid line increasing with field shows the energy at which electrons can tunnel into the well from the ground state of an adjacent well. The solid symbols therefore represent sequential resonant tunnelling. The calculated and experimental fields for tunnelling are compared in table 9.1. The calculations support the allocation of peaks X and Z to tunnelling into $E_4$ and $E_5$, and suggests that the current step Y is related to tunnelling into the barrier resonance $E'_2$.

The effect of varying the band-offsets and the non-parabolicities was also studied. The calculated fields for the tunnelling transitions considered above are shown in table 9.2 for values of the Ga$_{0.47}$In$_{0.53}$As conduction band non-parabolicity of 0, 1.70 eV$^{-1}$ and 2.26 eV$^{-1}$. The figures in bold type show the closest agreement with the experimental data. The non-parabolicity appears to be greater at
Figure 9.6. Transmission of double and single barrier at a field of $10^5$ V·cm$^{-1}$. 
Figure 9.7. Energy levels in double barrier as a function of field.
energies of ≈ 0.5 eV than that measured at low energies, as expected from the curvature of the bands.

The dependence on the conduction band discontinuity is shown in table 9.3 for $\Delta E_c$ equal to 0.44 eV (Weiner et al., 1985), 0.50 eV (People et al., 1983) and 0.53 eV (Sugiyama et al., 1986). As expected, the bound levels are only weakly dependent on the band offset. The
Table 9.3. Calculated fields for SRT for different band-offsets.

<table>
<thead>
<tr>
<th>Transition</th>
<th>Calculated field $(10^5 \text{ V/cm}^{-1})$</th>
<th>Experimental field $(10^5 \text{ V/cm}^{-1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta E_C$ (eV)</td>
<td></td>
</tr>
<tr>
<td>$E_1$ to $E_2$</td>
<td>0.28 0.29 0.29</td>
<td>0.3</td>
</tr>
<tr>
<td>$E_1$ to $E_3$</td>
<td>0.67 0.69 0.69</td>
<td>0.7</td>
</tr>
<tr>
<td>$E_1$ to $E_4$</td>
<td>1.11 1.13 1.14</td>
<td>(1.15 - 1.2)</td>
</tr>
<tr>
<td>$E_1$ to $E_5$</td>
<td>1.54 1.59 1.62</td>
<td>1.45</td>
</tr>
<tr>
<td>$E_1$ to $E'_1$</td>
<td>1.07 1.22 1.27</td>
<td>(1.3 - 1.35)</td>
</tr>
<tr>
<td>$E_1$ to $E'_2$</td>
<td>1.27 1.39 1.49</td>
<td></td>
</tr>
</tbody>
</table>

barrier resonances however, since they are measured with respect to the conduction band-edge in the wells, vary significantly. Thus the step Y might be associated either with tunnelling into $E'_1$ ($\Delta E_C \approx 0.53$ eV) or into $E'_2$ ($\Delta E_C = 0.45$ eV).

Hysteresis and bistability have been observed in double barrier devices, and interpreted as intrinsic effects due to the feedback of the electrostatic field from space-charge regions (Goldman et al, 1987). The problem can be tackled using a self-consistent approach (Berkowitz and Lux, 1987). However, hysteresis can also arise due to an internal series resistance (Tsuchiya and Sakaki, 1986). In the MQW device studied in this section, the hysteresis is observed close to the field range for coupling between well and barrier resonances. Tunnelling into the barrier resonances will have an effect on the charge accumulation in the wells due to the decrease in the capture rate. Work is in progress to understand the effect of the coupling on the charge in the wells and on the hysteresis.

The above results show that $\text{Al}_{0.46}\text{In}_{0.52}\text{As}/\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ is an
interesting material system for tunnelling experiments. The large conduction-band barrier reduces background current from thermoionic emission, while the light effective mass and large non-parabolicity in Al$_{0.48}$In$_{0.52}$As (compared to Al$_{0.3}$Ga$_{0.7}$As) enhance the tunnelling probability. The strength of the barrier resonances is dependent on the ratio of effective masses in the barriers and wells, which is $\approx 1.8$ in Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As and $\approx 1.4$ in Al$_{0.3}$Ga$_{0.7}$As/GaAs.

9.2 Resonant Zener Tunnelling.

At high reverse bias, the dark current of these Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW diodes exhibits features due to resonant Zener tunnelling (RZT). Narrow band-gap semiconductor P-N junctions exhibit large leakage currents at high reverse-bias due to band-to-band tunnelling (see chapter 2). In a MQW structure, the tunnelling current is modified by the quantisation of the energy levels in the wells. Significant tunnelling current will only occur at fields such that there is overlap in energy between conduction and valence subbands.

The devices were mounted, contacted and cooled as described in the previous section. Figure 9.8 shows the reverse-bias current-voltage characteristics at temperatures between 4 K and 240 K. At high temperatures, thermally-generated electron-hole pairs dominate the dark current at low bias. Above a reverse bias of about 20 V the current increases strongly due to tunnelling of electrons across the band-gap. At a reverse bias of $\approx 35$ V, avalanche breakdown occurs. As the temperature decreases, the rate of thermally-generated electron-hole pair's decreases strongly, whereas the tunnelling component is only weakly temperature dependent. Two inflections are observed in the tunnelling current. These occur at reverse bias of $\approx 24$ V and $\approx 28.5$ V at the lowest temperatures, and at slightly higher bias as the
Figure 9.8. Reverse-bias current-voltage characteristic of D418 at various temperatures. The current below 20 V for the 4 K trace is due to stray leakage paths.
temperature increases. This corresponds to fields of \( \approx 2.6 \) and \( \approx 3.0 \times 10^5 \) V cm\(^{-1}\).

At these fields the potential drop across one period of the MQW is of the order of the band-gap, which suggests resonant tunnelling across the band-gap between bound states in the valence- and conduction-band quantum wells. Tunnelling (at the zone centre) will occur when the potential drop across a period \( (|e|F_a) \) is equal to the separation of the electron sub-band \( (E_n^{(e)}) \) and the hole sub-band \( E_m^{(h)} \):

\[
|e|F_a = E_g + E_m^{(h)} + E_n^{(e)} \tag{9.4}
\]

where \( E_g \) is the band-gap of Ga\(_{0.47}\)In\(_{0.53}\)As. This tunnelling process will occur simultaneously in all 35 periods of the MQW if the field is uniform. The electron-hole pair created escape from the wells by tunnel-assisted thermal emission and are collected by the contacts, giving rise to the measured leakage current.

The energy levels in the quantum wells were calculated at fields of 2.6 and 3.0 \( \times 10^5 \) V cm\(^{-1}\), corresponding to the experimental current steps. The electron levels (table 9.4) were calculated using the tunnelling resonance technique discussed above, with a conduction-band offset of 0.5 eV and a non-parabolicity of 1.70 eV\(^{-1}\). The heavy-hole levels (table 9.5) were calculated assuming parabolic bands and a valence-band offset of 0.2 eV, under the following conditions:

1. Tunnelling resonance calculation with \( m_{hh} = 0.465 \) \( m_0 \) (Ga\(_{0.47}\)In\(_{0.53}\)As) and \( m_{hh} = 0.8 \) \( m_0 \) (Al\(_{0.48}\)In\(_{0.52}\)As),

2. Tunnelling resonance calculation with \( m_{hh} = 0.6 \) \( m_0 \) (Ga\(_{0.47}\)In\(_{0.53}\)As) and \( m_{hh} = 0.6 \) \( m_0 \) (Al\(_{0.48}\)In\(_{0.52}\)As),

3. Infinite triangular well approximation (equation 3.8) with \( m_{hh} = 0.465 \) \( m_0 \) (Ga\(_{0.47}\)In\(_{0.53}\)As).
Table 9.4. Calculated electron energy levels for RZT.

<table>
<thead>
<tr>
<th>Transition</th>
<th>Field ($10^5$ V·cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td>$E_1^e$ (meV)</td>
<td>-22</td>
</tr>
<tr>
<td>$E_2^e$ (meV)</td>
<td>105</td>
</tr>
</tbody>
</table>

Table 9.5. Calculated heavy-hole energy levels for RZT.

<table>
<thead>
<tr>
<th>Transition</th>
<th>Field ($10^5$ V·cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td>$E_1^{hh}$ (meV)</td>
<td>(1) -106</td>
</tr>
<tr>
<td></td>
<td>(3) -88</td>
</tr>
<tr>
<td>$E_2^{hh}$ (meV)</td>
<td>(1) -40</td>
</tr>
<tr>
<td></td>
<td>(3) -23</td>
</tr>
</tbody>
</table>

Table 9.6. Calculated light-hole energy levels for RZT.

<table>
<thead>
<tr>
<th>Transition</th>
<th>Field ($10^5$ V·cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td>$E_1^{lh}$ (meV)</td>
<td>-57</td>
</tr>
</tbody>
</table>

The effect of the mass variation is small, but the infinite well calculation overestimates the confinement energy since there can be significant penetration of the wavefunction into the barrier due to the high fields and low barrier. The light-hole levels were estimated (table 9.6) assuming parabolic bands and masses $m_{lh} = 0.0503m_0$. 
The band-gap of Ga$_{0.47}$In$_{0.53}$As at 4 K is taken as 0.821 eV. The calculated values of the field for tunnelling from $E_1^{(hh)}$ into $E_1^{(e)}$ (shown in figure 9.9(a)) and from $E_1^{(hh)}$ into $E_2^{(e)}$ (figure 9.9(b)) are 2.5 and 2.8 x 10$^5$ V cm$^{-1}$ respectively. Observation of tunnelling into higher electron sub-bands is obscured by avalanche breakdown. The fields for tunnelling from the light-hole states are approximately 2.7 and 3.0 x 10$^5$ V cm$^{-1}$. The experimental values of 2.6 and 3.0 x 10$^5$ V cm$^{-1}$ therefore lie between these two sets of calculated values. Light holes are expected to have a higher tunnelling probability due to their lighter mass. A more accurate calculation of the valence-band energy levels, including light-hole non-parabolicities and band-mixing effects, is required, together with a more accurate calculation of the field in the structure, including depletion of the contact layers.

The change with temperature in the measured RZT fields is due to the temperature-dependence of the band-gap (equation 4.2). Thus the estimated value of the bias for RZT into $E_2^{(e)}$ changes by ≈2.5 V between 4 K and 240 K. The RZT features become less clear at high temperatures due to the increasing contribution to the total dark current from thermally-generated electron-hole pairs (however, the second step is still visible in the derivative of the dark current at room temperature). The tunnelling is in competition with thermal emission over a barrier of $E_g + E_1^{(h)} + E_1^{(e)}$, or ≈0.7 eV. Thus RZT is
Figure 9.9(a) RZT from the lowest valence sub-band to the first conduction sub-band.

(b) RZT into the second conduction sub-band.
observed at higher temperatures than SRT where the barrier is \( \approx 0.4 \) eV.

As a result of energy and momentum conservation, SRT can only occur when the bottoms of the sub-bands of adjacent wells are coincident in energy. However, in RZT, lateral momentum \((k_\perp)\) can be conserved at all values of the field greater than that required to match \(E_{1(e)}\) and \(E_{1(h)}\). This is due to the different dispersion relations for the conduction- and valence-band bound states, as shown schematically in figure 9.10. The left side of the figure shows the dispersion relation for heavy holes in the \(m=1\) sub-band, and also the dispersion for the conduction-band barrier layer. The total height of the tunnelling barrier at a particular value of the lateral momentum \((B(k_\perp))\) is the vertical height between these curves. The right-hand side shows the \(n=1\) and \(n=2\) conduction sub-bands in an adjacent well.

The heavy-hole sub-band of the adjacent well is superimposed (dashed line) for clarity. In (a), the applied field causes the heavy-hole \(m=1\) energy level to line up with the \(n=1\) electron level. Lateral momentum is conserved at the zone centre and tunnelling can occur. The barrier height is shown by \(B(k_\perp=0)\). At a higher bias (b), lateral momentum can be conserved at a particular value of \(k_\perp\neq 0\). The barrier height is increased to \(B(k_\perp=0) + (\hbar^2 k_\perp^2/2m^*)\), where \((1/m^*)=(1/m_{hh})+(1/m_e)\). At higher bias, tunnelling is possible into both the \(n=1\) (at large \(k_\perp\)) and \(n=2\) sub-bands.

Although RZT can occur at all values of \(k_\perp\), the tunnelling probability away from the zone centre is greatly reduced by the increase in the effective barrier height, by a factor (Moll, 1964)

\[
\exp \left[ \frac{-\epsilon_\perp}{\epsilon_o} \right] \quad \text{where} \quad \epsilon_o = \frac{\sqrt{2|e|FM}}{2m^*V^\frac{1}{2}} \tag{9.5}
\]

Thus the tunnelling probability is significant only for carriers within
Figure 9.10. Conservation of lateral momentum for RZT. On the left is the in-plane dispersion relation for the heavy-hole sub-band, and the conduction band in the barrier layers. The right hand side shows the first two conduction sub-bands, at different values of the applied field.
a few meV of the zone centre. The tunnelling current is expected to consist of a series of peaks at applied fields such that zone-centre tunnelling between bound states can occur. The peaks will be strongly asymmetric due to the decreasing barrier width with increasing field. This explains the higher tunnelling current for the second experimental peak, since the barrier width \(E_g/(eF)\) decreases from 342 Å (\(n=1\)) to 283 Å (\(n=2\)) while the barrier height is roughly constant.

When the potential drop across one well exceeds \(\Delta E_V\), there are regions in the narrow-gap layer which are not confined by the valence-band barrier. Tunnelling from these regions into \(E^+(e)\) can occur at reverse-bias greater than \(\approx 26\) V, giving rise to the bulk-like background tunnelling current superimposed on the RZT peaks.

9.3 Transport in an Undepleted Ga\(_{0.47}\)In\(_{0.53}\)As/InP MQW structure.

In this section, carrier transport at room temperature in an undepleted Ga\(_{0.47}\)In\(_{0.53}\)As/InP MQW photodiode structure is studied. Steps in the photocurrent, dark current and capacitance are associated with charge collection by successive depletion of the quantum wells.

The structures were grown by Fanish and Sumski using gas source MBE (see section 4.3.5). The MQW’s were grown in the unintentionally doped N-region of a P\(^+\)-N-N\(^+\) structure. A 1 μm N\(^+\)-InP buffer layer was grown on the Sn-doped InP substrate, followed by a highly doped 1.7 μm Ga\(_{0.47}\)In\(_{0.53}\)As absorption layer and a 2500 Å Ga\(_{0.47}\)In\(_{0.53}\)As undoped spacer layer. The latter was intended to prevent diffusion of dopants into the MQW’s from the substrate. The MQW’s consisted of 10 periods of 300 Å Ga\(_{0.47}\)In\(_{0.53}\)As wells and 500 Å InP barriers. The top p-type contact was a 1.5 μm InP window layer to allow absorption of long-wavelength light by the MQW. The layers were processed into mesa devices of 100 μm diameter by chemical etching with Br-Methanol, and
ohmic contacts were made by evaporation and alloying as described in chapter 5. The structures are shown schematically in figure 9.11. The sides of the mesa were coated with Merck HTR3 polyimide for passivation and physical protection. The processed devices exhibited a soft breakdown with a dark current of 10 µA at a reverse bias of 10 V at room temperature.

The diodes were illuminated with white light from a microscope lamp filtered through a 0.95 µm long pass filter, to ensure carriers are photogenerated in the narrow band-gap quantum wells. Figure 9.12(a) shows the room-temperature current as a function of reverse bias $V_r$, with and without illumination. The photocurrent increases in steps as the reverse bias is increased. Six steps are seen before the dark current exceeds the photocurrent and obscures any further structure. The increase in photocurrent at each step is approximately constant. These steps are equally spaced when the photocurrent is plotted against $(V_r + V_{bi})^{1/2}$ (figure 9.12(b)). $V_{bi}$ is the built-in potential ($\approx$0.65 V), measured by plotting the square of the depletion region width $W(V_r)$ against $V_r$ on a Miller feedback depletion profiler. This clearly shows that the steps occur at multiple increments of the depletion width, which is proportional to $(V_r + V_{bi})^{1/2}$ for a single-sided abrupt junction. Each increment corresponds to the superlattice period. This indicates that the steps are caused by the electric field punching through into successive wells. Similar results were obtained when 1.55 µm light from a monochromator was focussed onto the mesa top with the photocurrent measured using a current-sensitive amplifier followed by a lock-in amplifier. The position of the steps did not change with light intensity (up to $I_{p0} = 100$ nA) or with chopping frequency. Figure 9.13(a) shows the photocurrent at a chopping frequency of 1 kHz and a light intensity of 4.25 nW. When the sample is illuminated with light
Figure 9.11. Schematic diagram of InP/Ga$_{0.47}$In$_{0.53}$As MQW photodiode.
Figure 9.12. Reverse-bias current-voltage characteristic of InP/Ga$_{0.47}$In$_{0.53}$As MQW photodiode.
Figure 9.13. (a) Photocurrent, (b) dark current and (c) free carrier density of InP/Ga$_{0.47}$In$_{0.53}$As MQW photodiode, plotted on a common reverse-bias scale.
from a He/Ne laser (which is absorbed in the top InP layer) no structure is observed and the photocurrent is constant with bias up to breakdown. A similar step-like structure is observed in the dark current, with the first three steps being observable at corresponding biases (figure 9.13(b)).

The carrier concentration profile was measured using a Miller feedback depletion profiler (Miller, 1972) (figure 9.14). As the reverse bias is increased from 0 to 10 V, the depletion width increases from 0.2 μm to 0.7 μm and the measured free carrier concentration oscillates between $2 \times 10^{16} \text{ cm}^{-3}$ and $8 \times 10^{16} \text{ cm}^{-3}$ with a period of 800 Å, equal to that of the MQW. The 7 peaks observed correspond to the 300 Å wide wells. The method of depletion profiling measures the free carrier density via the movement of free carriers into and out of the depletion layer as an a.c. voltage modulation is applied. The spatial resolution is limited by Debye screening with a characteristic length given by

$$L_D = \sqrt{\frac{k_B T \varepsilon_S}{e^2 N}}$$

(9.6)

which is $\approx 200$ Å at a doping of $5 \times 10^{16} \text{ cm}^{-3}$. This smears out the charge distribution so that abrupt changes in the free carrier distribution cannot be resolved to better than several Debye lengths (Johnson and Panoussis, 1971). This smearing is of the order of the well widths in our MQW, hence the high and low free carrier densities in figure 9.14 do not represent the true densities in the wells and barriers. In order to obtain these values it would be necessary to simulate the profiling measurement using a numerical solution of Poisson's equation, and to vary the density to give the best fit to experimental data (Whiteaway, 1983). The measured profile shows that the free carriers in the barriers spill over into the wells, similar to the process which
Figure 9.14. Measured carrier density profile of InPGa0.47In0.53As MQW photodiode.
occurs in modulation doping.

The measurement of the depletion profile shows that the current transport of electrons in the undepleted wells (by thermionic emission) is fast compared to the 1 MHz measurement frequency (Lang et al., 1986). If the sample is illuminated with white light, the peaks increase due to photogenerated holes which are trapped in the undepleted quantum wells.

The photocurrent steps can be explained as follows. In a P-N junction in bulk material, minority carriers generated outside the depletion region can be collected by the field, if they are generated within a diffusion length ($\approx$ a few microns) of the depletion edge. In heterostructure devices, diffusion currents can be suppressed by trapping of carriers at the heterojunction interfaces. This is shown in the low quantum efficiency at low bias in SAM-APD's.

The band diagram of the MQW diode is shown in figure 9.15, in the ideal case in which all the free carriers are in the wells and the barriers are depleted by carrier transfer into the wells. Electron-hole pairs will be generated in the wells by thermal processes and by photoionization across the band gap. In the undepleted region of the MQW, the carriers are localized in the wells due to the potential barriers. Thus holes generated in this region cannot diffuse to the depletion region and contribute to the current, but instead recombine with electrons. The depletion length at zero bias is $\approx$0.1 $\mu$m for a carrier concentration of $5 \times 10^{16}$ cm$^{-3}$. Thus the first period is depleted at zero bias. At the point when the depletion edge punches through into the second well (figure 9.15(a)), carriers generated in the well can be collected by the field: the effective width of the absorbing region is thus two wells. As the bias is increased and the depletion edge sweeps through the second well, the photocurrent and the thermal generation
Figure 9.15. Band-structure of partially-depleted MQW photodiode.
component of the dark current (which dominates the total dark current at low bias) remain constant (figure 9.15(b)), until the depletion edge reaches the next barrier. Since the barrier has no free charges, the depletion edge immediately punches through into the next well (figure 9.14(c)), where it can 'collect' the charge previously localized in the well. The suppression of the diffusion current by localization gives rise to the current steps since the absorption length ($\sim 1 \mu m$) is significantly greater than the depletion width in the bias range 0 to $-10 \, V$. For the case of electron injection into the depletion region the quantum efficiency is constant with bias, since the rate of transport through the undepleted region is faster than the rate of recombination with minority holes.

In our data the photocurrent does not increase abruptly as soon as punch-through into the well occurs. Figure 9.13 shows the correlation between the structure in the photocurrent and dark current plotted as a function of bias and the doping profile plotted at a corresponding depletion depth. The features in the current can be seen to coincide with the periods of the MQW. The photocurrent is constant while the depletion edge is sweeping through a given barrier because carriers are only photogenerated in the wells (note that a finite voltage is required to sweep through the barriers since not all the carriers have transferred into the wells). However, when the well is reached the current increases gradually with the field until it levels off at some point within the well. As the chopping frequency is increased, the current increase becomes more gradual and a higher field is required for it to level off. This suggests that the efficiency of collection of carriers from the well is reduced by recombination processes within the well. In the absence of recombination, all the carriers generated in a given well will be collected as soon as the depletion edge punches
through into that well. In practice, to achieve unity collection efficiency, the average field in the well must be sufficiently high that the rate of thermionic emission over the barriers exceeds the recombination rate.

Above about -5 V, the room temperature dark current is dominated by other mechanisms (surface leakage, tunnelling, etc.) so the structure is no longer observed. For the photocurrent measurements, structure is obscured above ≈-8 V by the presence of gain effects caused by quantum well impact ionization (see chapter 8).

Transport in this structure was studied further by Lang et al. (1987) using the technique of admittance spectroscopy (Losee, 1975). The capacitance and conductance at zero bias were measured as a function of frequency and temperature. A step was observed in the capacitance. At low frequencies and high temperatures, the capacitance measured was $\varepsilon_s A / W_{HT}$, where $A$ is the junction area and $W_{HT}$ is the width of the depleted region. This requires that the response time of the current flowing in the undepleted region of the quantum wells is fast on the time scale of the measurement frequency. At low temperatures or high frequencies, the capacitance was given by $\varepsilon_s A / W_{LT}$, where $W_{LT}$ is the width of the entire MQW region. Transport in the quantum wells is frozen out and the entire MQW is effectively an insulator. The step in the capacitance occurs when the measurement frequency is equal to the RC time constant, where $C$ is the capacitance of the depletion layer and $R$ is the series resistance which is determined by thermally-activated transport in the undepleted region. Lang et al. measured the activation energy for both P and N-type samples, and deduced a value of the conduction-band offset $\Delta E_C = (0.42 \pm 0.02) \cdot \Delta E_g$. This is in good agreement with values obtained by other workers, as discussed in section 4.2.2.
CHAPTER 10

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER WORK.

This thesis has reported experimental investigations of electrical transport in multiple quantum well structures. The results have a number of implications for the design of heterojunction photodiodes and for the understanding of transport in heterojunction devices.

10.1 Performance of long-wavelength MQW avalanche photodiodes.

It was shown in chapters 6, 7 and 9 that high quality multiple quantum wells in Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As and InP/Ga$_{0.47}$In$_{0.53}$As can now be grown by a number of different epitaxial techniques. Photodiodes fabricated from these structures can exhibit low leakage currents comparable to bulk Ga$_{0.47}$In$_{0.53}$As diodes. These devices have a high room-temperature responsivity in the near-infrared due to avalanche gain. The pulse response at high fields is not limited by carrier trapping at the heterojunctions, and a gain-bandwidth product of $\approx 40$ GHz was measured in the Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As devices. The devices are thus suitable for use in high bit-rate detector systems, without the need for compositional grading of the heterojunction interfaces. Further work may clarify the mechanisms for emission from the well at high fields, and the role of carrier-carrier scattering.

MQW APD's are attractive because of the enhancement of the ionization rates claimed in such structures. The enhancement was not observed in this work because of the difficulty of making accurate measurements in our structures. Several suggestions for carrying out these measurements are made in the following section.
The performance of these devices as detectors for use in optical communications systems is limited mainly by the dark current, which arises from Zener tunnelling at high fields. This problem appears to be unavoidable in diodes with Ga$_{0.47}$In$_{0.53}$As high field regions, and may limit the potential applications of MQW APD's fabricated in these material systems.


A number of problems occur in the measurement of band-to-band impact ionization rates in MQW APD's, in addition to those encountered for bulk materials. The following points should be taken into account in future experimental investigations of these material systems.

(i) Pure carrier injection. Mixed injection by absorption of recombination radiation can be important, as shown in the results of chapter 7. For a wide band-gap absorption region, the recombination radiation is strongly absorbed in the quantum wells even at zero field. Narrow band-gap absorption regions should therefore be used, in which case the sub-bandgap radiation can still be absorbed in the quantum wells at high fields due to electroabsorption. It is therefore essential to minimise the recombination of minority carriers in the absorption regions. This can be achieved by making the absorption layers thinner than a few diffusion lengths (typically about a micron). The use of narrow band-gap absorption regions also means that pure injection can be achieved with back illumination without the need for the etching of wells in the substrate, which is transparent in the near infrared.

(ii) Field-independent internal efficiency. As shown in section 9.3, the quantum efficiency of MQW photodiodes can be strongly dependent on the field due to the suppression of the diffusion current by carrier
localisation in undepleted wells. The MQW's must therefore be entirely within the high field region. Doping of the MQW region by diffusion of dopants from the contact regions during growth or subsequent processing must be avoided by incorporating undoped spacer layers between the contact layers and the quantum wells.

(iii) Constant electric field profile. The low-doped P+-I-N+ diode allows the study of field-dependent effects such as geometrical resonances in the ionization rates, resonant tunnelling between wells and intersubband optical transitions, which are also indicators of high quality layers.

(iv) Low dark current. Large dark currents in heterojunction devices can lead to high concentrations of carriers trapped at the interfaces. These carriers can be impact ionized across the discontinuity, but this will obscure band-to-band impact ionization.

10.3. Impact ionization across the band-edge discontinuity.

Experimental evidence for this effect has been presented in Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As, GaSb/AlSb and InP/Ga$_{0.47}$In$_{0.53}$As MQW photodiodes. The single carrier-type nature of the multiplication was demonstrated in a structure with compositionally graded interfaces. The performance of these devices was limited by the large dark currents which provided the carrier reservoirs in the wells. If the technological problems of making selective contacts to the quantum wells can be overcome, then the three-terminal APD described in chapter 9 may offer low-noise performance, providing a true solid-state analogue of the photomultiplier tube. Unlike Staircase APD's, which must be fabricated in a material system where the conduction-band offset is greater than the ionization threshold energy, the three terminal APD can be fabricated in GaAs/AlGaAs or an InP-based material system. Due
to the low threshold energies compared to band-to-band ionization, the applied field can be smaller and tunnelling currents may be less of a problem. The performance at high frequencies may be limited by RC time constants.

10.4. Tunnelling in MQW's.

The large number of tunnelling transitions observed in the Al$_{0.48}$In$_{0.52}$As/Ga$_{0.47}$In$_{0.53}$As MQW's described in chapter 9 shows that this material system is very promising for tunnelling devices. Work is in progress to understand the origin of the hysteresis effects and the role of charge accumulation and barrier resonances, and to measure the tunnelling time associated with sequential resonant tunnelling.

10.5. Transport in undepleted quantum wells.

Studies of the characteristics of InP/Ga$_{0.47}$In$_{0.53}$As MQW's in chapter 9 showed the importance of carrier transport in the undepleted wells. In particular, it should be noted that constant capacitance-voltage curves measured at low temperatures and high frequencies do not imply that the MQW structure is fully depleted. Measurement of the activation energy of this transport has enabled an accurate determination of the band-offsets by admittance spectroscopy. This method could be extended to measure band-offsets in other material systems.
REFERENCES.


Electron Devices ED-30, p.381.


Phys. 52, p.1015.


Gossard, A.C., Wiegmann, W., Miller, R.C., Petroff, P.M., and Tsang,


Kane, E.O., 1966. In Semiconductors and Semimetals", edited by
Lang, D.V., Panish, M.B., Capasso, F., Allam, J., Hamm, R.A., Sergent,
6, p. 3625.
Electronics QE-17, p.275.
Levine, B.F., Choi, K.K., Bethea, C.G., Walker, J., and Malik, R.J.,
Magnea, N., Petroff, P.M., Capasso, F., Logan, R.A., and Foy, P.W.,
46, p.1161.
Masselink, W.T., Klein, M.V., Sun, Y.L., Chang, Y.C., Fisher, R.,
Mattera, V.D., Capasso, F., Allam, J., Hutchinson, A.L., Dick, J.,
McIntyre, R.J., 1986. Private communication.
Menendez, J., Pinczuk, A., Gossard, A.C., English, J.H., Werder, D.J.,
Rev.B 26, p.7101.


Nicholas, R.J., Portal, J.C., Houibert, C., Perrier, P., and Pearsall,
Pearsall, T.P., Capasso, F., Nahory, R.E., Pollack, M.A., and