José Maria Longras Figueiredo

Optoelectronic Properties of Resonant Tunnelling Diodes



Departamento de Física
Faculdade de Ciências da Universidade do Porto
Abril 2000

José Maria Longras Figueiredo

Optoelectronic Properties of Resonant Tunnelling Diodes



Tese submetida à Faculdade de Ciências da Universidade do Porto para a obtenção do grau de Doutor em Física

Departamento de Física
Faculdade de Ciências da Universidade do Porto
Abril / 2000

Optoelectronic Properties of Resonant Tunnelling Diodes

© José Maria Longras Figueiredo, 2000

ADF-UCEH, Universidade do Algarve, Campus de Gambelas, 8000-062 FARO, Portugal
URL: http://w3.ualg.pt/~jlongras/ E-mail: jlongras@ualg.pt

Esta tese é dedicada aos meus pais
António e Maria do Carmo
e à minha tia Maria

This thesis is dedicated to my parents
António and Maria do Carmo
and to my aunt Maria

The author would like to thank the Department of Electronics and Electrical Engineering of the University of Glasgow for making this work possible.

O autor agradece à *Fundação para a Ciência e Tecnologia* e ao *Programa PRAXIS XXI* o apoio financeiro através da Bolsa de Doutoramento PRAXIS XXI /BD/2871/94:

Data de início: 01-01-1995

Duração: 48 meses

País de Acolhimento: Reino Unido

Instituições de Acolhimento:

Faculdade de Ciências da Universidade do Porto

Department of Electronics and Electrical Engineering, University of Glasgow (de Junho de 1995 a Dezembro de 1998)

Instituição que Atribui o Grau: Universidade do Porto

Título do Plano de Trabalho: Propriedades Optoelectrónicas de Díodos Túnel Ressonantes

Área Científica: Ciências e Engenharia de Materiais

Acknowledgments

I am indebted to my supervisor at the University of Porto, Dr. António Pereira Leite, who took the necessary steps for this work to be possible, for his advice and encouragement, and for his help on the thesis preparation that made this dissertation a better document. I am very grateful for the opportunity he gave me to attend several international conferences.

The Centro de Física do Porto and the Departamento de Física da Universidade do Porto are to be thanked for providing computer facilities, and library. I am grateful to Mr. João Viana Lopes, Dr. José F. Mendes, and Dr. João Lopes dos Santos for introducing me to Lyx[1] and for their help.

I wish to express my gratitude to my supervisor at the Department of Electronics and Electrical Engineering of the University of Glasgow, Professor Charles N. Ironside, for proposing the Ph.D. project and accepting me as research student in his group, for his continuous guidance and support. I would also like to thank him for his continuous belief in this work, and for the freedom he allowed me during this research.

I would like to thank Professor Richard De La Rue for the support he gave me as research student in the Optoelectronics Research Group, for his encouragement and friendship during my stay in Glasgow.

I am extremely grateful to Professor Colin R. Stanley and Dr. Adam R. Boyd for useful discussions on many occasions as well as for providing high quality MBE material.

Special thanks are due to Dr. Karen McIlvanney and Dr. Ian Thayne for their help on many aspects of the experimental work, their encouragement and friendship. I am also grateful to Dr. Scott G. McMeekin of the University of Cardiff Wales for his help and advice.

I would like to express my gratitude to all my fellow research students and research assistants within the Optoelectronics Research Group for their kind and prompt help, demonstrated on many occasions.

I wish to acknowledge Professor John Marsh's Group, the Ultra-Small Structures Group, the Dry Etching Group, the Mechanical and the Electronic Workshops, and the Clean Rooms team at the University of Glasgow.

I am extremely grateful to the University of Glasgow. Special thanks are due to the technical staff, for their help and support that contributed to overcome the difficulties I encountered at Glasgow.

I would like to sincerely thank all those not mentioned and that by their hard work and kindness made the completion of the work reported possible.

Finally, and most of all, I am indebted to my family and Cristina for their support, encouragement and good moments throughout the period of study.

[1] Lyx is Copyright (©) 1995 by Matthias Ettrich, 1995-1998 Lyx Team (`http://www.lyx.org`). Lyx provides an "almost-WYSIWYG" interface to LaTeX.

... a useful solid-state device is one which can be used in electronic applications or can be used to study the fundamental physical parameters.

S. M. Sze, Physics of Semiconductor Devices,
Chapter 11, John Wiley, New York 1961

Sumário da Tese

Propriedades Optoelectrónicas de Díodos Túnel Ressonantes

José Maria Longras Figueiredo

Esta tese descreve o estudo das propriedades optoelectrónicas de guias de onda semicondutores unipolares incorporando um díodo túnel ressonante de dupla barreira de potencial (DBQW-RTD), implementado, sucessivamente, nos sistemas materiais semicondutores AlGaAs/GaAs e InGaAlAs/InP, e com comportamento análogo ao do DBQW-RTD (resistência diferencial negativa, NDC). Os sistemas materiais empregues permitem a operação nos comprimentos de onda 900 nm e 1300/1550 nm, respectivamente. Este díodo unipolar combina o confinamento óptico com o ganho eléctrico e as propriedades de banda - larga do DBQW-RTD. A investigação concentrou-se na demonstração do seu potencial como modulador óptico baseado no efeito de Franz-Keldysh: díodo túnel ressonante modulador por electro-absorção (RTD-EAM). Um método fiável de fabricação do RTD-EAM foi estabelecido para ambos os sistemas materiais.

O RTD-EAM baseado em AlGaAs/GaAs consiste num guia de onda unipolar AlGaAs-GaAs-AlGaAs contendo um DBQW-RTD em GaAs/AlAs. Este dispositivo apresenta NDC com uma razão corrente de pico - corrente de vale de 1.6, densidade de corrente de pico até 13.5 kcm^{-2} e tensão de pico no intervalo 1.5-3.2 V. Da característica corrente-tensão e do comportamento espectral do dispositivo, estimou-se uma profundidade de modulação óptica da ordem de 14 dB. A modulação por electro-absorção foi caracterizada usando uma "streak camera", medindo-se profundidades de modulação até 18 dB em RTD-EAMs de AlGaAs/GaAs com área activa de 4 μm $\times$ 200 μm. Foi também demonstrada modulação óptica associada à auto-oscilação do RTD-EAM, observando-se profundidades de modulação até 11 dB. A partir da análise da operação no modo de oscilação de relaxação, é de esperar uma largura de banda superior a 30 GHz. O estudo das propriedades dos dispositivos baseados no sistema AlGaAs/GaAs permitiu a elaboração de um modelo de operação simples, que constituiu uma ferramenta de grande utilidade no desenho das estruturas baseadas em InGaAlAs/InP apresentando NDC significativa.

O RTD-EAM implementado no sistema material InGaAlAs, com rede cristalina ajustada à do InP, opera no comprimento de onda 1560 nm. A configuração deste dispositivo corresponde a um guia de onda unipolar de InAlAs-InGaAlAs-InP incorporando, na região do núcleo de InGaAlAs, um DBQW-RTD de $In_{0.53}Ga_{0.47}As/AlAs$. O RTD-EAM de InGaAlAs apresenta superior NDC à dos dispositivos de GaAs (razão corrente de pico - corrente de vale até 7 e densidade de corrente de pico de 18 $kAcm^{-2}$). Um sinal de baixa frequência de 1 V de amplitude induz profundidades de modulação até 28 dB no comprimento de onda de 1565 nm. A resposta em electro-absorção apresenta uma alteração de 5 dB na absorção, induzida por variação de 1 mV na tensão de operação na região NDC. Foi observada uma alteração de 13 dB na absorção para variações de tensão entre os extremos da região de NDC.

Esta configuração do dispositivo mostrou o seu potencial para aplicações em sistemas optoelectrónicos de comunicação de alta frequência.

Abstract of the Thesis

Optoelectronic Properties of Resonant Tunnelling Diodes

José Maria Longras Figueiredo

This thesis reports an investigation of the optoelectronic properties of unipolar semiconductor optical waveguides incorporating a double barrier quantum well resonant tunnelling diode (DBQW-RTD), implemented successively in the AlGaAs/GaAs and the InGaAlAs/InP material systems, and showing typical DBQW-RTD behaviour (negative differential conductance, NDC). The material systems employed allow operation at wavelengths around 900 nm and 1300/1550 nm, respectively. This unipolar diode combines waveguide optical confinement with the electrical gain and potential wide bandwidth properties of the DBQW-RTD. Research concentrated on the demonstration of its potential as an optical modulator based on the Franz-Keldysh effect: the resonant tunnelling diode electro-absorption modulator (RTD-EAM). A reliable RTD-EAM fabrication process was established for both material systems.

The AlGaAs/GaAs RTD-EAM consists of an unipolar AlGaAs-GaAs-AlGaAs waveguide embedding a GaAs/AlAs DBQW-RTD. This device presented NDC with a peak-to-valley current ratio around 1.6, peak current densities up to 13.5 $kAcm^{-2}$ and peak voltages in the range $1.5 - 3.2\,V$. From the devices current-voltage characteristic and spectral behaviour, an optical modulation depth of around 14 dB was estimated. The electro-absorption modulation was characterized using a streak camera, and modulation depths up to 18 dB were measured in $4\,\mu m \times 200\,\mu m$ active area AlGaAs/GaAs RTD-EAMs. Optical modulation associated with RTD-EAM self-oscillation was also demonstrated, showing modulation depth up to 11 dB. From analysis of the operation in the relaxation oscillation mode, a bandwidth higher than 30 GHz is to be expected. Knowledge of the electrical properties of the devices based on the AlGaAs/GaAs system allowed a simple operation model to be established, which was employed to help the design of InGaAlAs/InP wafers presenting significant NDC.

The RTD-EAM implemented in the InGaAlAs material system, lattice matched to InP, operated at around 1560 nm. This device configuration consists of an unipolar InAlAs-InGaAlAs-InP waveguide incorporating, in the InGaAlAs core region, an $In_{0.53}Ga_{0.47}As$/AlAs DBQW-RTD. The InGaAlAs RTD-EAM showed larger NDC (peak-to-valley current ratio up to 7 and peak current density as high as 18 $kAcm^{-2}$) than the GaAs device. A low frequency electrical signal with an amplitude of 1 V induced an optical modulation depth as high as 28 dB at around 1565 nm. The electro-absorption response showed a change in absorption of 5 dB for a 1 mV variation in the bias voltage within the NDC region. A total change in absorption of 13 dB was observed as the RTD-EAM bias voltage was swept through the NDC region.

This device configuration proved its potential for applications in high frequency optoelectronic communication systems.

Sommaire de la Thèse

Propriétés Optoélectroniques de Diodes Tunnel Résonnantes

José Maria Longras Figueiredo

Cette thèse traite des propriétés optoélectroniques de guides d'ondes optiques à base de semiconducteur unipolaire incorporant une diode tunnel résonnante à puits quantiques double barrière (DBQW-RTD). Cette structure a été appliquée avec succès aux systèmes AlGaAs/GaAs et InGaAlAs/InP et a montré un comportement typique RTD (conductance différentielle négative, NDC). La longueur donde des deux systèmes semiconducteurs employés se situe autour de 900 nm et 1330/1550 nm respectivement. Cette diode unipolaire allie le confinement optique du guide d'onde avec le gain électrique et les propriétés éventuelles de large bande passante des DBQW-RTD. Cette étude a été concentrée sur la démonstration de ses possibilités en tant que modulateur optique basé sur l'effet Franz-Keldysh: le modulateur électro-absorbant à diode tunnel résonnante (RTD-EAM). Un procédé de fabrication fiable de RTD-EAM a été établi pour les deux systèmes étudiés.

Le modulateur RTD-EAM AlGaAs/GaAs est constitué d'un guide d'onde unipolaire AlGaAs-GaAs-AlGaAs enchâssant une structure DBQW-RTD AlGaAs/GaAs. Ce dispositif présente une conductance différentielle négative avec un rapport de courant maximum/minimun d'environ 1.6, des densités de courant maximum allant jusqu'à 13 $kAcm^{-2}$ et des tensions maximum dans la gamme 1.5 - 3.2 V. A partir des caractéristiques courant- tension et de la réponse spectrale des dispositifs, on a pu estimer la profondeur de modulation optique à environ 14 dB. La modulation électro-absorbante a été caractérisée en utilisant une "streak camera", et des profondeurs de modulation allant jusqu'à 18 dB ont été mesurées sur des modulateurs RTD-EAM AlGaAs/GaAs de zone active 4 μm $\times$ 200 μm. La modulation optique associée à l'auto-oscillation du modulateur a également été démontrée, montrant une profondeur de modulation allant jusqu'à 11 dB. A partir de l'analyse du fonctionnement en mode oscillation de relaxation, une bande passante plus large que 30 GHz est attendue. La connaissance des propriétés optiques des dispositifs basés sur le système AlGaAs/GaAs a permis d'établir un modèle de fonctionnement simple, qui a aidé à concevoir des wafers InGaAlAs/InP présentant une conductance différentielle négative significative.

Le modulateur RTD-EAM réalisé dans le système InGaAlAs, de même paramètre de maille que InP, opère autour de 1550 nm. La configuration de ce dispositif est constituée d'un guide d'onde unipolaire InAlAs-InGaAlAs-InP incorporant, dans le coeur InGaAlAs, une diode tunnel résonnante à puits quantiques double barrière $In_{0.53}Ga_{0.47}As$/AlAs. Le modulateur RTD-EAM InGaAlAs présente une conductance différentielle négative plus grande que le dispositif équivalent dans le système GaAs (rapport de courant maximum/minimun allant jusqu'à 7 et densités de courant maximum allant jusqu'à 18 $kAcm^{-2}$). Un signal électrique à basse fréquence et d'amplitude 1 V induit une profondeur de modulation optique jusqu'à 28 dB aux environs de 1565 nm. La réponse en électro-absorption a montré une modification en absorption de 5 dB pour une variation de tension de 1 mV dans la région de conductance différentielle négative. Une

modification maximale d'absorption de 13 dB a été observée quand la tension balaie la totalité de la zone de conductance différentielle négative.

Nous avons donc prouvé qu'une telle configuration pour ce dispositif est potentiellement intéressante pour des applications dans le domaine des systèmes de communication optoélectronique à haute fréquence.

Publications and Presentations

Journal Publications

1. **J. M. L. Figueiredo**, C. R. Stanley, A. R. Boyd, C. N. Ironside, S. G. McMeekin, and A.M. P. Leite, "Optical modulation in a resonant tunnelling relaxation oscillator", Applied Physics Letters, Vol. **74**, No. 9, 1 March 1999, 1191-1199.

2. **J. M. L. Figueiredo**, A. R. Boyd, C. R. Stanley, C. N. Ironside, S. G. McMeekin, and A.M. P. Leite, "Optical modulation at around 1550 nm in a inGaAlAs optical waveguide containing a InGaAs/AlAs resonant tunnelling diode," Applied Physics Letters, Vol. **75**, No. 22, 29 November 1999, 3443-3445.

3. **J. M. L. Figueiredo**, A. R. Boyd, C. R. Stanley, C. N. Ironside, and A. M. P. Leite, "Optelectronic switching in semiconductor waveguides incorporating a resonant tunneling diode," Photonics in Switching 1999, TOPS Vol. **32**, P. Prucnal and D. Blumenthal, Editors, Optical Society of America, Washington DC, 2000.

Conference Proceedings

1. **J. M. L. Figueiredo**, C. N. Ironside, A. M. P. Leite and C. R. Stanley, "Integration of a RTD with an Optical Waveguide to Form a High Speed Electroabsorption Modulator," Proceedings of IEE Colloquim on Optoelectronic Integration and Switching, IEE Reference No. 1997/372, paper 6/1-5.

2. **J. M. L. Figueiredo**, C. N. Ironside, A. M. P. Leite and C. R. Stanley, "High Speed Electroabsorption Modulator by Integration of a RTD with an Optical Waveguide," Proceedings of 5th IEEE International Workshop on High Performance Electron Devices for Microwave and Optoelectronic Applications (EDMO), 1997 IEEE Catalog No. 97TH8305, pp. 352-357. **Student poster paper prize.**

3. **J. M. L. Figueiredo**, A. M. P. Leite, C. R. Stanley and C. N. Ironside, "Integration of a resonant tunneling diode and an optical waveguide to form an electroabsorption modulator with a large bandwidth to drive voltage ratio," Proceedings of Conference on Lasers and Electro-Optics (CLEO), Vol. 6, 1998 OSA Technical Digest Series, pp. 263-264.

4. **J. M. L. Figueiredo**, C. N. Ironside, C. R. Stanley and A. M. P. Leite, "High speed optical waveguide modulator based on resonant tunnelling effect," Proceedings of Conference on Lasers and Electro-Optics - Europe (CLEO/Europe), 1998 IEEE Catalog No. 98TH 8326, pp. 20.

5. **J. M. L. Figueiredo**, C. N. Ironside, A.R. Boyd, C. R. Stanley, and A. M. P. Leite, "High frequency optical modulation in a resonant tunneling transmission line oscillator," Proceedings of 6th IEEE International Workshop on High Performance Electron Devices for

Microwave and Optoelectronic Applications (EDMO), 1998 IEEE Catalog No. 98TH8345, pp. 240-243.

6. **J. M. L. Figueiredo**, C. N. Ironside, A. R. Boyd, C. R. Stanley, A. M. P. Leite and S. G. McMeekin, "Self-oscillation in a resonant tunnelling diode electroabsorption modulator," Proceedings of Conference on Lasers and Electro-Optics (CLEO), 1999 OSA Technical Digest Series (Optical Society of America, Washington DC, 1999), pp. 125.

7. **J. M. L. Figueiredo**, A. R. Boyd, C. R. Stanley, C. N. Ironside, and A. M. P. Leite, "Optelectronic switching in semiconductor waveguides incorporating a resonant tunneling diode," Proceedings of The Integrated Photonics Research Conference (IPR99), Optical Society of America, 18-23 July 1999, Santa Barbara, USA, paper JWB3.

8. **J. M. L. Figueiredo**, A. M. P. Leite, C. N. Ironside, A. R. Boyd, C. R. Stanley, and S. G. McMeekin, "Optical modulation at around 1550 nm in a InGaAlAs optical waveguide containing a InGaAs/AlAs resonant tunnelling diode," Proceedings of the Quantum Electronics & Photonics Confence (QE-14), 6-9 Sept. 1999 Manchester, UK, paper 16.2.

9. **J. M. L. Figueiredo**, A.M. P. Leite, C. N. Ironside, C. R. Stanley, S. G. McMeekin, K. Bouris, D. G. Moodie, "Resonant tunnelling diode electro-absorption waveguide modulator operating at around 1550 nm," Proceedings of Conference on Lasers and Electro-Optics (CLEO) 2000, Optical Society of America, 7-12 May 2000, San Francisco, USA.

Papers in Preparation

1. V. E. Lyubchenko, N. V. Alkeev, C. N. Ironside, **J. M. L. Figueiredo**, C. R. Stanley, "Microwave properties of InGaAlAs resonant tunnelling optical modulators,"

2. S. G. McMeekin, K. Bouris, P. J. Tasker, **J. M. L. Figueiredo**, C. N. Ironside, D. Moodie, "Optoelectronic Response of a Resonant Tunneeling Diode Electro-Absorption Modulator at 1550 nm,"

3. J. F. Martins-Filho, R. E. de Araújo, A. S. L. Gomes, **J. M. L. Figueiredo**, C. R. Stanley, and C. N. Ironside, "40 dB Photodetection Gain in a Resonant Tunneling Diode Optical Waveguide,"

Submited Patent

Patente GB 9912178.2: *RTD-EAM Optical Modulator*, **J. M. L. Figueiredo**, C. N. Ironside and C. R. Stanley, University of Glasgow, UK, 1999.

Contents

1 Introduction **21**

1.1 Optoelectronic Semiconductor Technology 22

1.2 Optical Communication Systems 24

1.3 Resonant Tunnelling Devices 26

1.4 RTD Electro-Absorption Modulator 27

1.5 Thesis Outline 29

2 Physics and Properties of Semiconductors **31**

2.1 Charge Transport in Semiconductors 31

2.1.1 Energy Bands 32

2.1.2 Carrier Concentrations and Doping 35

2.1.3 Carrier Transport Mechanisms 38

2.1.4 The *pn* and Metal-Semiconductor Junctions 46

2.2 Heterostructures 52

2.2.1 Molecular Beam Epitaxy 53

2.2.2 Heterojunctions and Band Engineering 55

2.2.3 Layered Structures: Quantum Wells and Barriers 57

2.2.4 Optical Field Confinement 61

2.3 Optical Properties of Semiconductors 65

2.3.1 Optical Interactions in Semiconductors 66

2.3.2 Optical Absorption and Radiative Recombination 69

2.3.3 Electro-Optic and Electro-Absorption Effects 74

2.3.4 The Franz-Keldysh Effect 77

2.4 Optical Modulation 82

2.4.1 Modulation Schemes and Modulation Implementations 83

2.4.2 Electro-Refraction Modulators 85

2.4.3 Electro-Absorption Modulators 86

2.4.4 AlGaAs and InGaAlAs Material Systems for Optical Modulation 91

2.5 Final Comments 94

3 RTD Electro-Absorption Modulator **95**
3.1 Double Barrier Quantum Well RTD . . . 95
3.1.1 Resonant Tunnelling . . . 96
3.1.2 Effect of an Applied Bias . . . 100
3.1.3 Negative Differential Conductance . . . 102
3.1.4 Current-Voltage Characteristic . . . 103
3.2 Physical Mechanisms in DBQW-RTDs . . . 104
3.2.1 Transport Processes . . . 105
3.2.2 Models of DBQW-RTDs . . . 106
3.2.3 Material Systems . . . 111
3.2.4 Structural Parameters . . . 114
3.3 Electronic Applications of RTDs . . . 116
3.3.1 Tunnel Diode . . . 117
3.3.2 RTD Small-Signal Equivalent Circuit . . . 120
3.3.3 RTD Stability Analysis . . . 123
3.3.4 High Frequency and Functional Applications . . . 128
3.4 Optoelectronic Applications of RTD Structures . . . 135
3.4.1 Sources and Photo-Detectors . . . 135
3.4.2 Optical Bistability . . . 137
3.4.3 Intersubband Infrared Detectors . . . 138
3.4.4 Intersubband Infrared Lasers . . . 139
3.5 RTD Electro-Absorption Modulator . . . 142
3.5.1 RTD-Optical Waveguide Integration Concept . . . 142
3.5.2 Electro-Absorption Effect in RTD-OWs . . . 143
3.5.3 RTD-EAM Small-Signal Equivalent Circuit . . . 147
3.5.4 RTD-EAM Expected Performance and Applications . . . 152
3.6 Final Comments . . . 155

4 RTD-EAM: Design, Fabrication and Testing **157**
4.1 RTD-EAM Design . . . 157
4.1.1 Waveguide Structure . . . 158
4.1.2 RTD-EAM Structure . . . 160
4.1.3 High Frequency Bonding Pads . . . 161
4.1.4 Design and Production of Masks . . . 164
4.2 RTD-EAM Fabrication . . . 166
4.2.1 Photolithography . . . 166
4.2.2 Metallization . . . 169
4.2.3 Etching . . . 173
4.2.4 SiO_2 Passivation and Second Level Metallization . . . 178
4.3 Post Processing and Packaging . . . 181

4.3.1 Back-Sample Lapping . . . 181
4.3.2 Device Separation . . . 181
4.3.3 Package and Device Connection . . . 182
4.4 Electrical Characterization . . . 184
4.4.1 dc and Low Frequency Testing . . . 184
4.4.2 High Frequency Testing . . . 185
4.5 Optical Characterization . . . 186
4.5.1 Waveguiding Characterization . . . 186
4.5.2 Loss Measurement . . . 187
4.5.3 Waveguide Transmission Spectrum Change . . . 188
4.5.4 High Frequency Characterization . . . 189
4.6 Final Comments . . . 191

5 AlGaAs Based RTD-EAM **193**
5.1 AlGaAs RTD-EAM Design and Fabrication Procedures . . . 193
5.1.1 Ridge Waveguide Design . . . 193
5.1.2 Wafer Specification . . . 194
5.1.3 Fabrication and Testing Procedures . . . 197
5.2 AlGaAs RTD-EAM Electrical Characterization . . . 200
5.2.1 dc I-V characteristics . . . 200
5.2.2 Relaxation Oscillation . . . 203
5.2.3 High Frequency Response . . . 206
5.3 AlGaAs RTD-EAM Optical Characterization . . . 206
5.3.1 Optical Modes . . . 206
5.3.2 Propagation Loss . . . 207
5.3.3 RTD-EAM Transmission Spectrum . . . 208
5.4 AlGaAs RTD-EAM Modulation Characterization . . . 211
5.4.1 Relaxation Oscillation Optical Modulation . . . 211
5.4.2 Direct Modulation . . . 212
5.4.3 Modulation Due to Injection Locking . . . 213
5.5 Final Comments . . . 216

6 InGaAlAs Based RTD-EAM **219**
6.1 InGaAlAs RTD-EAM Design and Fabrication Procedures . . . 219
6.1.1 $In_{1-x-y}Ga_xAl_yAs$ System . . . 219
6.1.2 Ridge Waveguide . . . 220
6.1.3 Wafer Specification . . . 220
6.1.4 Fabrication and Testing Procedures . . . 222
6.2 InGaAlAs RTD-EAM Electrical Characterization . . . 225
6.2.1 dc Characterization . . . 225

6.2.2 High Frequency Characterization 226
6.3 InGaAlAs RTD-EAM Optical Characterization 228
6.3.1 InGaAlAs Waveguide Bandgap and Guided Modes 228
6.3.2 Propagation Loss 230
6.3.3 Waveguide Transmission Spectrum 231
6.3.4 Electro-Absorption Modulation 233
6.4 Final Comments 234

7 Conclusions and Suggestions for Future Work **237**
7.1 Conclusions 237
7.2 Suggestions for Future Work 239

I References **241**

Chapter 1

Introduction

Semiconductor devices are key components in advanced communication systems that process digital data at high rates, or handle analogue signals at high frequencies. To achieve high speed/high frequency capability a semiconductor device must have its operation based on charge carriers that respond strongly and rapidly to changes in the applied electric field and/or to the incident optical field. In Optoelectronics, a high speed device is either one that can generate, modulate or detect light at microwave frequencies or one that can handle very short pulses (that is, pulses with appreciable Fourier components in the microwave or millimetre-wave region, or above).[1]

Double barrier quantum well resonant tunnelling diodes (DBQW-RTDs) [1][2] have been designed to have cut-off frequencies well into the millimetre-wave (30 - 300 GHz) and tera-hertz (above 300 GHz) ranges [3]; high speed applications, including high frequency signal generation [4][5][6], high frequency signal detection [7] and high speed switching [8][9], have been demonstrated.

This thesis reports a novel optoelectronic approach for light modulation and detection at microwave and millimetre-wave frequencies based on a unipolar diode heterostructure implemented by incorporating a DBQW-RTD within a moderately doped optical waveguide core: the resonant tunnelling diode electro-absorption modulator (RTD-EAM).

This chapter contains a short overview of the state of the art of the Optoelectronics Semiconductor Technology and of Optical Communication Systems, introduces Resonant Tunnelling Devices and presents the RTD Electro-Absorption Modulator.

[1] In this context, high speed in Electronics and Optoelectronics respects to either high capacity for information transmission or ultrafast transient behaviour, implying typical operation above 1.0 GHz.

1.1 Optoelectronic Semiconductor Technology

Optoelectronics corresponds to the interrelation between electrical and optical phenomena, and results from the synergy of various fields of Science and Engineering: Optics, Electronics, Materials and Microtechnologies [10]-[21]. It covers diverse topics, such as light generation, modulation, transmission, routing, and detection, which, aside from their scientific relevance, are the basis for advanced applications in Communications and Signal Processing. Implementation of these functions relies on the optical, electro-optical and acousto-optic properties of materials such as semiconductors, silica glass, dielectric and ferroelectric single crystals, and polymers [16]. In many cases, functions like light generation, modulation, detection and routing are mainly implemented with recourse to semiconductor materials, where electrons and photons can interact efficiently. Optoelectronic semiconductor devices, such as light emitters, waveguides, intensity or phase modulators and/or detectors, have found widespread applications ranging from communications and radar to imaging systems, displays, compact disc systems, information storage, barcode reader systems, sensors, laser printers, chemistry and materials processing [15]. Optical fibre communication systems, optical interconnects and high power infrared laser diodes represent the major driving forces behind the development of optoelectronic semiconductor devices.

Semiconductor Optoelectronics work started in the early 1950s with silicon-based solar cells. The late 1950s saw the discovery of a new class of semiconductor materials: the three-five (III-V) family [17].[2] Almost at the same time, the idea of making solid solutions of these binary compounds (ternary compounds; the quaternary compounds appeared in the 70s) emerged. The first heterojunction lasers were obtained with gallium arsenide (GaAs) based materials (AlGaAs) in 1969, and the first indium phosphide (InP) based laser was demonstrated in 1975. At the present time, device properties are no longer strictly bound by those of the bulk, naturally occurring materials; instead, layered material structures can now be produced and tailored to specific device applications [22][23].

Optoelectronic components rely simultaneously on the optical and electronic properties of the materials they are fabricated from. In their most versatile form, optoelectronic systems are built up from components that are very different in functionality - light emitters, waveguides, intensity or phase modulators, and/or detectors, among others. Each one of these specific components requires a different material composition, often through heteroepitaxial structures with accurately controlled parameters, to achieve optimum performance [13][15][19].

Until the 80s, Liquid Phase Epitaxy (LPE) and Vapour Phase Epitaxy (VPE) were the major industrial techniques for producing LEDs, lasers and solar cells. The need for complex, well controlled multilayered structures resulted in the development of two new techniques: Metal Organic Vapour Phase Epitaxy (MOVPE) and Molecular Beam Epitaxy (MBE).[3] The general feeling is that, to date, MOVPE is more advanced for meeting these requirements when pro-

[2] Semiconductor families are defined by those materials that share the same lattice constant or by materials that may be combined, as alloys, to lattice match a third material [11].

[3] The traditional LPE and VPE techniques have almost disappeared from the laboratories, but are still quantitatively dominant in industry [17].

duction of sophisticated optoelectronic devices is involved, while MBE is often chosen for new demonstrations in the laboratory.

Progress in crystal growth, specially through molecular beam epitaxy (MBE), and advances in micro-fabrication technologies, allow the fabrication of artificial semiconductor structures constituted by layers, a few nanometers thick, of different materials. These include silicon (Si) based materials (e.g. GeSi/Si), and GaAs based materials (e.g. AlGaAs/GaAs). The most important semiconductors for high speed applications are Si, and the III-V family, which includes GaAs, InP and their related compounds and solid solutions (e.g., ternary compounds as AlGaAs and InGaAs, and quaternary compounds as InGaAsP and InGaAlAs). For many optoelectronic applications the systems of interest are direct bandgap semiconductor materials, such as the AlGaAs/GaAs, InGaAlAs/InP and InGaAsP/InP families [17]. High quality semiconductor substrates are only available for Si, GaAs and InP [15] (quartz substrates are also available and used for some applications [16]). Apart from silicon based solar cells and specific long wavelength applications based on II-VI compounds, most of the high speed optoelectronic devices are built on GaAs, GaP or InP substrates [15][17].

Strained heteroepitaxy (lattice mismatched semiconductor growth) results in the incorporation of built-in strain, leading to important effects on the electronic and optoelectronic properties of the material that can be exploited for production of high performance devices [22][24]. In principle, it can solve the substrate availability problem (allowing the possibility of growing GaAs on Si, for example) [17]. As a result of the strain, the cubic symmetry is no longer present in face centred cubic structure (fcc) based semiconductor crystals. An important consequence is a decrease of the density of states mass in the valence band (a reduction by up to a factor of 3) [22]. In quantum wells, strain can restore the degeneracy lost due quantum confinement, producing a very high density of hole states at the bandedge, which can be exploited in optical modulation [22][24]. The motivations for using strain in quantum well lasers are bandgap tuneability (up to 150 meV), reduced threshold current (the reduced hole mass allows the inversion condition to be reached at a lower injection density), laser emission with tailored polarization (TE, TM, or unpolarized light) [24], reduction in Auger rates (as a consequence of the reduced hole masses), and improved laser reliability (by suppressing defect migration into the active region) [23][24].

The appearance of optical fibres with very low attenuation, Fig. 1.1, led to the development of studies on ternary, then quaternary, compounds of the III-V family (specially InGaAlAs and InGaAsP on InP), aimed at obtaining optoelectronic device characteristics matched to the properties of these silica fibres [17]. Minimum loss below 0.2 dB/km at 1550 nm and zero dispersion at 1300 nm of standard single-mode optical fibres have caused the spectral windows at 1300 nm and 1550 nm to become dominant in today's optical telecommunication systems [14]. It is for this reason that the compound semiconductors InGaAsP and InGaAlAs, lattice-matched to InP and covering this wavelength range, have gained enormous importance over the past 15-20 years in providing the material basis for the optoelectronic components used in those systems [25]. The GaAs-based mixed crystal system Ga(In)AlAs, matching the 800-1100 nm wavelength

range, is predominantly utilized in optical interconnection applications [15], because of its high degree of technological maturity, specific superior material properties and the short transmission distances involved.[4]

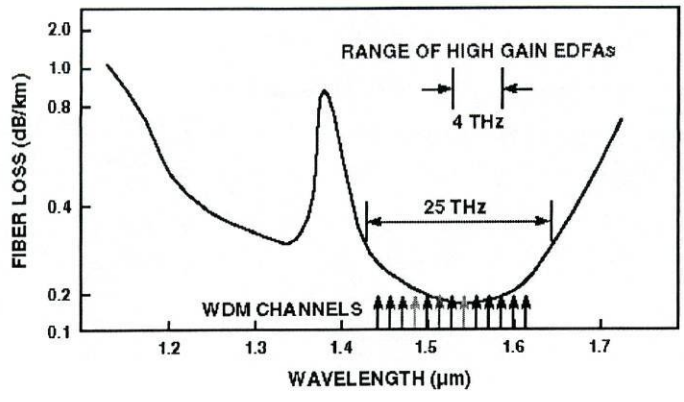


Figure 1.1: Attenuation in optical fibres as a function of the wavelength. (EDFAs: erbium doped fibre amplifiers; WDM: wavelength division multiplexing.)

Optoelectronic devices based on silicon and III-V semiconductor materials are processed or fabricated using many of the technologies that form the basis of silicon integrated circuit production [20][21]. A particularly ambitious goal is the production of optoelectronic integrated circuits (OEIC) [15], monolithically combining different active and passive devices in the same chip.

1.2 Optical Communication Systems

Basic communication systems perform the task of information transfer between two points using modulation of a high frequency carrier wave [10][14]. The information is normally coded in the form of an electronic signal exhibiting a one-to-one correspondence to the modulated parameter of the carrier wave. The modulation rate is, in general, much lower than the carrier frequency. Limitations to the performance of communication systems arise mainly from insufficient bandwidth or excessive noise level [14]. As lower frequency bands become more crowded, there is pressure to utilize higher carrier frequencies for emerging applications and systems, which require higher bandwidth channel capacity, higher frequency signal sources, modulators and receivers.

As the bit rate increases, the biggest problems in electronic communication systems are chip-to-chip connection, the limited bandwidth of components and crosstalk effects [10][14]. When the switching speed rises, so does the size of the output buffers necessary to provide fast voltage changes, and electrical power dissipation increases correspondingly [19]. Furthermore, problems with the maintenance of good signal integrity along the interconnecting lines become more acute at high switching speeds, as any mismatch between the line impedance and the line load can

[4]This is also true for high-power laser diodes, being extensively developed for use as pump sources for optical fibre amplifiers and solid-state lasers, for materials processing, and other applications [15].

cause severe "ringing" at the signal edges, making electrical signal transmission a bottle-neck in electronic systems [10][19]. All signal lines have to be treated as transmission lines at high frequencies and this complicates the process of printed circuit board design and production, particularly where multiple chips are being driven from one signal line [10][26]. In the future, communication between and within electrical boards will be made optically, using fibres, planar waveguide structures and free-space transmission [15][19]. Optoelectronic fibre links between computers have already been put into practical use, for example.

Optoelectronic interconnection and communication technologies are being widely studied due to their inherent parallelism (simultaneous action on many beams or wavelengths), and low crosstalk effects [14]. Moreover, all-optical transparent networks, which include all optical switching, are extensively pursued at the present time [19]. Advanced optical communication systems/networks will take advantage of the fibre-optic wide-bandwidth and low loss, Fig. 1.1, together with the interference free signal distribution supported by lightwave carriers, and will be mainly based on intensity modulation and direct detection [15][19]. To become practical, these systems require low cost high bandwidth optoelectronic devices, in order to create sufficient performance advantages over other types of communication systems, such as satellite links where high frequencies can also be employed. Direct broadcast satellite and other direct satellite-to-customer links are currently being used.

Present trends are to increase the transmission capacity by rising the bit rate (>10 Gbit/s) and by exploiting *optical frequency division*, *time division*, and *wavelength division* multiplexing techniques [14][19], which utilize different technologies to achieve this goal. Multiplexing techniques allow multiple communication channels to share the same link, increasing its capacity [14][15]; all three multiplexing technologies can be combined for maximum information transfer by fibre optics. However, to make fibre optic communication systems widely used, low cost optoelectronic devices requiring low drive voltage and consuming a small switching power are needed. Devices for such systems are, among others, stable distributed Bragg reflector lasers, wavelength-tuneable lasers and laser arrays, high speed lasers, fast and stable soliton sources, semiconductor optical amplifiers, optical fibre amplifiers (erbium doped fibre amplifiers, EDFAs), broadband modulators, low loss optical waveguides, wavelength-(de)multiplexers, wavelength-converters, space switches, and high efficiency balanced photo-receivers [14][15].

In addition, linear modulation in analogue systems has been actively considered, such as in optical community antenna television systems and radio-on-fibre (fibre-optic distribution of microwave and millimetre-wave signals) subscriber systems using subcarrier multiplexing in the access network [14].[5] Currently, the bandwidth of the total system is mainly limited by the performance of its microwave-optical and optical-microwave converters [14][15]. These systems demand high speed fibre-optic links, which require optoelectronic devices operating in the 40-100 GHz range, including switches, modulators and full optical logic ports [10][14].

Presently, great emphasis is placed on system development and definition of standards that

[5]The process of combining multiple microwave subcarriers into a modulated optical carrier is known as subcarrier multiplexing [14].

will form the foundation of a new information industry with the emphasis on integrated services [15]. Electro-absorption modulators (EAMs) are most promising as external optical modulators or optical switches for future multimedia large capacity optical communication systems [13][19], because of their small size and low power consumption.

1.3 Resonant Tunnelling Devices

The rapid progress in crystal growth and microfabrication technologies over the past two decades has led to the development of nanometre-scale semiconductor electronic/optoelectronic structures [22][23], where the wave-nature of electrons becomes relevant. As the size of a device scales down to that of an electron wavelength, quantum effects take over and new device concepts are needed beyond those used classically [22][30]. Nano-electronics research goes back to the work on resonant tunnelling [1], quantum wells and superlattices [22]. Significant examples of such devices are multiquantum well (MQW) lasers [24], MQW modulators and MQW photo-detectors [13][19], as well as electronic components based on quantum effects such as resonant tunnelling diodes [11].

Resonant tunnelling based devices, which utilize electron-wave resonance in potential barriers, have emerged as one of the most important testing grounds for modern theories of transport physics, and are central to the development of new types of semiconductor nanostructures [26][31]. Resonant tunnelling in semiconductor double/dual potential barriers was first demonstrated by Chang, Esaki and Tsu in 1974 [1][27], and since then has become a topic of great interest, investigated both from the standpoints of quantum physics and of its application in functional quantum devices [2][26][30]. Over the past two decades, RTDs have received a great deal of attention, following the pioneering work by Tsu, Esaki and Chang: well over a thousand research papers have been written on various aspects of this seemingly simple device.

Despite its simple structure, the resonant tunnelling diode (RTD) is indeed a good laboratory for electron-wave experiments, which can reveal various manifestations of quantum transport in semiconductor nanostructures, such as single electron tunnelling [2], and enable the study of more complex and advanced quantum mechanical systems [11][30]. Significant accomplishments have been achieved in terms of RTD device physics, modeling, fabrication technology, and circuit design and applications [26][31][32]. From the practical application point of view, the challenge is to reduce variations in peak current and voltage of RTDs across the wafer and from wafer to wafer [33], as well as fabrication dependent parasitic impedance of devices [32].

RTDs have two distinct features when compared with other semiconductor devices, from an applications point of view: their potential for very high speed operation and their negative differential conductance (NDC) [2]. The former feature arises from the very small size of the resonant tunnelling structure along the direction of carrier transport. The NDC corresponds to electric gain in active devices which can be applied in signal generation, detection and mixing, in multivalued logic switches at extremely high frequency, as well as in low-power amplifiers, local

oscillators and frequency locking circuits [8][9][26]. The RTD can also generate multiple high frequency harmonics, extending well into the submillimetre-wave band [2][11].

Besides this high speed potential, the negative differential conductance (NDC) makes it possible to operate RTDs as so-called functional devices, enabling circuits to be designed on different principles from those applied to conventional devices. Functional applications, such as a one-transistor static random acess memory, signal processing circuits with significantly reduced number of devices and multiple-valued memory cells using RTDs have been proposed and demonstrated [26][32]. These functional applications are highly promising since RTDs, with their simple structure and small size, can be integrated with conventional devices, such as field effect and bipolar transistors [26].

Most of the work on resonant tunnelling devices has been entirely concentrated on electronic high speed applications [2][11]. However, use of the resonant tunnelling (RT) effect in the optical and infrared regions has been proposed by several groups [3][28][29], due to its functional characteristics, high speed response and low rf power consumption. The optical applications utilize interband transitions (bandgap transitions) [34]-[39], whereas the infrared region employ intersubband transitions (transitions within the same band) [28][29].

Interband absorption at photon energies close to but smaller than the bandgap can be induced in a depletion region adjacent to the DBQW [39]. However, the interaction volume between the RTD depletion region and the optical radiation field is, in general, very small, as the depletion layer width ($\leq$100 nm [4]) in a typical RTD is much smaller than the optical wavelength ($\sim$1000 nm). This limitation is overcome if the light propagates in a direction parallel to the tunnelling plane, therefore increasing the interaction volume [41]. The same applies to intersubband transitions; a radically different type of semiconductor laser, the quantum cascade (QC) laser based on the resonant tunnelling effect, has been recently developed [42].

1.4 RTD Electro-Absorption Modulator

Embedding a RTD within an optical waveguide core can enhance the interaction length between its active region and the guided light. S. G. McMeekin *et al* [41][43] demonstrated electro-absorption modulation (at around 900 nm) from a 900 MHz CW signal and a 2 ns pulse signal with an optical response time of 500 ps, in an AlGaAs/GaAs optical waveguide containing a GaAs/AlAs double barrier quantum well RTD, via the Franz-Keldysh effect.

This thesis reports further developments of this concept [44], and presents the design and demonstration of the modulator in the InGaAlAs material system, lattice matched to InP, operating at around 1550 nm: the resonant tunnelling diode electro-absorption modulator (RTD-EAM) [45]. The device consists of an unipolar InAlAs-InGaAlAs-InP waveguide incorporating an InGaAs/AlAs DBQW-RTD in the core region (InGaAlAs). This waveguide modulator configuration is shown schematically in Fig. 1.2, together with the refractive index n and the Γ-conduction band energy (E_c) profiles of the semiconductor waveguide.

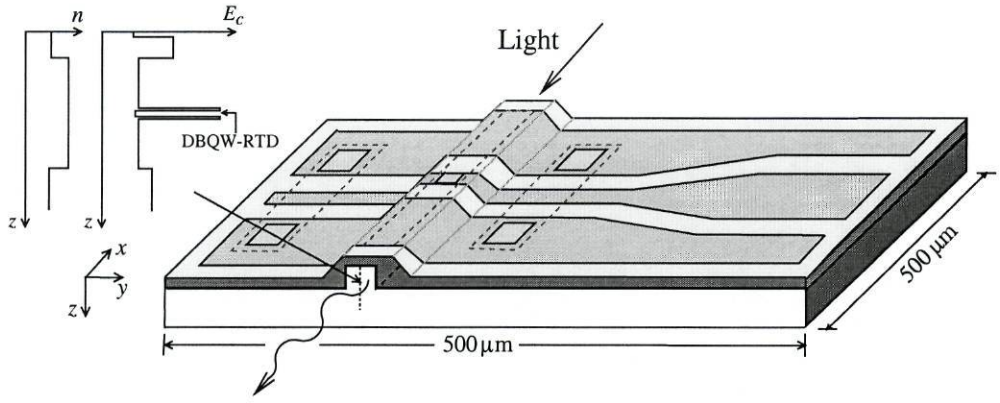


Figure 1.2: Scheme of a RTD-EAM.

The presence of a RTD within the optical waveguide introduces high non-linearities into the electric field distribution across the waveguide core, which becomes strongly dependent on the bias voltage due to accumulation and depletion of electrons on the emitter and collector sides of the RTD, respectively. Modulation relies on a depletion region that is formed in the waveguide core, on the collector side of the DB-RTD; depending on the bias condition, a substantial part of the diode terminal voltage may be dropped across this region, inducing a large electric field distribution in the depleted part of the waveguide core. This electric field shifts the waveguide transmission bandedge to longer wavelengths via the Franz-Keldysh effect [41], changing the transmission characteristics of the waveguide.

A small high frequency ac signal (<1 V) can induce high speed RTD switching. This produces substantial high speed modulation of the waveguide optical absorption coefficient and, therefore, light modulation at photon energies slightly lower than the core bandgap energy. The modulation depth can be considerable because, under certain conditions, the RTD operation point switches well into the two positive differential resistance portions of the current-voltage characteristic [46], with a substantial part of the terminal voltage dropped across the depleted region on the collector side [45]. Compared to conventional *pn* electro-absorption modulators,[6] the advantage of the RTD-EAM is that, when dc biased close to the RTD negative differential conductance (NDC) region, the device behaves as an optical waveguide electro-absorption modulator integrated with a wide bandwidth electrical amplifier [44].

This new modulator type, implemented in the InGaAs-InAlAs material system lattice matched to InP, appears to be a promising route towards high speed, low rf power consumption, opto-electronic (rf-optical and optical-rf) converters. The RTD-EAM could be an important element in these systems due to its electrical gain and potential wide bandwidth (up to 60 GHz). Other applications to be considered are optical injection locking of free-running microwave or millimetre-wave remote local oscillators based on RTDs, which could also be useful for clock recovery through generation of an electric signal locked to the optical signal bit rate. Its capabil-

[6]The electrical bandwidth of reverse-biased *p-i-n* junction electro-absorption modulators can be limited to ~20-30 GHz by the creation of carriers that must be swept out of the active region [13][15].

ity to operate in the self-oscillation mode can be used to implement self-oscillating modulators, which are of interest for WDM systems as the sidebands produced by the modulator can provide a comb of frequencies at the appropriate channel spacing.

1.5 Thesis Outline

The aim of the work reported in this thesis was to investigate the optoelectronic properties of optical waveguides containing a resonant tunnelling diode (RTD) for light modulation at microwave and millimetre-wave rates. The major application areas could be microwave and millimetre-wave fibre-radio systems, mobile broadband systems and broadband wireless access.

The contents of this thesis are arranged in seven chapters, which are outlined here.

Chapter 2 introduces a background of III-V semiconductor physics, including a description of carrier transport mechanisms and electronic behaviour in heterostructures. Relevant optical properties, such as refractive index and absorption, are addressed, together with optical confinement and optical modulation.

Chapter 3 presents the basics of the resonant tunnelling diode structure, discussing its high frequency performance. A novel electro-absorption modulator based on integration of a RTD with an optical waveguide is introduced.

Chapter 4 describes the general processes involved in fabrication and characterization of the RTD-EAM, as used in this work, from wafer and waveguide designs to photolithography, etching, metallization, passivation, packaging, and optoelectronic characterization.

Chapter 5 contains the description of high speed electro-absorption modulation of light at wavelengths around 900 nm with RTD-EAMs implemented in the GaAs/AlGaAs material system, by direct electrical excitation and due to self-oscillation.

Chapter 6 is devoted to electro-absorption modulation of light with wavelength around 1550 nm with RTD-EAMs based on the InGaAlAs/InP material system.

Finally, Chapter 7 summarizes the contents and the main results obtained in this research, and finalizes with a discussion of the conclusions drawn from the experiments, together with suggestions for future work in the field.

Chapter 2

Physics and Properties of Semiconductors

Semiconductor device operation typically involves some form of response of the materials physical properties to external perturbations, which are usually electric or electromagnetic fields, and the device performance depends upon how charged carriers respond to these external stimuli. In this chapter, the basic semiconductor concepts and notations, which will be used throughout the thesis, about crystal structure, energy bands, carrier concentrations, doping, and current transport mechanisms are introduced. Device building blocks, such as homojunctions, heterojunctions, potential barriers, quantum-wells, superlattices, and metal-semiconductor contacts will be summarily discussed. The physical phenomena upon which electronic and optoelectronic semiconductor devices are based, and the interaction of light with semiconductor structures and its applications, will also be introduced, with particular emphasis on the aspects relevant to high frequency operation of electro-absorption devices and the specific implementation of the RTD-EAM.

2.1 Charge Transport in Semiconductors

The electronic and optical properties of semiconductors are determined mainly by their crystalline structure [48]-[53]. Elemental semiconductors, such as Si, present diamond lattice crystalline structures, while compound semiconductors, such as GaAs, AlAs, InP, have Zinc Blend structure (face centred cubic structure, *fcc,* with two non-identical atoms per basis) [48]-[53]. For GaAs the arrangement of the two species is alternate: Ga occupies the original sites of the *fcc* lattice, while As occupies the tetrahedral sites, to give a *unit cell* that contains four atoms of each species.[1] A complete description of the semiconductor electrical and optical properties is rather involved [12][50]. Due to the cubic symmetry of the semiconductor crystals, useful information can be

[1]The length of each side is the lattice constant a ~0.565 nm in GaAs at 300 K (1 mono-layer, ML, corresponds approximately to 0.283 nm).

obtained, however, using approximations [50]; a standard simplification that allows treatment of the dynamics of carriers in a semiconductor as though they were free, rather than using the complicated band structure of the semiconductor, known as the effective mass approximation [22], will be used throughout this text.

2.1.1 Energy Bands

To understand the optoelectronic properties of semiconductors, such as absorption due to electronic transitions in the presence of an incident optical wave, it is necessary to know the electronic band structure [12]. The electron wavefunction in the semiconductor must satisfy the Schrödinger equation [12]:

$$\left[\frac{\left(i\hbar\vec{\nabla}+e\vec{\mathcal{A}}(\vec{r};t)\right)^2}{2m}+U(\vec{r};t)\right]\Psi(\vec{r};t)=i\hbar\frac{\partial}{\partial t}\Psi(\vec{r};t), \tag{2.1}$$

where $\Psi(\vec{r};t)$ is the particle wave function, $U(\vec{r})$ is the potential energy, $\vec{\mathcal{A}}(\vec{r};t)$ is the vector potential (representing magnetic fields), e is the elementary charge, and m is the carrier mass in the semiconductor. It is convenient to develop the discussions in terms of monoenergetic wave functions $\Psi(\vec{r};t)=\psi(\vec{r})\exp(-iEt/\hbar)$; for the moment, it will also be assumed that there are no magnetic fields, so that the vector potential $\vec{\mathcal{A}}(\vec{r};t)$ can be set to zero [12].

The time-independent form of the Schrödinger equation [54] is obtained,

$$\left[-\frac{\hbar^2}{2m}\nabla^2+U(\vec{r})\right]\psi(\vec{r})=E\,\psi(\vec{r}). \tag{2.2}$$

Because of the symmetry and periodicity of the semiconductor crystals, the potential function $U(\vec{r})$ is periodic, leading to a periodic electron density across the crystal, and consequently the solutions of Eq. 2.2 have a component with the same periodicity as the crystal [50]. The Bloch theorem states that the electron wavefunction in a crystal is the product of a function $u(\vec{r})$ with the same periodicity as the background potential with a plane wave $e^{i\vec{k}\cdot\vec{r}}$, which typically varies much more slowly:

$$\psi(\vec{r})=u(\vec{r})\exp\left(i\vec{k}\cdot\vec{r}\right), \tag{2.3}$$

with $E=\hbar^2\left|\vec{k}\right|^2/2m$, and $u(\vec{r})$ being known as the Bloch function [50].

The electronic and optical properties of semiconductors are determined by the atoms outermost electrons, which correspond to electrons staying in orbitals either of the s-type or of the p-type. Consequently, the Bloch functions in semiconductors are adequately described by s- and p-type like functions [53], or their combination.

The wavefunction solutions for the electronic states provide the electron energy as a function of the plane wave vector $\vec{k}$ and the unit cell periodic part of the wave function $u(\vec{r})$. The magnitude of the vector $\vec{k}$ is related to the wavelength of the electron wave function (de Broglie wavelength) by $\left|\vec{k}\right|=\frac{2\pi}{\lambda}$; the quantity $\vec{p}=\hbar\vec{k}$ is called the crystal momentum [50].

The energy solutions correspond to regions in the $\vec{k}$ space where there are allowed states (*allowed energy bands*) separated by regions where allowed states cannot exist (*forbidden energy bands*) [50]. The energy-momentum ($E - \vec{k}$) relation is called the *energy band-structure* of the semiconductor. Figure 2.1 represents the energy bands along different $\vec{k}$ directions of Si, GaAs and AlAs.

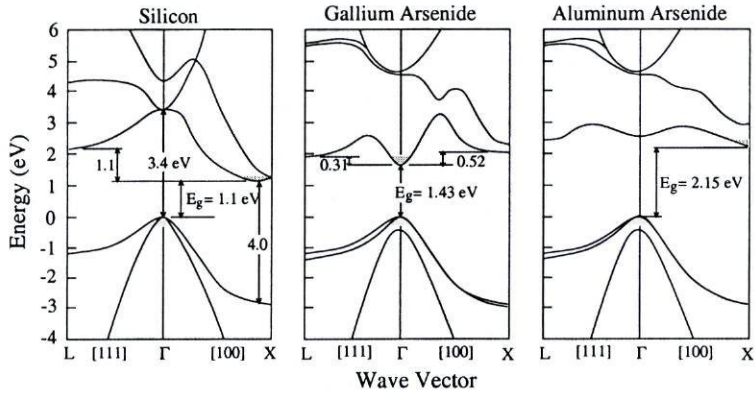


Figure 2.1: Energy-band structures of Si, GaAs and AlAs [11].

The energy bands that are normally filled with electrons at 0 K in semiconductors are called the *valence bands,* while the upper unfilled allowed bands are called the *conduction bands*. The separation between the energy of the lowest conduction band and that of the highest valence band is called the *bandgap energy,* E_g, which is one of the most important parameters in semiconductor physics [49]. The energy difference between the vacuum level and the highest occupied electronic state is the *work function,* and the energy difference between the vacuum level and the bottom of the lowest conduction band is the *electron affinity* [23].

At 0 K, the electrons are bound in their respective lattice points (forming the valence band); consequently, they are not available for conduction.[2] However, at finite temperature some of the electrons are thermally excited, breaking the bonds with the lattice atoms, which gives rise to conduction electrons that can participate in current flow. An electron deficiency is left in the covalent bond, which may be filled by one of the neighbouring electrons, resulting in a shift of the deficiency location. This valence band unoccupied state (electron deficiency) behaves as a particle similar to an electron but with symmetrical charge, and is called *hole* (h) [49]. If the electron momentum in the valence band is $\vec{k_e}$, the corresponding hole momentum is $\vec{k_h} = -\vec{k_e}$ [49]. Under an applied electric field, all conduction electrons move in the direction opposite to the electric field, resulting that the vacancy states move in the field direction. Thus, the hole responds to an electric field as if it had a positive charge.

For the face centred cubic structure, which is the crystaline structure for most semiconductors, the relevant $\vec{k}$-vector values are confined to the *first Brillouin zone* [50],[3] centred at

[2]When an allowed band is completely filled with electrons, the carriers in that band cannot contribute to any current unless they move to an empty band, because the electrons, being fermions, can only move into an empty state.

[3]The Brillouin zone corresponds to a polyhedron in $\vec{k}$-space.

$\vec{k} = (0,0,0)$ - the Γ-point; other important high symmetry points in the Brillouin zone are the X-point, corresponding to $\vec{k} = (\frac{\pi}{a}, 0, 0)$, and the L-point, corresponding to $\vec{k} = (\frac{\pi}{a}, \frac{\pi}{a}, \frac{\pi}{a})$, where a is the lattice constant. The *bandedges* (minimum or maximum) of the allowed bands of most semiconductors occur at $\vec{k}$-values corresponding to these points. The top of the valence band of most semiconductors occurs at the Γ-point, where the wavefunction character is p-type. The bottom of the conduction band in some semiconductors occurs also at the Γ-point, with a wavefunction purely s-type; such semiconductors (like GaAs, InP) are called *direct bandgap* semiconductors [50]. In other semiconductors, such as Si, Ge, AlAs, the bottom of the conduction band does not occur at the Γ-point, but at a certain other point $\vec{k} \neq (0,0,0)$, and they are called *indirect bandgap* semiconductors [50]; an electron transition from the valence band to the conduction band in these materials requires not only an energy variation ($> E_g$) but also some change in the crystal momentum $\vec{k}$ [50].

In general, the valence band structure close the bandedge is more complicated than the conduction band, Fig. 2.1, and the situation becomes worse if the cubic symmetry is broken [50]. In the zinc blend structures, it consists of four branches when spin is neglected (each branch is doubled when spin is taken into account). Three of the four branches are degenerate at the Γ-point, and form the upper edge of the valence band structure. Along a given direction, the top of the valence bands and the bottom of the conduction bands can be approximately fitted by parabolic bands with different curvatures $\partial^2 E/\partial k^2$ [23]. Typical semiconductor energy-momentum ($E - \vec{k}$) relations near the top of the highest valence bands and the bottom of the lowest conduction band are schematically shown in Fig. 2.2. In this figure the bottom of the lowest conduction band is designated by E_c, and the top of the highest valence band by E_v.

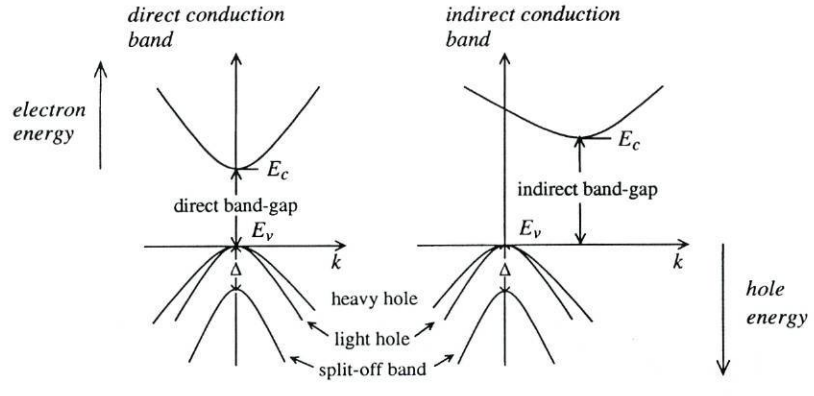


Figure 2.2: Schematic representation of the valence, direct- and indirect-gap conduction bands.

In most semiconductor devices the performance is controlled by the bandedge states at the bottom of the lowest conduction band and at the top of highest valence band. The top valence bands can be approximately fitted by two parabolic bands with different curvatures: the *heavy-hole* band (the wider band, with smaller $\partial^2 E/\partial k^2$) and the *light-hole* band (the narrower band, with larger $\partial^2 E/\partial k^2$) [50]. For most semiconductors, the energy of the *split-off band*, Δ (Fig. 2.2), is quite large, and therefore there is a negligible number of holes in this band. As a result,

when analysing semiconductor properties it is sufficient to consider only the heavy-hole and the light-hole valence bands [50]. The terms heavy-hole and light-hole should not always be associated with the effective mass of the heavy- and light-hole states, respectively. Under certain conditions (e.g., when the semiconductor is under strain) the valence band mass can be altered considerably. It is then possible for the light-hole mass to become actually heavier than the heavy hole mass in some directions [24].

Near the conduction bandedge (occurring at $\vec{k} = \vec{k_0}$) it is possible to represent the conduction band structure by

$$E(\vec{k}, \vec{k_0}) = E(\vec{k_0}) + \hbar^2 \sum_{i=x,y,z} \frac{(k_i - k_{0_i})^2}{2m_e}, \tag{2.4}$$

where the subscript i represents the x, y, z components of $\vec{k}$ and $\vec{k_0}$ [50]. For a direct bandgap material, $\vec{k_0} = (0,0,0)$, and if the band structure is isotropic, as with most direct bandgap semiconductors, the relation becomes

$$E(\vec{k}) = E_c + \frac{\hbar^2 k^2}{2m_e}. \tag{2.5}$$

E_c is the conduction bandedge energy. The zero of energy is conventionally taken at the top of the highest valence band. Energy in the conduction band is conventionally defined to be positive when measured upward, and the valence band energy is positive when measured downward, Fig. 2.2. The quantity m_e, the electron *effective mass* in the semiconductor, is defined by the relation

$$\frac{1}{m_e} = \frac{1}{\hbar^2} \frac{d^2E}{dk^2}, \tag{2.6}$$

which, in some semiconductors (such as Si), is dependent on crystal orientation [49].[4] A similar expression holds for the effective mass of the bandedge states near the top of the valence band.

The effective mass in the valence band is, in general, higher than the conduction band electron effective mass. In the case of GaAs, $m_e = 0.067m_0$ for electrons, for light-holes (lh) $m_{lh} = 0.082m_0$ and for heavy-holes (hh) $m_{hh} = 0.45m_0$ [12]; m_0 represents free electron mass.

2.1.2 Carrier Concentrations and Doping

In semiconductors the bonding force between atoms has, in general, a covalent nature.[5] At 0 K, the electrons are bound in their respective lattice points, but at finite temperature, thermal excitation may break the bonds, giving rise to electrons that can participate in current conduction. In fact, the conduction properties of semiconductors are sensitive to temperature, illumination, external fields, and minute amounts of impurity atoms.

---

[4] The effective mass is, in general, tensorial, with components m_{ij} defined as $\frac{1}{m_{ij}} = \frac{1}{\hbar^2} \frac{\partial^2 E(\vec{k})}{\partial k_i \partial k_j}$ [12].

[5] Covalent bonding can occur between atoms of the same element or between atoms of different elements that have similar outer-shell (valence) electron configurations.

Intrinsic Semiconductors [49]

In intrinsic (pure) semiconductors at a finite temperature, electrons are scattered from the valence band into the conduction band, and consequently there are electrons in the conduction band and holes in the valence band which can transport current.

The probability of an electronic state with energy E to be occupied by an electron is given by the Fermi-Dirac distribution function, f_e,

$$f_e(E) = \frac{1}{1 + \exp\left(\frac{E - E_{F_i}}{k_B T}\right)}, \tag{2.7}$$

and the probability of an electronic state with energy E to be empty, the hole distribution function, f_h, is given by

$$f_h(E) = 1 - f_e(E) = \frac{1}{1 + \exp\left(-\frac{E - E_{F_i}}{k_B T}\right)}, \tag{2.8}$$

where k_B is the Boltzmann constant, T is the absolute temperature, and E_{F_i} is the *Fermi Level* energy (the Fermi level energy is such that the probability of occupation is exactly 0.5), which can be determined from the *charge neutrality condition.* The Fermi level energy for an intrinsic semiconductor is given by

$$E_{F_i} = \frac{E_c + E_v}{2} + \frac{3k_B T}{4} \ln\left(\frac{m_h}{m_e}\right). \tag{2.9}$$

At room temperature, the second term is small and the Fermi level generally lies very close to the middle of the bandgap.

The number of allowed electronic states in a given band per unit volume per unit energy around an energy E is the band *density of states,* which depends on the dimensionality of the system and its respective *energy vs. momentum* relation, $E(\vec{k})$. For a bulk semiconductor and assuming that the energy momentum relation near the bandedges has the simple parabolic form $E(k) = \frac{\hbar^2 k^2}{2m}$, the density of states takes the form

$$\rho_{3D}(E) = \frac{\sqrt{2}\,(m_{dos})^{\frac{3}{2}}}{\pi^2 \hbar^3} \sqrt{E}, \tag{2.10}$$

where m_{dos} represents the *density of states* effective mass. For direct bandgap semiconductors, the conduction band density of states effective mass is identical to m_e. A similar expression holds for the valence band, but the heavy-hole and the light-hole masses have to be considered instead. The difference in effective mass for electrons and holes shows that the density of states in the valence band is almost 20 times greater than in the conduction band.

For an intrinsic semiconductor, the density of electrons in the conduction band, n, is equal to the density of holes in the valence band, p, $p = n \equiv n_i$, where n_i is the effective intrinsic carrier

density. The electron density in the conduction band is given by

$$n = \int_{E_c}^{\infty} \rho_{3D}(E) f_e(E) dE = N_c \exp\left(-\frac{E_c - E_{F_i}}{k_B T}\right), \tag{2.11}$$

where $N_c = 2\left[\frac{m_e k_B T}{2\pi\hbar^2}\right]^{3/2}$ represents the effective density of states in the conduction bandedge.

A similar expression holds for the hole density p in the valence band. For GaAs, at room temperature (300 K), N_c is 4.7×10^{17} cm^{-3}, and N_v is 7×10^{18} cm^{-3}, where $N_v = 2\left[\frac{m_h k_B T}{2\pi\hbar^2}\right]^{3/2}$. The relation $p \cdot n = n_i^2$ is called the *mass action law*, and is valid for any semiconductor under thermal equilibrium. The intrinsic carrier density is then

$$n_i = \sqrt{p \cdot n} = \sqrt{N_v \cdot N_c} \exp\left(-\frac{E_g}{2k_B T}\right), \tag{2.12}$$

where $E_g = E_c - E_v$ (the energy bandgap). For GaAs at 300 K, n_i =1.79 $\times$ 10^6 cm^{-3} (whereas in metals, carrier concentrations are $n_i \sim 10^{21}$ cm^{-3}). Intrinsic (pure) semiconductors have low concentration of carriers that can contribute to a current. Semiconductors become most useful when doping is used to alter, in a controllable manner, the density of charge carriers.

Extrinsic Semiconductors [49]

The conductivity of a semiconductor can be increased by replacing a fraction of an element constituting the semiconductor with another element, known as *impurity or dopant.* Dopants are elements that either possess an extra electron in the outer shell compared to the host semiconductor, or instead have one less electron, the dopant being called *donor* or *acceptor*, accordingly. For a *n-type* semiconductor, the material is doped with a donor impurity, and for a *p-type* the dopant is an acceptor.

When a semiconductor is intentionally doped with donor impurities, additional allowed energy states are created in the bandgap region near the conduction bandedge, the *donor impurities energy levels.* The same considerations apply for an acceptor, with the exception that *acceptor impurity energy levels* are created in the forbidden gap, near the valence bandedge. In heavily doped n-type materials, all valence-band states and some extent of conduction band states are filled, with the Fermi level energy being shifted upwards into the conduction band. On the other hand, a heavily doped p-type semiconductor exhibits a downward shift of the Fermi energy into the valence band, so that the entire conduction band and some states lying near the top of the valence band are empty.

At low temperatures, the donor electrons in n-type semiconductors are tied to the donor sites and this effect is known as *carrier freeze out.* At room temperature, the donor impurities are ionized, freeing electrons into the conduction band. Such electrons can carry current and modify the electronic properties of the semiconductor, leaving the donor atoms positively charged. The ionization energy for donors (measured from the conduction bandedge) in GaAs is around -0.007 eV. Under a complete ionization condition, the number of extra electrons in the conduction band

is roughly equal to the donor concentration, N_d^+, and the electron density in the semiconductor is $n = n_i + N_d^+$.

Doping causes the balance between electron and hole concentrations to be broken, that is, $n - p = \Delta n \neq 0$. However, the law of mass action still holds (the relation is not valid only at high doping levels [50]), with the electron and hole densities being

$$n = \frac{1}{2}\Delta n + \frac{1}{2}\sqrt{\Delta n^2 + 4n_i^2} \tag{2.13}$$

$$p = -\frac{1}{2}\Delta n + \frac{1}{2}\sqrt{\Delta n^2 + 4n_i^2}, \tag{2.14}$$

with

$$\frac{\Delta n}{n_i} = 2\sinh\left(\frac{E_{F_n} - E_{F_i}}{k_B T}\right), \tag{2.15}$$

where E_{F_n} is the Fermi energy level corresponding to the doped semiconductor (the subscript n refers to the n-type semiconductor).

The Joyce-Dixon approximation [53] relates the Fermi level energy with the carrier concentration:

$$E_{F_n} = E_c + k_B T\left[\ln\left(\frac{n}{N_c}\right) + \frac{1}{\sqrt{8}}\frac{n}{N_c}\right]. \tag{2.16}$$

A similar expression can be written for holes (p).

At room temperature, the donors and acceptors are essentially all ionized. Three regions can be identified: the freezeout regime discussed above; the saturation regime, which corresponds to complete ionization of the impurity, where the charge carrier density is essentially equal to the impurity density; the third regime occurs at high temperatures, where the intrinsic carrier density dominates.

The band-structure discussed so far applies to the case where all the electrons are in the valence band and the conduction band is empty. The introduction of extra electrons in the conduction band, or holes in the valence band, induces changes in the $E(\vec{k})$ relation. An electron (e) - hole (h) pair is coupled through the Coulombic interaction forming an *exciton* [12][19]. In bulk semiconductors, excitons are usually observable optically only in high purity samples, due to their rather small binding energies [19]. However, when quantum confined, the exciton binding energy is greatly increased; moreover, this binding energy can be controlled easily by simple electronics or optics, allowing the use of excitonic transitions for high speed modulation of optical signals [12], as well as for optoelectronic switches [19].

The presence of excess electrons in the conduction band and/or holes in the valence band shifts the bandgap energy, decreasing it slightly, and this phenomenon is often called *bandgap renormalization* [53], which is relevant with heavily doped materials.

2.1.3 Carrier Transport Mechanisms

Electrical conduction in semiconductors relies on several physical mechanisms, such as drift and diffusion of carriers, their generation/recombination and tunnelling effects. In the study

of semiconductor components such as diodes and transistors, the characteristics of the devices are described by the *voltage-current* relations [49]. The standard equations governing the carrier transport and the electrostatic potential distribution are the *drift-diffusion* and *Poisson equations*, respectively [49].

Drift Current [11]

At finite temperatures, electrons in the semiconductor are scattered as a result of collisions with lattice atoms, impurity atoms, and other scattering centres. The average distance and time between collisions are designated *mean free path* (l_{cp}) and *mean free time* (τ_{ct}), respectively. This random motion of electrons gives no net contribution to the current. However, when an electric field $\vec{\mathcal{E}}$ is applied to a semiconductor, a force $\vec{F} = -e\vec{\mathcal{E}}$ will accelerate the electrons in the direction opposite to the field. A steady state is established in which the electrons have an additional velocity component, superimposed upon their thermal motion, which is called the *drift velocity*, $\vec{v_d}$. The momentum $m_e\vec{v_d}$ gained by the electron is equal to $-e\vec{\mathcal{E}}\tau_{ct}$, giving

$$\vec{v_d} = -\left[\frac{e\tau_{ct}}{m_e}\right]\vec{\mathcal{E}} = -\mu_n\vec{\mathcal{E}}, \tag{2.17}$$

where $\mu_n = e\tau_{ct}/m_e$ represents the electron *mobility*. (A similar expression can be written for holes in the valence band.) In the quantum mechanical picture, the mobility is given by

$$\mu_{k_i} = -\left[\frac{e\tau_{ct}}{\hbar^2}\right]\frac{d^2E}{dk_i^2}, \tag{2.18}$$

where the second derivative of the energy is evaluated in the $\vec{k_i}$-space direction of carrier motion in the bandedge.[6]

In the presence of an electric field there is a flow of charge across the semiconductor due to electron and hole densities, n and p, respectively, given by

$$\vec{J_n} = -ne\vec{v_d} = \frac{ne^2\tau_{ct}}{m_e}\vec{\mathcal{E}} \tag{2.19}$$

$$\vec{J_p} = +ep\vec{v_d} = \frac{pe^2\tau_{ct}}{m_h}\vec{\mathcal{E}}. \tag{2.20}$$

The conductivity $\sigma = \vec{J}/\vec{\mathcal{E}}$, $\vec{J} = \vec{J_n} + \vec{J_p}$, of the semiconductor is, then:

$$\sigma = \frac{ne^2\tau_{ct}}{m_e} + \frac{pe^2\tau_{ct}}{m_h} = e\left(n\mu_n + p\mu_p\right). \tag{2.21}$$

The electron mobility in pure GaAs is $\sim$8500 $\mathrm{cm^2V^{-1}s^{-1}}$; the hole mobility is around 400 $\mathrm{cm^2V^{-1}s^{-1}}$.

[6] It can be shown that an ideal high-speed semiconductor should have a small, direct energy gap [11]. However, if the band-gap becomes too small, electron-hole pairs may form spontaneously, creating a leakage current that appears as a "dark current" in an optical device, or premature breakdown in a high-field condition [11].

Negative Differential Mobility [49]

At low values of the applied electric field $\mathcal{E}$ the electron drift velocity v_d is proportional to the field. At sufficiently high fields, the *differential mobility*, $\frac{\partial \mu_n}{\partial \mathcal{E}}$, in typical semiconductors becomes vanishingly small and the drift velocity reaches a value referred to as the *saturation drift velocity*, v_{sat}. Figure 2.3 presents the drift velocity - electric field characteristic of most common semiconductors.

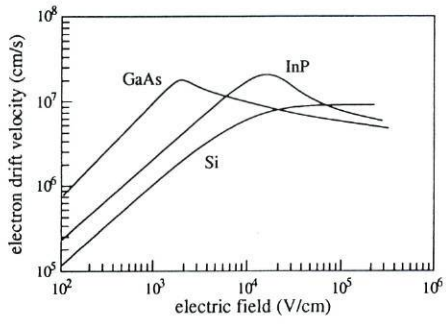


Figure 2.3: Velocity-field characteristic for carriers in the most used semiconductors [11].

In some semiconductors, such as silicon, the saturated drift velocity is the maximum drift velocity. However, certain semiconductors, notably direct-gap III-V compounds, display an anomalous peak in the drift velocity at a certain field strength (*threshold electric field*, $\mathcal{E}_{th}$), as shown in Fig. 2.3, above which they exhibit *negative differential mobility*. While velocity saturation occurs at sufficiently high fields in all semiconductors, the negative differential mobility is a property associated with a conduction band structure in some direct-gap III-V compound semiconductors. These display several heavy-mass satellite valleys [local minima of the energy dispersion $E(\vec{k})$] that are not too far separated in energy from the Γ-valley, which is responsible for conduction at low field values. Electrons energized by the electric field scatter from a high mobility, lower energy valley (Γ: $\vec{k} = \vec{0}$) into higher energy, low mobility valleys (e.g., L: $\vec{k} \neq \vec{0}$), leading to a decrease in the average electron velocity and, consequently, to *negative differential conductance* (NDC): *Gunn effect.*

Velocity Overshoot [11]

For field values below the threshold $\mathcal{E}_{th}$, the electrons remain in thermal equilibrium. However, if the device dimensions in the direction of charge flow are small enough, the carriers may pass through a high field region so quickly that few scattering events occur, being accelerated to higher velocities before relaxing to the equilibrium velocity: the *transient velocity overshoot effect.* The motion is termed *ballistic*, with the degree of velocity overshoot dependent on how long the carrier can avoid scattering; the overshoot will shorten the electron transit time through a high field region. Figure 2.4 shows the calculated electron velocity in submicron field-effect transistors as a function of carrier position, z, within the channel, making clear that the characteristic overshoot length (i.e., the distance of electron travel at velocity above the steady-state value) in GaAs can

be as long as 1 μm.

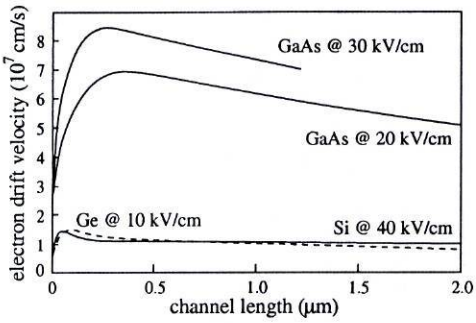


Figure 2.4: Velocity overshoot dependence on the carrier position within the channel [11].

For longer distances, carriers should regain thermal equilibrium and thus revert to velocity values cited in Fig. 2.3. If there were no collisions at all, the carrier velocity would be determined by the conservation of energy, $mv^2/2 = e\phi(z)$ [11], where $\phi(z) = Ez$ is the electrostatic potential.

Scattering Mechanisms [11][49]

Scattering effects influence carrier mobility and hence semiconductor conductivity, limiting the performance of both electronic and optoelectronic devices. Two important scattering mechanisms are *lattice scattering* and *impurity scattering.* Lattice scattering results from thermal vibrations of the lattice ions around their positions of minimum energy at any temperature above absolute zero. In quantum mechanics these modes of lattice vibration are called *phonons.* The periodicity of the lattice causes the phonon energy to have a band structure in much the same way as the electrons. Two kinds of lattice vibrations, denoted by *acoustic* and *optical* phonons, are possible.[7] Additionally, there are two *transverse* and one *longitudinal* modes of vibration for each kind of lattice vibration. These vibrations disturb the lattice periodic potential and allow energy to be transferred between carriers and the lattice, this becoming the dominant scattering mechanism at high temperatures; as a consequence, the mobility decreases with increasing temperature.

In compound semiconductors with two different atoms in the basis of the crystal structure a dipole momentum exists [53], and an extremely important scattering potential arises additionally from optical phonons. When optical vibrations take place, the effective dipole in the unit cell vibrates causing polarization fields from which the electron scatters: *polar optical phonon scattering* [53]. In materials like GaAs and some other III-V semiconductors, optical polar phonon scattering is the dominant scattering process in the Γ-valley at high field or room temperature, with scattering rates of $10^{12} - 10^{13} s^{-1}$ [53]. Optical polar phonon scattering can determine the ultimate bandwidth of devices, such as lasers and Schottky diodes.

Impurity scattering results when a charge carrier travels past an ionised dopant impurity, as its path will be deflected by the Coulomb interaction. Its probability depends on the total

[7]The acoustic phonons can be characterized by vibrations where the atoms in the basis of the unit cell (i.e., Ga and As for GaAs) vibrate with the same sign of the amplitude. In the case of optical phonons, the two atoms on the unit cell vibrate in opposing motion.

concentration of ionised impurities, and decreases with an increase in temperature, as this causes the carriers to travel faster and be less affected by the ionised impurities. For GaAs with $n \sim 2 \times 10^{16}$ cm^{-3}, the electron mobility decreases to ~4000 cm^{2}V^{-1}s^{-1}[11].

Carrier Injection and Diffusion current [49]

In the operation of electronic and optoelectronic devices, electrons and holes are injected into a semiconductor region either by an external biasing circuit or by absorption of light, usually under non-equilibrium conditions, $p \cdot n > n_i^2$ (*excess carrier injection*). In some cases, optical excitation for example, the carriers can be injected at energies much greater than the thermal energies ($\sim k_B T$), and are called *hot carriers.*

When carriers are injected into a semiconductor, they will gain or lose energy until they are in thermal equilibrium with the lattice: *carrier thermalization.* The *carrier thermalization relaxation time* is an important parameter for high speed performance of electronic and optoelectronic devices that depend upon carrier injection at high energies. Polar optical phonon scattering is responsible for relaxation of hot electrons injected into a semiconductor. Hot electrons in GaAs lose energy at the rate of about 0.16 eV/ps due to the emission of optical phonons, typically, with carrier relaxation time around 1 ps [11].

Whenever there is a gradient in the concentration of a species of mobile particles, the particles diffuse from the regions of high concentration to the regions of low concentration according to the third law of thermodynamics: *carrier diffusion.* For example, the excess carriers generated in the optically illuminated region of a semiconductor will diffuse towards the portions not affected by light.

An electron gradient concentration, $n(z,t)$, in the absence of an external electric field, gives rise to a diffusion current density:

$$J_{n,dif_n}(z,t) = e\frac{l_{cp}^2}{2\tau_{ct}}\frac{dn(z,t)}{dz} = eD_n\frac{dn(z,t)}{dz}, \tag{2.22}$$

where D_n is the diffusion coefficient for electrons, l_{pc} is the electron mean free path between collisions, and τ_{ct} is the mean collision time (a similar expression applies to holes). In many electronic devices the charge moves under the combined influence of an applied electric field and a concentration gradient. The total electron current is then:

$$J_n = -e\mu_n n(z,t)E + eD_n\frac{dn(z,t)}{dz}. \tag{2.23}$$

The relation between the mobility and the diffusion coefficient, is the *Einstein relation*:

$$\frac{D_n}{\mu_n} = \frac{k_B T}{e},$$

which is valid at equilibrium for the non-degenerate case.

Similar expressions can be written for holes in the valence band. It must be stressed that,

when using these expressions, the mobility is not constant in general, but only at low electric fields.

Impact Ionization [11]

In the transport mechanisms considered in the previous sections, the carriers (electrons and holes) remain in their respective band during transport. At very high electric fields, this is not always true and the semiconductor can suffer a *breakdown*, in which the current has a "runaway" behaviour due to *impact ionization*. Under a very high electric field, the electron acquires so much energy that it can scatter a valence band electron into the conduction band, originating an extra electron in the conduction band and one hole in the valence band. Highly energetic holes can undergo a similar process. The initial current J_i becomes $M \times J_i$, where M is a multiplication factor that depends upon the impact ionization rate, which varies with the material bandgap and the applied field. The critical electric field - breakdown field, $\mathcal{E}_{bdown}$ - is reached, for a given semiconductor, when the impact ionization rate approaches 10^4cm^{-1}, that is, there is about one impact ionization event when a carrier travels a distance of one micron. ($\mathcal{E}_{bdown}$ is $\simeq 4 \times 10^5$ V/cm for GaAs and $\simeq 2 \times 10^5$ V/cm for $In_{0.53}Ga_{0.47}As$ [53].) Once the applied bias becomes so large that $\mathcal{E} \cong \mathcal{E}_{bdown}$, the impact ionization process becomes dominant and the current rapidly increases due to carrier multiplication and control over the current is lost.

In narrow bandgap materials, breakdown can also occur by a mechanism which involves band to band *tunnelling* of electrons [11][22], which will be discussed next. This process is important in avalanche devices [49].

Impact ionization generation-recombination processes usually depend on the incident current densities instead of the carrier concentration, with the ionization coefficients related to the electric field in the ionization region.

Recombination and Generation [12]

The basic generation-recombination processes in semiconductors can be classified as *radiative* or *non-radiative*. The radiative mechanisms involve the creation or annihilation of photons. The non-radiative transitions do not involve photons; they may involve the interaction with phonons or/and exchanges of energy and momentum with other carriers.

The recombination and generation processes can be discussed in terms of band-to-bound state transitions and band-to-band transitions (interband transitions). Interband recombination, where an electron-hole pair recombines, is, in general, a radiative process; the recombination rate, R, of electrons and holes should be proportional to the product of the electron and hole concentrations:

$$R_n = R_p = C_c\, n \cdot p, \tag{2.24}$$

where C_c is a capture coefficient.

The generation rates being $G_n = G_p = g$, give the net recombination rate:

$$R = R_n - G_n = R_p - G_p = C_c n \cdot p - g. \tag{2.25}$$

If there is no external perturbation, such as electrical or optical injection of carriers, the generation rate is written as

$$g = C_c n_0 \cdot p_0, \tag{2.26}$$

where n_0 and p_0 are the electron and hole concentrations in thermal equilibrium ($n_0 \cdot p_0 = n_i^2$).

Auger recombination/generation occurs when an electron (or hole) is emitted/captured with the released/required energy taken/supplied by another electron or hole. It corresponds to electron (or hole) excess energy transfer to either an electron or a hole after their recombination, without producing photons. Auger processes are relatively unimportant in semiconductors with bandgaps larger than approximately 1.4 eV (e.g., GaAs, AlGaAs, and InP), but are a serious non-radiative mechanism in narrow bandgap materials (such as $In_{0.53}Ga_{0.47}As$, $E_g \sim 0.8$ eV), and are thus a serious hindrance for development of long wavelength lasers [18].

Tunnelling [11]

In Quantum Mechanics, the term *tunnelling* describes a host of phenomena associated with the existence of finite solutions of a particle wave function in regions where the potential energy is higher than the particle total energy [54].

In a semiconductor, an electron is constrained to move within a given band as long as it does not absorb enough energy to be scattered into the next higher energy band. However, under a strong electric or magnetic field, electrons can tunnel through the bandgap into the next band: *Zenner tunnelling* [23]. Under an applied electric field, the semiconductor bandeges are tilted because of the potential energy $e\mathcal{E}z$ from the electric field $\mathcal{E}$, as indicated in Fig. 2.5.

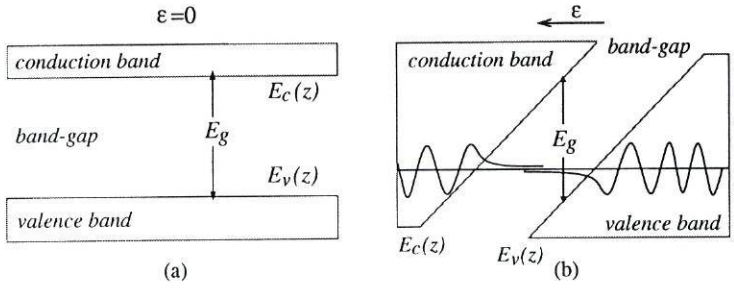


Figure 2.5: Bands of a one dimensional system under an applied electric field. Considering a charge $-e$ in a uniform electric field $\mathcal{E}$ along z, its potential energy is $U(z) = e\mathcal{E}z$ [23].

The probability of tunnelling from the valence band into the conduction band under high field conditions may be approximated by considering tunneling through a triangular potential barrier dependent on z and constant in time [23]. For a crude estimate of the tunnelling rate, the gap barrier is taken to be of height $U_0(z) = E_g$, decreasing linearly from E_g at $z = 0$

to 0 at $z = d = E_g/e\mathcal{E}$ (the barrier thickness). The decay wave number in the gap is then $\kappa \approx \sqrt{2mE_g/\hbar^2}$, and the probability of tunnelling, $T(\mathcal{E})$, can be derived in the WKB (Wentzel-Kramers-Brillouin) approximation as:

$$T(\mathcal{E}) \propto \exp\left[-2\int_0^d \kappa z\right] = \exp\left[-\frac{2\sqrt{2mE_g^3}}{e\mathcal{E}\hbar}\right], \tag{2.27}$$

where m is the effective mass of the tunnelling carrier. The tunnelling rate may become significant if the field is large or the bandgap is narrow. Taking n as the number of valence electrons tunnelling through the gap and v_{th} as the thermal velocity of the electrons, the tunnelling current density can be written as

$$J_{tun} = env_{th}T(\mathcal{E}). \tag{2.28}$$

Three basic conditions are necessary for tunnelling: (1) occupied energy states at the incidence region; (2) unoccupied energy states at the same energy levels as in (1) at the exit region; and (3) low potential barrier height and short tunnelling distance so that there is a finite transmission coefficient.

Tunnelling is very common in semiconductor devices: ohmic contacts to heavily doped material are possible precisely because of this effect [49]; a heavily doped p^+-n^+ junction achieves *negative differential conductance* (NDC), over a small range of forward bias voltages [49], by the mechanism of carrier tunnelling through a quantum-mechanical barrier.

Poisson's and Current Continuity Equation [12]

The Poisson's equation can be derived from the Maxwell's equations using the *quasi-electrostatic fields approximation.* For devices with a dc or low frequency bias, i.e., when the time variation is slow, Faraday's law is given by $\nabla \times \vec{\mathcal{E}} \simeq \vec{0}$, and in this case the solution of the electric field can be put in the form of the gradient of an electrostatic potential ϕ, $\vec{\mathcal{E}} = -\nabla\phi$. Gauss's law (Poisson's equation) in a semiconductor becomes

$$\nabla \cdot (\epsilon \nabla \phi) = -\rho, \tag{2.29}$$

with ρ being the charge density given by $\rho = e\left(N_d^+ - N_a^- + p - n\right)$, where N_d^+ and N_a^- are the ionized donor and acceptor concentrations, p and n are hole and electron concentrations, respectively. In the case of a n-type semiconductor invariant in the x and y directions, the Poisson's equation becomes

$$\frac{d}{dz}\left(\epsilon(z)\frac{d\phi}{dz}\right) = -e\left[N_d^+ - n(z)\right]. \tag{2.30}$$

This fundamental equation relates the charge distribution with the electric field.

Carrier charge conservation, incorporating generation (G) and recombination (R) effects,

results in the current continuity equations for electrons and holes:

$$\frac{\partial n}{\partial t} = G_n - R_n + \frac{1}{e}\nabla \cdot \vec{J_n} \tag{2.31}$$

$$\frac{\partial p}{\partial t} = G_p - R_p - \frac{1}{e}\nabla \cdot \vec{J_p}. \tag{2.32}$$

2.1.4 The *pn* and Metal-Semiconductor Junctions

Junctions between differently doped semiconductors, between metals and semiconductors, and between different semiconductors, provide the building blocks of devices [11][12]. This section reviews the fundamentals of semiconductor *pn* and semiconductor-metal junctions, pointing to their optoelectronic applicattions.

The *p-n* Junction [49]

The *n*-type and *p*-type regions of a *pn* junction can be formed in intrinsic semiconductor materials by selectively introducing slight deviations from stoichiometry with dopants or impurities. The fabrication techniques used to form *p*- and *n*-type regions of a junction involve i) epitaxial procedures (to be discussed), where the dopant species are simply switched-in at a particular instant during growth; ii) ion-implantation, in which the dopant ions are implanted at high energies into the semiconductor; iii) diffusion of dopants into an oppositely doped semiconductor. The *p-n* junctions ii) and iii) are obviously not as abrupt as in the case of epitaxial techniques. Fig. 2.6 gives the schematic view of a *pn* junction and the corresponding band diagram in the absence of applied bias.

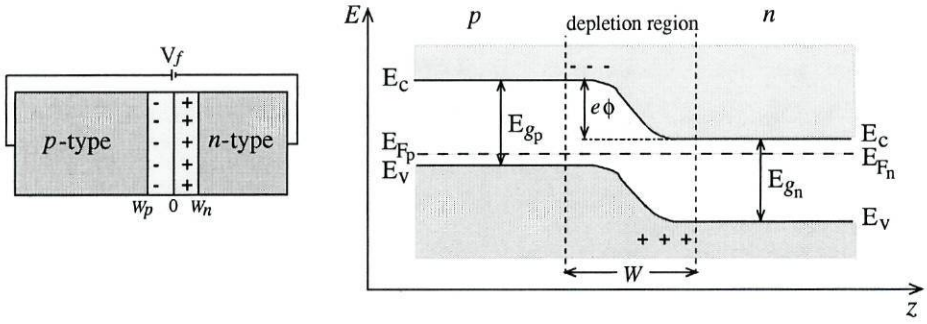


Figure 2.6: The *pn* junction and its band profile structure under equilibrium.

Three regions can be identified: a) the *p*-type region, where the material is neutral and the bands are flat, as the density of acceptors N_A exactly balances the density of holes; b) the *n*-type region, also neutral and where the density of immobile donors N_D exactly balances the electrons density; c) the *depletion* region, which extends a distance W_p into the *p*-region and a distance W_n into the *n*-region, and where mobile charges do not neutralise the charges of the donors and acceptors. The depletion region is the result of electron diffusion from the *n*-side into the *p*-side, and hole diffusion from the *p*-side into the *n*-side, due to the difference in electron and hole

densities across the junction. The space charge distribution formed causes a potential difference ϕ (the contact potential barrier) and hence an electric field at the junction, which prevents further diffusion from occurring; a drift current exists which conterbalances the diffusion current, and an equilibrium configuration is reached where there is no net current flow.

The energy of an electron at the bottom of the conduction band is lower by $e\phi$ on the n side of the junction than on the p-side, Fig. 2.6. The opposite occurs for a hole in the valence band. The depletion region ($-W_p < z < W_n$) in the vicinity of the junction has a much higher resistivity than any other part of the semiconductor, and most part of any applied voltage will be dropped across this region.

When a forward bias voltage $V = V_f$ is applied to the crystal,[8] the potential barrier at the junction becomes $\phi - V_f$, Fig. 2.7. As a result, the electric field across the depletion region decreases, and the depletion width $W = W_p + W_n$ becomes narrower. The system is no longer under thermal equilibrium, and two quasi-Fermi energy levels, E_{F_p} and E_{F_n}, are assigned to the n- and p-regions, respectively, with $E_{F_n} - E_{F_p} = eV_f$, as represented in Fig. 2.7. As the potential profile is altered by the applied bias, the carrier profile changes accordingly, affecting the diffusion current component. Under forward bias, there will be a supply of minority carriers on each side of the junction, with the probability of carrier diffusion across the junction proportional to $\exp(eV_f/k_BT)$. Majority electrons (holes) on the $n(p)$-side surmount the potential barrier to migrate to the $p(n)$-side; regardless of the bias, electrons or holes that come into the depletion region are swept out and contribute to the same current, independently of the field.

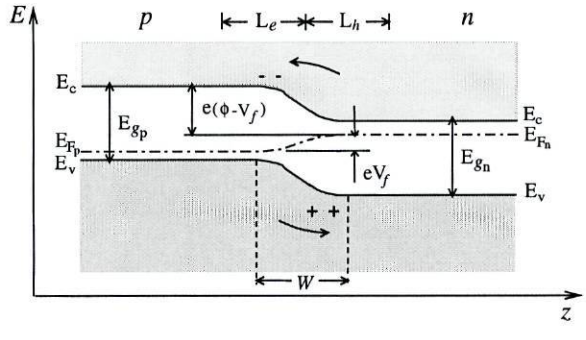


Figure 2.7: Schematic diagram of a pn junction band profile under forward bias, V_f .

Balance between drift and diffusion currents no longer exists, giving rise to a net current flow; the total current density due to both carriers can be expressed as

$$J = e\left(\frac{D_h p_n}{L_h} + \frac{D_e n_p}{L_e}\right)\left[\exp\left(\frac{eV_f}{k_BT}\right) - 1\right], \tag{2.33}$$

where p_n and n_p are the hole density on the n side and the electron density on the p side, respectively, and L_e and L_h are the electron and hole diffusion lengths, Fig. 2.7. The width W

[8] In forward bias V_f, the p-side is at a positive potential with respect to the n-side.

of the depletion region is given by

$$W = \sqrt{\frac{2\epsilon\,(\phi - V_f)}{e}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)}. \tag{2.34}$$

The diode equation, Eq. 2.33, which gives the current through a *pn* junction under forward bias ($V = V_f > 0$), also applies under reverse bias ($V = V_r = -V_f < 0$) condition, Fig. 2.8. However, under reverse bias the current saturates to a generally small value $J_0 = -e\left(\frac{D_h p_n}{L_h} + \frac{D_e n_p}{L_e}\right)$, typically less than 1 μA [49].

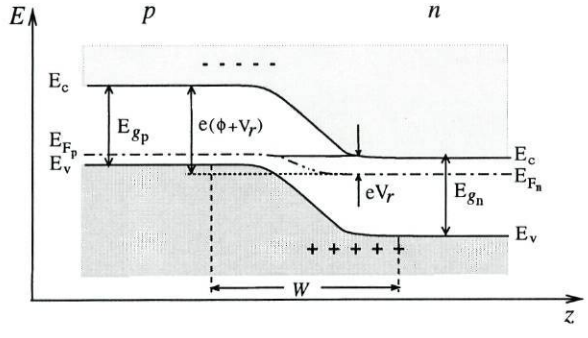


Figure 2.8: A *pn* junction band profile under reverse bias, V_r.

The current flow through the *pn* junction shows non-linear rectifying behaviour, Fig. 2.9; the diode current becomes significant (~mA) at the *cut-in voltage*, V_c, which is ~0.7 V for Si diodes and ~1.2 V for GaAs diodes [49]. As the reverse bias is increased, the diode current may abruptly run away *(breakdown)* as a consequence of impact ionization or Zener tunnelling; the electric field (voltage) at breakdown is known as the *breakdown field (voltage, V_B)*.

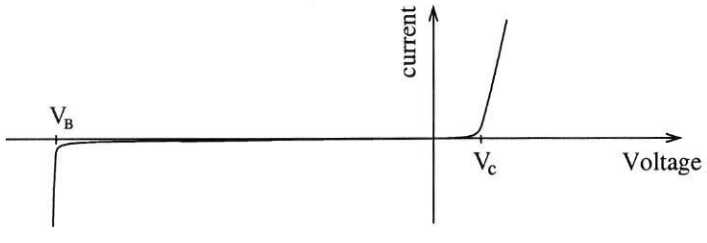


Figure 2.9: Diagram of the *pn* junction voltage-current I-V characteristic.

Semiconductor *pn* junctions have many optoelectronic applications, such as in *light emission, detection*, and *modulation* [15]. In a *pn* junction under forward bias, minority carriers are injected, electrons into the *p* side and holes in the *n* side, which can *radiatively* recombine with the majority carriers (typical recombination time around 10 ns), yielding photons with energy close to the semiconductor bandgap energy E_g. The resulting electro-luminescent device is a *light-emitting diode* (LED), which should use a semiconductor material with a direct gap.[9] In

[9]Radiative transitions in indirect gap materials require the assistance of phonons and so have emission yields many orders of magnitude lower than in direct gap semiconductors.

heavily doped (degenerate) junctions, application of a forward bias voltage results in carrier population inversion, leading to optical gain (by stimulated emission) [24]. If a feedback mechanism is provided (using the reflectivity at the crystal facets, gratings, or external mirrors), *laser oscillation* is obtained. AlGaAs/GaAs lasers emit at around 840 nm; InGaAsP and InGaAlAs materials have been developed for longer infrared wavelength lasers, 1.3 μm and 1.55 μm [25]. Semiconductor lasers can be modulated at much higher frequency (>20 GHz) than LEDs [18].

An important application of reverse biased *pn* structures is in *photo-detection* [15]. When light impinges upon a *pn* junction depletion region, many of the electron-hole pairs generated are split and the charge carriers swept rapidly by the electric field (electrons into the n-region and holes into the p-region), and then collected at the contacts. The photocurrent is independent of the applied bias, away from the breakdown regime (contrary to the case of the avalanche photodiode, APD). Since the electrons and holes contributing to the photo-current move under high electric fields, the response can be very fast. The foremost issue in detector design is the choice of a material which has a good absorption coefficient at the relevant optical frequencies. For applications where GaAs/AlGaAs ($E_g \sim 1.45$ eV) sources are used, Si detectors with absorption lengths of several micra are adequate, unless high speed is required. Fast response requires direct band gap materials, so that the active absorption length (depletion region) can be brought down to a micron or less to decrease the carriers transit time; InGaAs detectors are now widely used for high speed long haul communication applications [14]. *Schottky* diodes, discussed below, have the fastest response due to their lower parasitics [11]. Ultra high speed photo-detectors have, of course, small diameters (below 12 μm in [60]).

Another important optoelectronic application of a reverse biased *pn* structure is as *optical modulator* [12][13], which will be discussed later in this chapter.

Metal-Semiconductor Junctions [20][49]

Metal-semiconductor junctions are needed to connect the active region of a device with external circuitry. Their current-voltage characteristics are usually classified as either *ohmic*, Fig. 2.10(a), where the current through the junction and the voltage across it obey a linear relationship (independent of the polarity of the applied bias), or *rectifying* (or *Schottky*), where the I-V characteristic depends on the polarity of the applied voltage, Fig. 2.10(b).

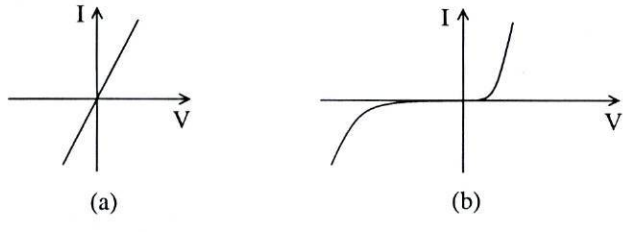


Figure 2.10: Current-voltage characteristic of (a) ohmic and (b) rectifying metal-semiconductor junctions.

Schottky Junction

When a metal is placed in intimate contact with a semiconductor, the valence and conduction bands of the semiconductor "bend" to make the Fermi levels in the metal and in the semiconductor equal (thermal equilibrium), as with the *pn* junctions. Figure 2.11 represents the metal-n type semiconductor junction band diagram. (Only the n-type semiconductor will be considered here.)

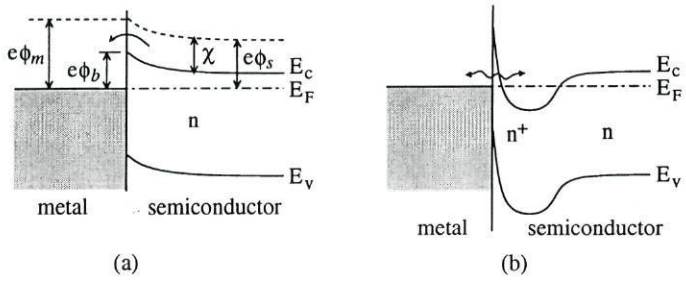


Figure 2.11: Metal-n semiconductor and metal-n^+ semiconductor junctions [20].

In the simplest approximation, the resulting potential barrier depends only on the work functions of the metal, ϕ_m, and of the semiconductor, ϕ_s, and on the electron affinity, χ. The potential barrier to electron flow from the semiconductor to the metal is $\phi_m - \phi_s$, and the potential barrier to electron flow from the metal is $\phi_b = \phi_m - \chi$ (the barrier height). Such metal-semiconductor junction is rectifying, and is known as *Schottky* junction. Its height can be experimentally measured, and is generally a function only of the semiconductor chosen. Typically, its value is close to 0.8 eV for GaAs, caused by the large density of surface states present on GaAs. It is difficult to make ohmic contacts to wide-gap semiconductors: a metal does not generally exist with a sufficiently low work function to yield a low barrier.

Carriers may traverse the potential barrier when their thermal energy is sufficient to pass over it, *thermionic emission*, Fig. 2.11(a), or by quantum mechanical *tunnelling* or *field emission*, Fig. 2.11(b), if the barrier is sufficiently narrow. A mixture of these mechanisms often occurs, and is referred to as *thermionic-field emission*. For most III-V semiconductors, bending of the bands at the surface is not as large as would be required to achieve thin barriers because of the problem of surface state pinning.

In all of these conduction regimes (thermionic, thermionic-field, or field emission), a potential difference must exist across the barrier to cause current flow. The metal-semiconductor contact is characterized by the specific contact resistance, r_c, defined as

$$r_c = \left[\frac{\partial J}{\partial V}\right]^{-1}_{V=0}, \tag{2.35}$$

per unit area, for current flow perpendicular to the contact (Ωcm^2 units; the contact resistance, R_c, is a function of device size, and is often given in Ωmm).

Thermionic emission is the major current flow mechanism in Schottky junctions formed on moderately doped GaAs. If V_f is the forward voltage applied across the metal-semiconductor

junction, the current density across the barrier is

$$J \propto \exp\left(-e\phi_b/k_BT\right)\left[\exp\left(eV/k_BT\right)-1\right], \tag{2.36}$$

giving (using Eq. 2.35)

$$r_c \propto \exp\left(\frac{e\phi_b}{k_BT}\right). \tag{2.37}$$

A small r_c requires low barrier height; Schottky junctions do not provide a good ohmic contact, and thermionic emission is thus seldom exploited.

Ohmic Junction

For a n-type semiconductor with $N_D \leq 10^{17}\text{cm}^{-3}$, the current is mainly due to thermionic emission, and r_c is essentially independent of doping. As the semiconductor doping level in the contact region increases, the width of the potential barrier decreases and tunnelling can occur, Fig. 2.11(b), and as a consequence the probability of electrons to tunnel across the barrier increases. The tunnelling current density across the junction is

$$J \propto \exp\left[-\frac{2}{\hbar}\sqrt{\frac{\epsilon m}{N_D}}\left(\phi_b - V\right)\right]. \tag{2.38}$$

From Eq. 2.35,

$$r_c \propto \exp\left[\frac{2\sqrt{\epsilon m}}{\hbar}\frac{\phi_b}{\sqrt{N_D}}\right], \tag{2.39}$$

which shows that, for a tunnelling contact, the specific contact resistance depends strongly on the barrier height and contact doping density, i.e., semiconductors with low barrier height and high doping concentrations must both be used to make a successful ohmic contact. The general technique for making an ohmic contact involves the realization of a heavily doped thin semiconductor layer, to obtain a metal-n^+-n junction, Fig. 2.11(b). In this case the resistance of the n^+ to n junction may be the limiting factor in lowering the contact resistance.

The required doping value for GaAs is on the order of 1×10^{19} cm^{-3}; however, low solubility of the n-type dopants in GaAs typically limits the value to 2×10^{18} cm^{-3} [20][56]. Various methods, such as shallow diffusion, alloy regrowth, in-diffusion of a dopant contained in the contact metals, and ion implantation can be used to obtain ohmic metal-n^+-n-GaAs junctions. The one usually adopted employs metal alloys which can highly dope the contact layer of the semiconductor [20]. For ohmic contacts on n-GaAs, a Au-Ge alloy (with 82% Au) is first evaporated onto the semiconductor; these metal-semiconductor junctions are then alloyed at the corresponding eutectic temperature into the semiconductor [20], in inert atmosphere (such as argon or nitrogen).

The contact resistances can be measured using the transmission line method [20][55].

2.2 Heterostructures

Artificial III-V and II-VI semiconductor materials have been investigated to increase the range of semiconductor properties available for applications [11][17]. Heterostructures are artificial semiconductor structures composed of two (or more) different types of material, where variations in composition are used to control the motion of electrons, holes and photons through *band engineering*, modifying the material electronic and optical properties [22]. Figure 2.12 shows the relationship between two important characteristics of semiconductor materials, the minimum energy bandgap and the lattice constant. The horizontal axis displays the crystalline lattice constant both in nanometre and as percentage mismatch to the silicon lattice constant of 0.5543 nm. The vertical axis represents the minimum energy bandgap (E_g) both in electron volts and as an effective wavelength $\lambda_g = 1239.8/E_g$. Light of wavelength λ smaller than λ_g will be absorbed by the semiconductor.

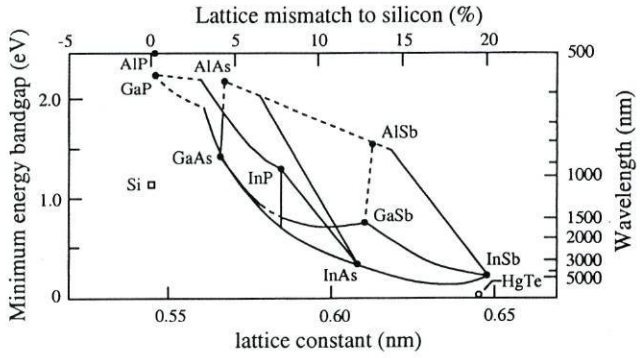


Figure 2.12: Energy bandgap versus lattice constant for some semiconductors [11].

The active regions of most heterostructures are typically a few nanometers thick, seat at/or close to single interfaces (*heterojunctions*) or involve many thin layers of different compositions. In general, when a heterojunction is, or belongs to, the active region, it must have a nearly perfect interface to perform well. A high quality heterostructure requires the materials to have the same crystal structure (or at least symmetry), and the two lattice constants to be nearly identical, if there is to be no strain in the final structure [15][17].[10] For example, an alloy such as GaAs-AlAs, which is generally denoted by $Al_xGa_{1-x}As$ (where x denotes the composition fraction of the chemically inter-changeable atom), and GaAs have an excellent lattice match: the lattice constant of $Al_xGa_{1-x}As$ changes less than 0.15% as a function of x, which makes possible to grow layers of GaAs and AlAs, or any of the intermediate alloys $Al_xGa_{1-x}As$, on top of one another, without significant stress at room temperature; InP is an alternative substrate, with two ternary alloys having the same lattice constant and direct bandgap: $In_{0.53}Ga_{0.47}As$ (abbreviated InGaAs) with energy bandgap 0.75 eV, and $In_{0.52}Al_{0.48}As$ (abbreviated InAlAs) with energy bandgap 1.439 eV.

[10]The initial effort was concentrated on materials of nearly identical lattice constant. Current applications require properties that can be met only by mismatched materials, using "strained layers" [24].

In layered heterostructures, films can be sufficiently thin, which requires a fabrication technology able to change the composition of successive layers very rapidly, preferably from one mono-layer to another [24]. Through epitaxial growth techniques, such as molecular beam epitaxy (MBE) and metal-organic chemical vapour deposition (MOCVD) [57], a range of artificial semiconductor heterostructures can now be tailored to the device needs; several building blocks of many present-day electronic and optoelectronic devices are layered structures grown by MBE and/or MOCVD [11][15].

The heterostructures employed in this work, optical waveguides incorporating a DBQW-RTD (AlGaAs/GaAs/AlGaAs on a "semi-insulating"-GaAs substrate and InAlAs/InGaAlAs/InP on a n^+-InP substrate), were grown by Molecular Beam Epitaxy in a Varian Gen II system by the MBE Group at the Department of Electronic and Electrical Engeering of the University of Glasgow, Scotland.

2.2.1 Molecular Beam Epitaxy

Molecular beam epitaxy is a heteroepitaxy crystal growth technique involving the reaction of one or more thermal beams of atoms or molecules with a crystalline surface (substrate) under ultra-high vacuum conditions [57]. It allows monolayer (~0.3 nm) control over the chemical composition, so abrupt changes in the optical and electrical properties of the semiconductor artificial crystals are attainable along the growth direction [22]; the crystalline orientation of the deposited layer is the same as that of the substrate [57]. Molecular beam epitaxial growth techniques allow engineering of the conduction and the valence bands of semiconductor heterostructures in one-dimension [22]; MBE can simultaneously achieve precise control over chemical composition and doping profiles, on single-crystal multilayer structures with thickness on the order of a few atomic layers [57]. The first DBQW-RTD showing negative differential conductance was grown by MBE [1].

The basic process of MBE is ultra-high vacuum evaporation [57]. A functional schematic of a MBE system is shown in Fig. 2.13(a). The apparatus consists of growth, auxiliary and load chambers, and each chamber has an associated pumping system. The auxiliary chamber may contain supplementary surface analytical tools not contained in the growth chamber, additional deposition equipment or other processing equipment. The load chamber is a smaller chamber connected to the growth chamber via a sample exchange load-lock, which permits maintenance of ultra-high vacuum in the growth chamber while changing substrates.

The growth chamber is shown in detail in Fig. 2.13(b); its main elements are: sources of molecular beams (K-cells) and shutters to occlude the molecular beams; substrate manipulator; gauge to measure chamber base pressure and molecular beam fluxes; a reflection high energy electron diffraction (RHEED) gun and screen to monitor film surface structure; and a quadrupole mass spectrometer to monitor specific background gas species or molecular flux compositions. The sources of molecular beams are effusion ovens where the elements that compose the heterostructure - here Ga, As, Al, and In - are evaporated; one or more effusion ovens may be

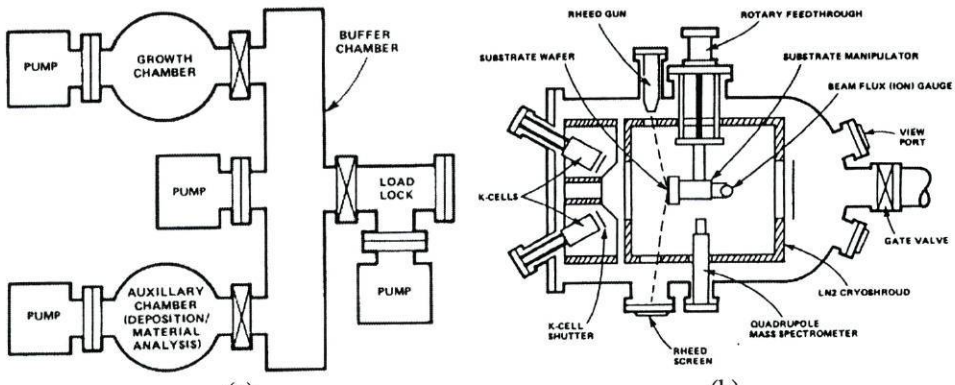


Figure 2.13: (a) Functional schematic of a basic MBE system. (b) Schematic cross-section of a typical MBE growth chamber [57].

used for dopants.[11] They are connected to the growth chamber via orifices directed towards the substrate, but shielded from it by shutters; growth commences when the shutters are opened.

The growth chamber is evacuated to ultra-high vacuum (UHV), typically 10^{-11} mbar or better; at this low pressure, the mean free path of the molecules between collisions is much larger that the width of the chamber (this is the Knudsen or molecular-flow regime of a gas) [57]. Molecules that emerge from the effusion ovens form a molecular beam, travelling in straight lines without collisions until they impinge on the substrate, or elsewhere. The flux of each element and hence the deposition rate is controlled through the temperature adjustment of each furnace. The heterostructures must be grown with extreme purity if their properties are not to be spoilt by contamination, so all materials used must have a high degree of purity. The background pressure in the chamber must be kept low to reduce contamination, as well as to ensure that the Knudsen regime holds. The flux from the effusion ovens must be uniform across the substrate, or there will be variations in composition across the wafer (which are minimised by rotating continuously the heated substrate during growth); the temperature of the furnaces must also be carefully controlled to keep the flux constant.

The morphology of the surface depends in a complicated way on substrate temperature (for example, AlGaAs and GaAs grow better under different conditions), which is also important in other ways [57]: defects will not have time to be removed by annealing if the substrate temperature is low, while unwanted diffusion will occur and blur the interface if the temperature is too high. Hence, the optimum growth conditions for each compound are met almost empirically. As mentioned, several in situ diagnostic techniques are used to monitor growth, including mass spectrometry for residual gas analysis and reflection high energy electron diffraction [57]; uniformity of the epitaxial layers is typically 1% in doping and 0.5% in thickness.

Advantages of MBE include highly abrupt junctions between different materials, good control

[11] The dopant used in the heterostructures employed in this work was Si, which is in group IV of the periodic table. It can act as a donor or acceptor in a group III-V compound (it also tends to become amphoteric). In practice, it usually acts as donor, but this can be changed by growing on a different surface orientation from the usual (100) or using very high concentrations (around 10^{19}cm^{-3}) [56].

over the thickness of layers, and reasonable reproducibility. Disadvantages include cost and the fact that the process does not scale well for production [15][57]: MBE is a slow process, growing materials at about 1 mono-layer per second or 1 μm per hour; the heated sources and substrate in a UHV environment place a tremendous demand on the vacuum systems used.

The next section discusses in more detail AlGaAs/GaAs and InGaAlAs heterojunction interfaces grown on GaAs and InP substrates, respectively.

2.2.2 Heterojunctions and Band Engineering

The fundamental reason for growing heterostructures is the opportunity they offer to manipulate the behaviour of electrons and holes through bandgap engineering [22].

The heterojunction between two materials, such as GaAs and AlGaAs, with $E_g^{GaAs} < E_g^{AlGaAs}$, is first considered. Significant bandgap differences can be obtained in the $Al_xGa_{1-x}As$ system, as the band gap is given by $E_{\Gamma g}(x)$ =1.424 + 1.247x eV for $x < 0.45$, and $E_{\Gamma g}(x)$=1.656 + 0.215x + 1.147x^2 eV for $x > 0.45$ [23]. Below the critical composition $x = 0.45$, the bandgap is direct; above this composition, the conduction band minima are at X points of the Brillouin zone ($E_{Xg}(x)$=1.900 + 0.125x + 0.143x^2 eV; $E_{Lg}(x)$=1.708 + 0.642x eV), while the valence band maximum remains at the Γ-point [23].

When two materials are joined, knowledge of band alignment at the junction is essential to predict and take advantage of the new structure properties. Although it is not easy to determine exactly the energy bands configuration even for the most-studied junction, GaAs-AlGaAs, a variety of models can be applied for design of specific structures [22][23]. Anderson's rule [23] states that the vacuum levels of the two materials making a heterojunction should be lined up, as in Figure 2.14(a). From Fig. 2.14(a), it results that $\Delta E_c \equiv E_c^{AlGaAs} - E_c^{GaAs} = \chi^{AlGaAs} - \chi^{GaAs}$ which, for GaAs ($\chi = 4.07$ eV) and for $Al_{0.3}Ga_{0.7}As$ ($\chi = 3.74$ eV), predicts $\Delta E_c = 0.33$ eV; the bandgap changes by $\Delta E_g \simeq 0.37$ eV, so $\Delta E_v = 0.04$ eV [23]. The fraction of the bandgap that has gone into the conduction band is $Q = \Delta E_c / \Delta E_g \approx 0.88$, according to this model; the value $Q \approx 0.60$ is now established for $x < 0.45$ [22].

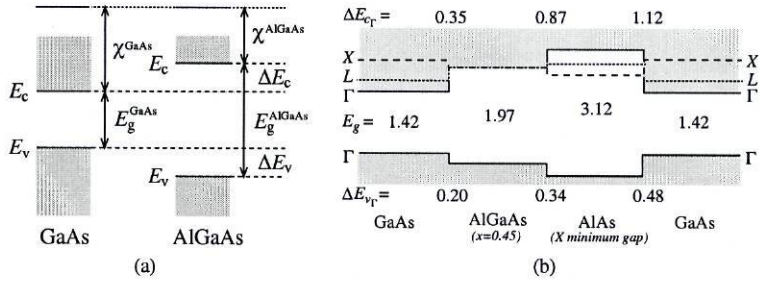


Figure 2.14: Alignment of bands at the heterojunction between GaAs and AlGaAs [23].

A difficulty can arise when the components of a heterostructure have band minima at different points in the Brillouin zone. Figure 2.14(b) shows the energies of the three lowest minima in the conduction band of $Al_xGa_{1-x}As$ for three values of x. If the Γ minimum is the lowest throughout

a $Al_xGa_{1-x}As$-GaAs structure, it is reasonable to neglect the higher energy bands, provided that the energy of the incident electron is well below the top of the higher $Al_xGa_{1-x}As$ barrier present. In the case of AlAs (for example, in a GaAs/AlAs/GaAs sandwich), it appears as a potential barrier if Γ-valleys are considered, but is a *potential well* from the point of view of the X-band, which lies lowest inside the barrier. This specific case will be discussed in more detail in Chapter 3. The AlGaAs-GaAs band alignment is called *type I* or *straddling* alignment [23].

A heterojunction of InP and $In_{0.52}Al_{0.48}As$, Fig. 2.15(a), shows *type II* or *staggered* alignment [23], where the bands favour electrons in InP but holes in InAlAs, thus a sandwich results in a potential well for one species but a barrier for the other. The other possibility is that the two bandgaps do not overlap at all, giving *type III* or *broken-gap* alignment [23]. The classical example is InAs-GaSb, also shown in Fig. 2.15(b). In this case the conduction band of one material overlaps the valence band of the other, by 0.15 eV in this example.

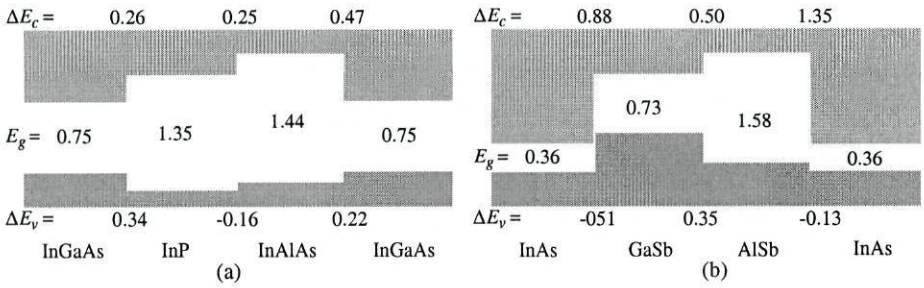


Figure 2.15: Band alignment in InGaAs-InAlAs-InP and InAs-GaSb-AlSb heterostructures [23].

Heterostructures also give control over the position of the extra carriers introduced by doping, allowing physical separation of the carriers from the ionized impurites [22][23]. This is illustrated, for a heterojunction between a layer of AlGaAs doped with donor impurites and an undoped GaAs layer, in Fig. 2.16.

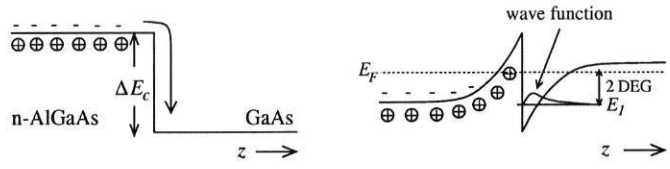


Figure 2.16: Conduction band discontinuity in a n-AlGaAs-undoped GaAs heterojunction [23].

Most of the released electrons in the n-AlGaAs region close to the interface diffuse into the GaAs region and cannot move back up to the AlGaAs conduction band because of the potential barrier due to the conduction band discontinuity. As a consequence, an electric charge distribution is created across the interface (positive ions in the AlGaAs region and electrons in the GaAs region), giving rise to an electrostatic triangular potential well $\phi(z)$ [22] which squeezes excess electrons in the GaAs material against the interface, in a region typically 10 nm wide, where they are trapped, forming a two-dimensional electron gas (2DEG) [22][23]: these electrons are free to move in the plane parallel to the interface, unhindered by the impurity atoms, with

much higher mobility than in bulk GaAs. This kind of structure permits the implementation of much faster devices, and is the basis of new electronic devices [11][22].

2.2.3 Layered Structures: Quantum Wells and Barriers

Research in the field of low dimensional layered semiconductor structures started in the early seventies with the strong impulse supplied by the prediction of new physical phenomena and novel devices, derived from the engineered conduction and valence bands [22].

A semiconductor structure shows *classical* electrical behaviour (*ohmic behaviour*) if its dimensions are much larger than each of these characteristic length scales [30]: (1) the carriers *de Broglie wavelength*, (2) the carriers *mean free path*, and (3) the *phase-relaxation length*, which is the distance that an electron travels before its initial phase is destroyed. Small structures, whose dimensions are much larger than those of the atoms but not large enough to be ohmic, are called *mesoscopic*. These length scales vary widely from one material/structure to another and are also strongly affected by temperature, electric and magnetic fields [30].

A brief survey of some of the structures that can be grown is presented next. (Only the behaviour of electrons in the conduction band is considered, but similar considerations apply to holes in the valence band, subject to the complications of the valence band).

Potential Barrier [23][54]

An example of an electron potential barrier is provided by the conduction band profile of a layer of AlGaAs surrounded by two thick layers of GaAs, as shown schematically in Fig. 2.17(a); the AlGaAs layer acts like a barrier because E_c is higher in $Al_xGa_{1-x}As$ than in GaAs, and the difference ΔE_c provides the barrier height U_0.

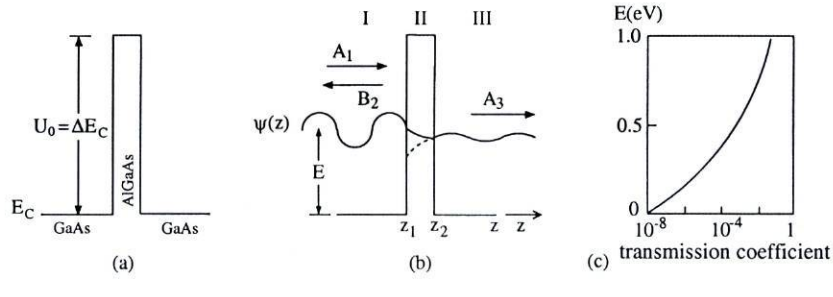


Figure 2.17: Schematic representation of tunnelling in a potential barrier.

Classical physics predicts that a particle with energy E, travelling towards a potential barrier of height U_0, will not be able to overcome the barrier, being reflected, unless it has sufficient kinetic energy to pass over the top. On the contrary, according to Quantum Mechanics, the particle has a finite probability to be transmitted or tunnel through the potential barrier.

A schematic representation of the wave functions across the barrier is shown in Fig. 2.17(b). Assuming an incident-particle from the left, represented in region I ($x < 0$) by an incident wave function (with amplitude A_1) and a reflected wave function (with amplitude B_1); in region

III ($x > d$, $d = z_2 - z_1$), by a transmitted wave function (with amplitude A_3); in region II ($0 < x < d$), the wave function contains components decaying in both directions, with a phase diference between them (with amplitudes A_2 and B_2, respectively).

The Schrödinger equation (Eq. 2.2) for the electron in any region can be written as

$$-\frac{\hbar^2}{2m_i}\left[\frac{d^2\psi_i(z)}{dz^2}\right] + U_i(z)\psi_i(z) = E\psi_i(z), \tag{2.40}$$

where the subscript i indicates the $i\,th$ region. The wave function solutions can be expressed as

$$\psi_i(z) = A_i \exp(jk_i z) + B_i \exp(-jk_i z), \tag{2.41}$$

where A_i and B_i are constants to be determined from the boundary conditions, and $k_i = \sqrt{\frac{2m_i(E-U_i)}{\hbar^2}}$.

Since the wave functions and their first spatial derivatives (i.e., $d(\psi_i/m_i)/dz = d(\psi_{i+1}/m_{i+1})/dz$) at each potential discontinuity must be continuous, solution of the corresponding equations gives the transmission coefficient across the barrier.

Assuming a constant effective mass across the three regions, the transmission coefficient is given by [2]

$$T(E) = \left(\frac{A_3}{A_1}\right)^2 = \left[1 + \frac{U_0^2}{4E(U_0 - E)}\sinh^2 k_2 d\right]^{-1}. \tag{2.42}$$

When $k_2 d \gg 1$, the transmission coefficient varies as [Fig. 2.17(c)]

$$T(E) \approx \frac{16E}{U_0}\exp(-2k_2 d) = \frac{16E}{U_0}\exp\left[-2d\sqrt{2m\frac{(U_0 - E)}{\hbar^2}}\right]. \tag{2.43}$$

For a given energy value E, to have a finite transmission coefficient the barrier must be thin, have a small height U_0, and a small effective mass m.

Single-barrier tunnelling has found widespread applications in semiconductor devices, such as heterojunction barrier varactor diodes [11][63][86].

Quantum Wells [23][54]

The opposite of the barrier structure is a *quantum well* (QW), Fig. 2.18, which is obtained by sandwiching a thin layer (a few nanometers thick) of a given bandgap material, for example GaAs, between thicker layers of a wider bandgap material, for example AlGaAs.

Electrons in the conduction band and holes in the valence band tend to move to the lower bandgap region, which acts as a well for electrons in the conduction band and for holes in the valence band. Carriers in a quantum well are confined only in the direction perpendicular to the layers, and allowed energy states in the growth direction are discrete and quantized.

Quantum wells in the GaAs-AlGaAs system are far from the idealized case of an infinite depth quantum well; the ideal "square" quantum-well consists of a finite region of space, of width

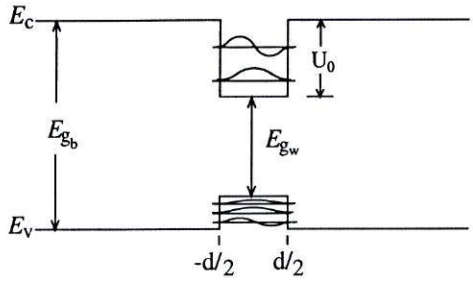


Figure 2.18: Quantum well schematic energy diagram.

d, bound by infinitely high barriers. For the infinite depth QW the allowed energy states can be analytically calculated, and are often used as a first approximation to the solutions of a real quantum well.

Assuming the growth direction as the z-direction, the electron has zero potential energy in the well region $-d/2 \leq z \leq d/2$ [i.e., $U_0(x, y, -d/2 \leq z \leq d/2) = 0$]. Schrödinger equation solutions for particle motion in the $z-$direction inside the well are identical to those for free space, where the potential energy is assumed to be zero:

$$\psi_{z_i}(z) = A_{z_i} \exp(+ik_{z_i}z) + B_{z_i} \exp(-ik_{z_i}z). \tag{2.44}$$

The infinitely high potential energy regions outside the well prevent the electron from straying beyond the well region. The wave function must remain well behaved across the whole structure; as in the barriers the Schrödinger equation takes the form

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{z_i}(z)}{dz^2} = (E - U)\,\psi_{z_i}(z), \tag{2.45}$$

for the term $U(z)\psi_{z_i}(z)$ to remain finite, $\psi_{z_i}(z) = 0$ must stand inside the barriers. Discontinuities in the wave function are not allowed, so $\psi_{z_i}(z)$ must vanish at the edge of the barriers; the boundary conditions are, therefore, $\psi_{z_i}(z = -d/2) = 0$ and $\psi_{z_i}(z = d/2) = 0$. The electron eigen-states and allowed energies are, then

$$\psi_{z_n}(z) = A_{z_n} \sin k_{z_n} z = A_{z_n} \sin \frac{n\pi}{d} z \tag{2.46}$$

$$E_{z_n} = \left(\frac{\hbar^2\pi^2}{2m\,d^2}\right) n^2, \tag{2.47}$$

i.e., the wave number k_{z_i} and the energy E_{z_i} in the z direction are quantized; the integer $n = 1, 2, 3, \ldots$ is a *quantum number* that identifies these eigen-states.[12]

The electrons are free to move in this x-y plane, where there is a continuous energy state

[12] R. Dingle *et al* (1974) performed optical measurements which showed directly the quantization of energy in QWs [22].

distribution:

$$E(x, y, n) = \left(\frac{\hbar^2 \pi^2}{2md^2}\right) n^2 + \frac{k_x^2 + k_y^2}{2m}. \tag{2.48}$$

In pratical quantum wells, when the separation between these energy levels is much larger than the thermal energy $k_B T$, the particles appear to be frozen into the ground state ($n = 1$) and no motion will be possible ouside the well plane.

A *quantum wire* is formed if carriers are confined in a two dimensional potential well (analogue of an optical fibre), where the carriers are free to move in one dimension only. When the electrons are restrained to a finite 3-D volume in space, a *quantum-box* or *quantum-dot* is formed [22].

Double-Barrier Quantum Well [2][23]

The high bandgap regions (barriers) on either side of the low bandgap layer (well) are usually made much thicker than the well region, resulting in *bound states* [54]. However, if the quantum well barriers are made thin enough and are surrounded by the same material as the well layer, electrons in the well can now tunnel through the barriers and the states are *quasi-bound* or *resonant states*, rather than true bound states [84][85]. The structure obtained forms a *double-barrier quantum well* (DBQW), Fig. 2.19.

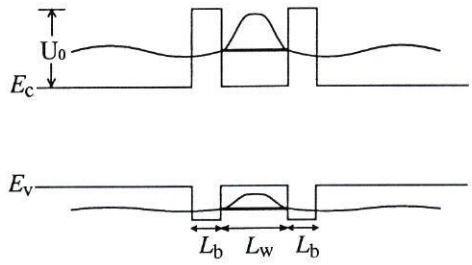


Figure 2.19: Double barrier quantum well.

Electron tunnelling through the barriers is strongly enhanced when its energy equals to one of the energy levels in the well - theoretically, a transmission probability of 1 is possible [87]. This corresponds to *resonant tunnelling* [1][27], and the DBQW structure acts as a filter such that only electrons with energy close to the resonance energies have high probability of crossing the two barriers. This unique electron tunnelling phenomenon was predicted and demonstrated in 1974 by Chang, Esaki and Tsu [1], for an AlGaAs/GaAs DBQW diode (the resonant tunnelling diode, RTD). RTDs show negative differential conductance at room temperature.

The next chapter contains a discussion, in greater detail, of the RTD structure and its applications in Electronics and Optoelectronics.

Superlattices and Multiple-Quantum Wells [22][23]

Alternating layers of wells and thin barriers, as shown in Fig. 2.20(a), form an artificial periodic one-dimensional structure, known as a *superlattice*, which imposes a second level of periodicity on

that given by the crystalline nature of the semiconductor. In a superlattice, the larger bandgap layers are thin enough so the wave function of the states in different wells do overlap, resulting in non-localised states which form energy *minibands*. An electron can now tunnel from one well into its neighbours. This new potential periodicity affects motion only along the direction of growth and is governed by Bloch's theorem (the period is, of course, longer than that of the crystal and the periodic potential is also weaker, with the result that the energy bands and gaps appear on a much smaller scale of energies).

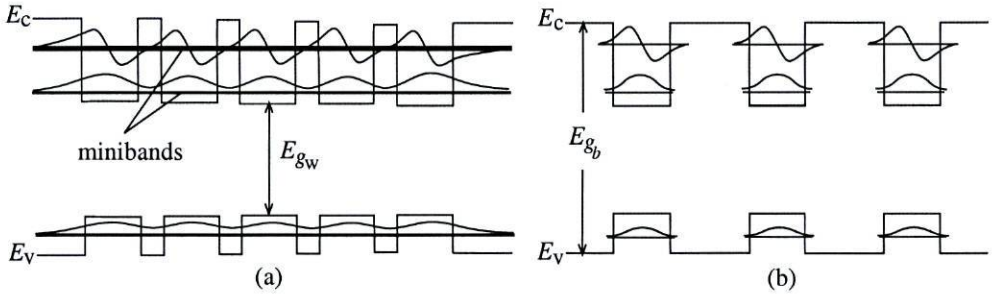


Figure 2.20: Superlattice and *multi-quantum well* schematics.

As the barriers between wells are made thicker, the minibands become narrower due to a lower wavefunctions overlap, which makes tunnelling between wells less important [22]. In the limit, in which wells are practically isolated from their neighbours , the superlattice is known as a *multi-quantum well* (MQW), Fig. 2.20(b), and the quantized states are localised.

Superlattices have been used to filter the energy of electrons, allowing only those with energy within the minibands to pass and reflecting those in the minigaps; absorption between the minibands can be used to detect infrared radiation [11]. It is also a common practice to grow a superlattice on top of the original substrate, then a buffer, and finally the desired structure; the interfaces trap many defects and impurities that would otherwise migrate up with growth and contaminate the active region of devices. MQW are widely used in optical applications such as in lasers [24], optical modulators [19] and detectors [15].

2.2.4 Optical Field Confinement

Propagation of electrons in mesoscopic systems has quite interesting similarities with the propagation of light, although the vector nature of light makes the calculations more complicated [12][59]. Electrons tend to be confined in regions of lower potential energy, while light can be confined to regions of higher refractive index, as represented in Fig. 2.21(a). When a simple double heterostructure guides light effectively its width is so large that quantization of electronic states is unimportant. The solution is to employ a *separate confinement heterostructure*, as indicated schematically in Fig. 2.21(b), where light and particles are confined separately, with a narrow inner well for charged particles and a broader outer region for light [24]. This structure is widely used in lasers to maintain an optimum trade-off between low optical power density, high *confinement factor*, and low barrier to charge transport [18][24].

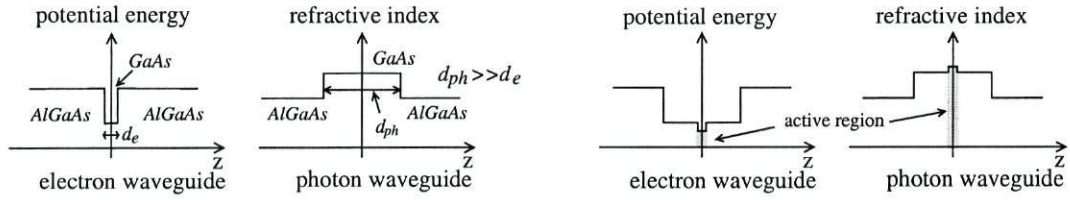


Figure 2.21: (a) Electron and photon confinement in semiconductor heterostructures [30]. The length scales of the confining structures are rather different, because the wavelength is around 1 μm for near infrared light but only about 50 nm for electrons. (b) *Separate confinement heterostructure* (SCH).

In semiconductor materials, the index can be varied by changing the number of carriers (homojunction) and/or changing the composition of the semiconductor (heterojunction), but strain and/or electrical fields can also be used to locally alter the refractive index [58]. The bandgap and the refractive index of a semiconductor are both functions of the alloy composition, making heterostructures the most widely used way to alter the properties of both electronic and optical states [22]. A smaller bandgap is usually associated with a larger refractive index, and therefore both charged particles and light waves will be confined into the GaAs region of a sandwich of AlGaAs/GaAs/AlGaAs, for example. By carefully controlling the composition, it is possible to grow different closely lattice-matched ternary and quaternary layers on GaAs and InP binary substrates [17]. Low optical loss, high quality optical waveguide structures require a good lattice match between the different layers and between the layers and the substrate, in order to avoid the formation of dislocations and strain at the interface and/or in the grown layer [15][58].

The propagation of photons is described by the Maxwell's equations [71]. In describing steady-state phenomena involving electromagnetic waves it is common to talk in terms of monochromatic waves (similarly with electron monoenergetic waves); the solution for a guided mode in a slab waveguide, Fig. 2.21(a), is obtained by matching solutions of the Maxwell's equations in the three different regions: the higher refractive index region, and the two surrounding lower refractive index regions. From the wave equations

$$\left(\nabla^2 + \omega^2\mu\epsilon\right)\vec{\mathcal{E}} = \vec{0} \tag{2.49}$$

$$\left(\nabla^2 + \omega^2\mu\epsilon\right)\vec{H} = \vec{0}, \tag{2.50}$$

and satisfying all the boundary conditions, the solutions (modes) for the electromagnetic field are obtained everywhere; here, $\vec{\mathcal{E}}$ and $\vec{H}$ are the electric and the magnetic components of the electromagnetic wave, respectively, ω represents the frequency of the electromagnetic field, μ is the magnetic permeability and ϵ is the electric permittivity of the medium. Two sets of transverse modes are allowed for this planar waveguide: *transverse electric field* (TE), and *transverse magnetic field* (TM). Transverse electric modes are polarized with the electric field

parallel to the interfaces; transverse magnetic modes have the magnetic field field along that direction.

For a planar symmetric heterostructure waveguide, such as AlGaAs/GaAs/AlGaAs, with core and cladding refactive indices n_{co} and n_{cl}, respectively, to guide ξ modes, the core thickness (d_c) must be [53]

$$d_c \simeq \frac{\xi\lambda}{2\sqrt{n_{co}^2 - n_{cl}^2}}. \tag{2.51}$$

Taking $\lambda \simeq 1\,\mu$m, $n_g = 3.590$ (GaAs) and $n_{cl} = 3.385$ ($Al_{0.3}Ga_{0.7}As$), the cut-off thickness of the second mode ($\xi = 2$) is $\simeq$0.4 μm, that is, for a waveguide core 0.4 μm thick only the fundamental mode ($\xi = 1$) is efficiently guided. The fraction of optical power in the waveguide core in the z-direction is the *optical confinement factor*, γ_{fz}, which is given by

$$\gamma_f = \frac{\int_0^{d_c} \mathcal{E}_{op}^2 dz}{\int_{-\infty}^{+\infty} \mathcal{E}_{op}^2 dz}, \tag{2.52}$$

where $\mathcal{E}_{op}$ represents the optical electric field magnitude [53].

However, the planar structure provides no transverse confinement of the light within the higher refractive index region. Just as electrons can be further confined to wires, light can be confined in semiconductor channel waveguides, which are used in many active and passive integrated optic devices, including lasers, modulators, switches and directional couplers [15]. This additional confinement is usually required to bring about desirable device characteristics, such as savings in drive power and drive voltage, and in addition compatibility with single mode fibre guides [15].[13] Waveguide channels can be made either ridged, or imbedded in the substrate, or result from strip-loading of planar guides, Fig. 2.22. The mode profiles of the guided light depend on the type of channel; Fig. 2.22 also shows schematic profiles of the fundamental mode for the three waveguide types. The optimum channel structure for a given application will depend on the light intensity profile required. For example, an optical waveguide modulator must be designed such that its mode profile presents a high overlap with the optical fibre mode profile, that is, any modulator optimisation should include good coupling efficiency (low *insertion loss*) with standard single-mode optical fibres [14], which requires the simultaneous consideration of both material and waveguide design optimisations.

In general, no exact analytic solutions for the modes of channel waveguides exists. Approximate calculations are usually employed for practical applications. The most popular methods are [59]: *Marcatili method*, *effective index method*, and *beam propagation method* (BPM). The BPM can compute the propagation of any beam in the waveguide structure; however, because BPM is a perturbation method, only structures with small index differences can be analyzed in principle.

[13] In monolithic integrated optics, all optical functions must be obtained in a single material. These functions can be broadly categorized as light generation, guiding, coupling, modulation and detection. III-V semiconductors and related compounds are the only materials to date in which all functions have been obtained [15].

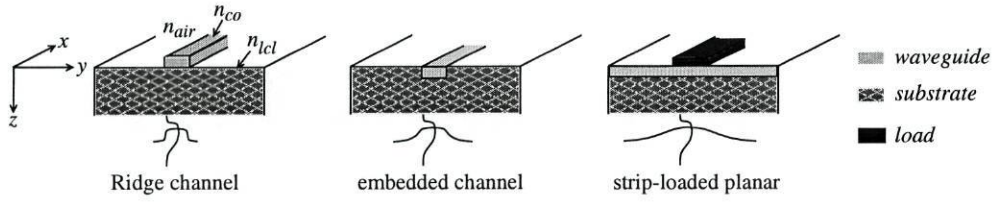


Figure 2.22: Diagrams of basic channel waveguide configurations [58].

The waveguides implemented in this work are ridged waveguides, usually made by etching away unprotected semiconductor material from planar layers by means of chemical, rf sputtering or ion-beam etching techniques [20][21]. Optical confinement (waveguiding) is achieved in the vertical direction perpendicular to the substrate plane by refractive index difference due to material compositionnal changes, and in the lateral direction due to the etched ridge. The ridge guide has air on either side of the waveguide channel, which results frequently in highly multimode guided light propagation because of the large dielectric discontinuity in the transverse directions [59]. The guided light profile typically fills the channel and has almost no field penetration into the air [13]. The waveguides used in this work were analysed using *FWave*,[14] a software based on a finite difference implementation of the vector electromagnetic wave equation solution.

In practical optical waveguides, an exponential attenuation of the optical power in the guided modes is observed, Fig. 2.23, which is characterized by the loss (extinction) coefficient, α. Loss sources in waveguides are generally attributable to three mechanisms [58][69]: *absorption*, *scattering*, and *radiation leakage*, Fig. 2.23.

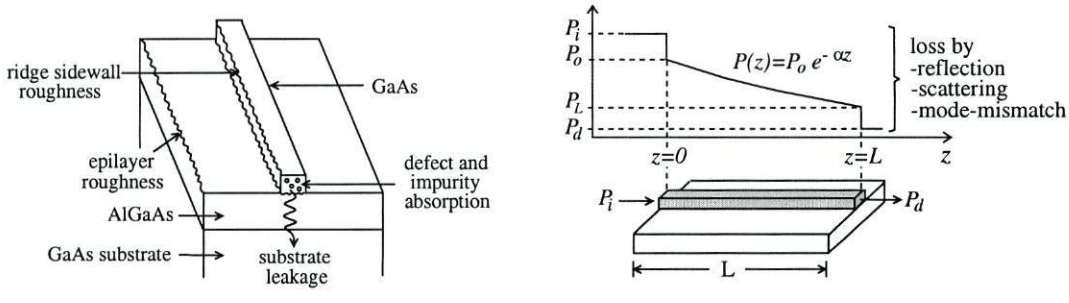


Figure 2.23: Loss mechanisms in ridge waveguides [69].

Absorption is usually the most important process in semiconductors and arises from imperfect material transparency, corresponding to photon annihilation [23]. In contrast, when photons are scattered or radiated, they change their direction of travel (as in Rayleigh and Mie scattering) and sometimes their energy (as in Brillouin and Ramam scattering), involving the transfer of optical energy from guided waves to nonguided radiation, thus constituting a loss as far as the total transmitted power is concerned [69]. Scattering arises from imperfect refractive index distributions (e.g., interface roughness), whereas leakage can occur even for ideally smooth

[14] M. R. S. Taylor, "FWave IV: a vector E-M wave solver," user manual, University of Glasgow, Glasgow, United Kingdom (URL: http://www.elec.gla.ac.uk/~michaelt/software/fwave/).

structures [59], as in the case of the AlGaAs/GaAs/AlGaAs based waveguides grown on GaAs substrates, where the core and the substrate have the same refractive index. To avoid radiation leakage from the core towards the substrate, the thickness of the substrate AlGaAs layer must be made thicker than the decay length of the waveguide fundamental mode [58][69].

The absorption loss of a given waveguide mode depends both on bulk material losses and the overlap of the modal distribution with the absorbing material [58][69]. Waveguides incorporating electrodes (metal and/or doped semiconductor), such as lasers and electro-absorption modulators, exhibit excess propagation loss due to electrode absorption above that obtained in waveguides without electrodes [69]. This loss can be minimized by reducing the overlap of the waveguide mode with the highly doped material but, in general, at the expense of device efficiency [13][15]. In devices driven by injection currents, separation of driving electrodes decreases the volume charge density achieved for a fixed injection current and hence the efficiency [13][15]. It should be noted that optical waveguide losses depend, in general, on the state of polarization of the guided light [69]. This may result simply from differences in the optical intensity distribution between polarizations, causing different overlap with scattering and absorptive material regions; polarization can also affect waveguide leakage losses [69].

Integrated optical channel waveguides, including both passive (i.e., power splitting from one to several channels) and active (such as electrical-to-optical signal conversion) types, require low propagation loss. However, photonic integrated circuits will benefit in many instances from the inclusion of an optical amplifier to compensate for coupling, splitting, and absorption losses. (In theory, use of integrated semiconductor preamps as receiver front ends can provide receiver sensitivities close to the quantum limit [14].)

2.3 Optical Properties of Semiconductors

Semiconductor optoelectronic devices depend upon the interaction of electromagnetic field with the electrons in the semiconductor, which can be described by scattering theory [12]. Photons can induce scattering within a band as well as between bands, Fig. 2.24.[15] Interband scattering, involving valence and conduction band states, is the most important mechanism for optical devices, such as lasers, modulators and photo-detectors [13][15].[16]

This section is devoted to a brief overview of the optical properties of direct bandgap III-V semiconductors, focusing on band-to-band and intraband absorption.

[15]Perturbations, such as ionised impurities and phonons, can also cause scattering of an electron from an initial Bloch state to another state.

[16]The exciton-photon interaction in semiconductor structures is also of great technical interest for high speed modulation and optical switching [13][19].

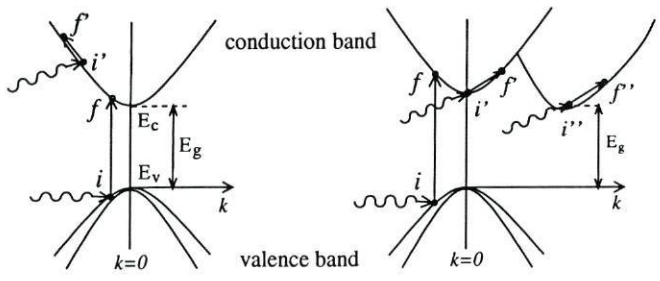


Figure 2.24: Electron scattering induced by an electromagnetic field [53].

2.3.1 Optical Interactions in Semiconductors

The semiconductor optical macroscopic properties, which ignore the details of the Bloch functions and bandstructure, are characterized by the electric permitivity, function $\widetilde{\epsilon_r}(\omega)$.[17] The refractive index and the extinction coefficient, which are related by the Kramers-Kronig relations, will be presented in terms of $\widetilde{\epsilon_r}(\omega)$ components, which can be determined from the microscopic properties of the semiconductor. Based on the perturbation theory [54], the physics behind devices operation, such as semiconductors optical detectors and light sources, will be briefly discussed.

Refractive Index and Extinction Coefficient [12][53]

In a macroscopic theory, the semiconductor response to an electric field, in the harmonic and linear regime, is characterized by its complex relative permitivity

$$\widetilde{\epsilon_r}(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega), \tag{2.53}$$

which describes the optical properties of the medium at all photon energies $E = \hbar\omega$ (ϵ_2 describes conduction, which is a dissipative process, whereas ϵ_1 describes polarization). The complex relative permitivity depends strongly on crystal structure [22].

The real and imaginary parts of the complex relative permitivity are related by the Kramers-Kronig relations:

$$\epsilon_1(\omega) = 1 + \frac{2}{\pi}\int_0^\infty \frac{\omega'\epsilon_2(\omega')}{\omega'^2 - \omega^2}d\omega' \tag{2.54}$$

$$\epsilon_2(\omega) = -\frac{2}{\pi}\int_0^\infty \frac{\epsilon_1(\omega')}{\omega'^2 - \omega^2}d\omega'. \tag{2.55}$$

In principle, information on ϵ_2 (ϵ_1) at all wavelengths should be included in the integrals. However, if one of the electronic transitions is dominant, the calculation of the integral is

[17] In the analysis of a semiconductor response to an optical field, the effect of the magnetic field component is usually much smaller than the effect of the electric field component, and will be ignored. It is assumed that the magnetic permeability is $\mu = \mu_0$.

markedly simplified; in particular, for photon energies in the transparency region of semiconductors or insulators, only the fundamental absorption edge has generally to be considered.

The complex refractive index, $\widetilde{n} = \sqrt{\widetilde{\epsilon_r}}$ [71], is defined as

$$\widetilde{n}(\omega) = n(\omega) + i\kappa(\omega) = [\epsilon_1(\omega) + i\epsilon_2(\omega)]^{\frac{1}{2}}, \tag{2.56}$$

where $n(\omega)$ is the real refractive index, and $\kappa(\omega)$ is the extinction coefficient, also known as the attenuation index. The extinction coefficient and the refractive index in the vicinity of the fundamental absorption edge depend strongly on crystal quality. The functions $n(\omega)$ and $\kappa(\omega)$ are known as the *dispersion* and *absorption curves* of the material, and are related to $\widetilde{\epsilon_r}$ by

$$\epsilon_1 = n^2 - \kappa^2 \tag{2.57}$$

$$\epsilon_2 = 2n\kappa. \tag{2.58}$$

In a lossy medium, where $\overrightarrow{J} = \sigma\overrightarrow{\mathcal{E}}$ (σ is the electrical conductivity), the wave equation for the electric field is:

$$\nabla^2\overrightarrow{\mathcal{E}} = \epsilon_1\epsilon_0\mu_0\frac{\partial^2\overrightarrow{\mathcal{E}}}{\partial t^2} + \sigma\mu_0\frac{\partial\overrightarrow{\mathcal{E}}}{\partial t}. \tag{2.59}$$

The general solution is of the form:

$$\overrightarrow{\mathcal{E}} = \overrightarrow{\mathcal{E}_0}\exp\left\{i\left(\overrightarrow{k_{op}}\cdot\overrightarrow{r} - \omega t\right)\right\}. \tag{2.60}$$

In terms of the real refractive index, $n(\omega)$, and the absorption coefficient, $\alpha(\omega)$, the electric field wave becomes (for propagation in the $\overrightarrow{r} = (0, 0, +z)$ direction):

$$\overrightarrow{\mathcal{E}} = \overrightarrow{\mathcal{E}_0}\exp\left\{i\omega\left[\frac{n(\omega)}{c}z - t\right]\right\}\exp\left(-\frac{2\pi\kappa(\omega)}{\lambda}z\right), \tag{2.61}$$

where λ is the vacuum wavelength of the light. The velocity of the wave is $c/n(\omega)$ and its amplitude is damped exponentially by the factor $\exp\left(-\frac{2\pi\kappa(\omega)}{\lambda}z\right)$.

In terms of macroscopic quantities, the optical absorption coefficient, $\alpha(\omega)$, in a medium, defined as $\alpha = \frac{\text{number of photons absorbed per unit volume per second}}{\text{number of photons injected per unit area per second}}$ (the fraction of photons absorbed per unit distance),[18] is:

$$\alpha(\omega) = \frac{4\pi}{\lambda}\kappa(\omega). \tag{2.62}$$

The reflection coefficient, $R(\omega)$, of a wave impinging on a medium characterized by real refractive index, $n(\omega)$, and extinction coefficient, $\alpha(\omega)$, is given by:

$$R(\omega) = \frac{[n(\omega) - 1]^2 + \kappa(\omega)^2}{[n(\omega) + 1]^2 + \kappa(\omega)^2}. \tag{2.63}$$

[18]The injected number of photons per unit area per second is the optical intensity $S = \hbar\omega c/n_r V$ (Wcm^{-2}) divided by the energy of photons $\hbar\omega$, where V represents the absorbing medium volume and c/n_r the light velocity in the medium.

Measurement of the reflection and absorption coefficients gives the optical constants $n(\omega)$ and $\kappa(\omega)$ of the material.

Electron-Photon Interaction [12][53]

Let the electromagnetic field applied to the semiconductor medium be represented by the vector potential $\vec{A}(\vec{r};t) = A_\omega \cos(\vec{q}\cdot\vec{r} - \omega t)\,\vec{u}$, where the unit vector $\vec{u}$ gives the polarization of the associated electric field $\vec{\mathcal{E}} = -\partial\vec{A}/\partial t$. Assuming weak optical signals ($\left|\vec{A}\right|$ is small), the Schrödinger equation (Eq. 2.1) takes the form

$$\left[\frac{\hbar}{2m_0}\nabla^2 + U_0 - \frac{e}{m_0}\vec{A}\cdot\hat{\vec{p}}\right]\Psi = i\hbar\frac{\partial\Psi}{\partial t}, \tag{2.64}$$

where $\hat{\vec{p}} \equiv (\hat{p_x},\hat{p_y},\hat{p_z}) = -i\hbar\nabla$ represents the momentum operator, and $\frac{e}{m_0}\vec{A}\cdot\hat{\vec{p}}$ is the perturbation term due to the electromagnetic field.[19] Expanding the perturbation term, gives

$$\frac{e}{m_0}\vec{A}\cdot\hat{\vec{p}} = \frac{eA_\omega}{2m_0}\left[e^{i(\vec{q}\cdot\vec{r}-\omega t)} + e^{-i(\vec{q}\cdot\vec{r}-\omega t)}\right]\left(\vec{u}\cdot\hat{\vec{p}}\right). \tag{2.65}$$

The first term represents the absorption of a photon by the electron, increasing the electron energy by $\hbar\omega$ and its momentum by $\hbar\vec{q}$; similarly, the second term describes the emission of a photon so that the electron loses both energy and momentum.[20] The photon momentum $\hbar|\vec{q}|$, for most energies of interest in solid state devices ($\hbar\omega \sim$0.1-2 eV; $|\vec{q}| \sim 10^6$ m^{-1}), is very small compared to the electron momentum in the crystal ($\left|\vec{k}\right| \approx \pi/a \sim 10^{10}m^{-1}$; $a \sim 0.5$ nm), and can be usually neglected when considering the electron-photon interaction. As a consequence, $\vec{q} = \vec{0}$ can be set in the equations and the electric field can be treated as constant in the electronic states region: *electric-dipole approximation.*

The net electron transition rate per unit volume from a state i (E_i) to another state f (E_f) due to scattering of an electron by the electromagnetic field is given by the *Fermi's golden rule*

$$W_{i-f} = \frac{2}{V}\frac{2\pi}{\hbar}\left(\frac{eA_\omega}{2m_0}\right)^2\sum_{i,f}|\vec{u}\cdot\vec{p_{fi}}|^2 \times [f(E_i) - f(E_f)]\,\delta(E_f - E_i \mp \hbar\omega), \tag{2.66}$$

with

$$|\vec{u}\cdot\vec{p_{fi}}| = \left|\left\langle\psi_f\left|e^{i\vec{q}\cdot\vec{r}}\left(\vec{u}\cdot\hat{\vec{p}}\right)\right|\psi_i\right\rangle\right| \tag{2.67}$$

(where $\mp$ applies to absorption and emission, respectively, and V represents the medium volume; the polarization dependence of the absorption coefficient is contained in the $(\vec{u}\cdot\vec{p_{if}})$ term).

In terms of the electron transition rate per unit of volume, Eq. 2.66, the steady state optical

[19] Even for an optical beam carrying 1 MW/cm^2 , the value of eA_ω/p is small enough so that perturbation theory can be used, and only the first order term in $\vec{A}$ will be considered (where $\vec{p}$ represents the electron momentum); in some cases, specially for forbidden transitions, the higher order terms containing $\vec{A}$ need to be included [12].

[20] In the quantum theory of radiation, the electromagnetic field is expressed in terms of creation and destruction operators [54].

absorption coefficient, α, in a medium (the fraction of photons absorbed per unit distance), becomes:

$$\alpha(\hbar\omega) = \frac{W_{i-f}}{(S/\hbar\omega)} = \frac{\pi e^2}{n_r c\epsilon_0 m_0^2}\frac{1}{\omega}\sum_{i,f}|\vec{u}\cdot\vec{p_{fi}}|^2 \times [f(E_i) - f(E_f)]\,\delta(E_f - E_i \mp \hbar\omega), \quad (2.68)$$

where $S = \hbar\omega c/n_r V$ is the optical intensity (the injected number of photons per unit area per second).

The element of the momentum matrix, in the dipole approximation, is given by

$$\vec{p_{fi}} = \left\langle \psi_f | \hat{\vec{p}} | \psi_i \right\rangle = \int \psi_f^* \hat{\vec{p}} \psi_i \, d^3\vec{r} = -i\hbar \int \psi_{c,k_f}^* \nabla \psi_{v,k_i} \, d^3\vec{r}. \quad (2.69)$$

Typical optical excitation is performed with visible or near-visible light wavelengths, which are much longer than the unit cell dimension in a semiconductor crystal, or the period of a superlattice, which set the scale of the electronic wave functions. Carrying out the differentiation (Eq. 2.69), the momentum matrix element in the dipole approximation is

$$\vec{p_{cv}} = \left\langle \psi_{c,k_f} | \hat{\vec{p}} | \psi_{v,k_i} \right\rangle = \hbar\vec{k_i} \int \psi_{c,k_f}^* \psi_{v,k_i} \, d^3\vec{r} - i\hbar \int u_{c,k_f}^* (\nabla u_{v,k_i}) \, e^{i\left(\vec{k_i} - \vec{k_f}\right)\cdot\vec{r}} \, d^3\vec{r}, \quad (2.70)$$

where u_{c,k_f} and u_{v,k_i} represent the conduction and valence band periodic parts of the Bloch functions.

Taking into account the small wavevector of the photon, $\vec{k_i} \simeq \vec{k_f}$, the left-hand side term of the momentum matrix element, $\int u_{c,k_f}^* (\nabla u_{v,k_i}) \, d^3\vec{r}$, is much greater than the right-hand side term, $\hbar\vec{k_i} \int \psi_{k_f}^* \psi_{k_i} d^3\vec{r}$, because this reduces to the integration of the product of orthogonal Bloch states, which is zero (or very small). For near bandedge transitions it will be assumed that $u_{c,k_i \approx 0}$ and $u_{v,k_f \approx 0}$ are given by their zone centre values, which are s-type for the conduction band (the s-type state $u_{c,0} = |s\rangle$ is a spherically symmetric state), and a combination of p-type states for the valence band ($u_{v,0} = |p_x p_y p_z\rangle$, p-states are anisotropic in space due to non-zero orbital angular momentum). From symmetry, only matrix elements of the form $\langle p_x|\widehat{p_x}|s\rangle = \langle p_y|\widehat{p_y}|s\rangle = \langle p_z|\widehat{p_z}|s\rangle = \vec{p_{cv}}$ are non-zero.

2.3.2 Optical Absorption and Radiative Recombination

Photon absorption and electron-hole radiative recombination in bulk and quantum well semiconductor structures are now considered.

Optical Absorption [12][53]

The value of the optical absorption coefficient, $\alpha(\hbar\omega)$, in a semiconductor can be considered to result from three separate processes, which are shown in Figure 2.25: direct (indirect) transitions from the valence band to the conduction band (interband transitions), with a contribution $\alpha_{vc}(\hbar\omega)$; intraconduction band absorption $\alpha_{cc}(\hbar\omega)$; and free-carrier absorption $\alpha_{fc}(\hbar\omega)$. The

α_{vc} and α_{cc} terms are relevant when $\hbar\omega \geq E_g$ and $\hbar\omega \geq \Delta E_{cc}$, respectively; the free-carrier absorption α_{fc} becomes significant at longer wavelengths.

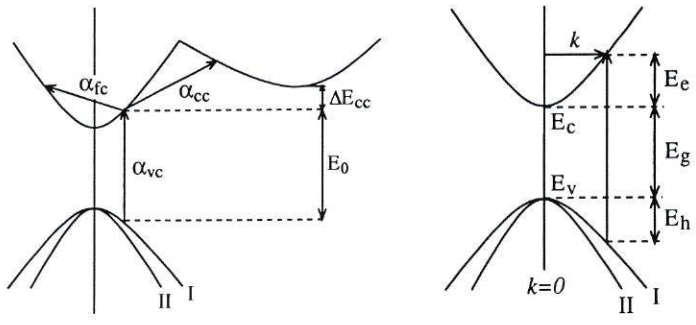


Figure 2.25: Optical absorption mechanisms [53].

Optical interband transitions in direct bandgap materials can occur satisfying both momentum and energy conservation laws, contrary to intraband transitions; however, it is possible to have strong intraband transitions in low dimensional semiconductor structures, which become quite exciting for long wavelength optical oplications, such as in quantum cascade lasers [28][29]. In bulk semiconductors, the intraband transitions must involve a phonon or some other scattering mechanism (ionized impurites, defects, etc.) to ensure momentum conservation. The free carrier absorption, as the absorption due to intraband transitions in bulk semiconductors is called, is particularly important for lasers and modulators, since it is responsible for losses in the guiding and cladding layers [18][13].

Interband Absorption in Direct Gap Semiconductors [50]

The selection rules for band-to-band transitions in direct gap semiconductors state that in the process of scattering of an electron from the valence band to the conduction band, energy and momentum have to be conserved. Thus, in the first order perturbation theory, electronic transitions due to photons are *vertical* in the $E - \overrightarrow{k}$ description. The hole energy in the valence band, the bandgap and the energy of the available state with the same $\overrightarrow{k}$ as the hole, determine the photon energy, Fig. 2.25:

$$\hbar\omega = E_c - E_v + \frac{\hbar^2 k^2}{2}\left(\frac{1}{m_e} + \frac{1}{m_h}\right) = E_g + \frac{\hbar^2 k^2}{2m_r}, \tag{2.71}$$

where m_r represents the electron-hole reduced effective mass, and $E_g = E_c - E_v$. The absorption coefficient is then given by

$$\alpha(\hbar\omega) = \frac{\pi e^2}{n_r c \epsilon_0 m_0^2}\frac{1}{\omega}\left|\overrightarrow{u}\cdot\overrightarrow{p_{cv}}\right|^2 N_{cv}(\hbar\omega), \tag{2.72}$$

with $N_{cv}(\hbar\omega)$ being the joint density of states, which is $N_{cv}(\hbar\omega) = \frac{\sqrt{2m_r^3}}{\pi^2\hbar^3}\sqrt{\hbar\omega - E_g}$.

In a bulk semiconductor, the absorption coefficient at $\hbar\omega = E_g$ has a zero value, and initially

increases as $\sqrt{\hbar\omega - E_g}$, but it has a $1/\sqrt{\hbar\omega}$ behaviour at high energies (where the density of states is not parabolic anymore.) Because of the degeneracy of the heavy-hole and light-hole states, there is no polarization dependence of the absorption coefficient near the bandgap region. In practice the absorption edge exhibits a complex structure with an exponential absorption tail (Urbach tail) extending to photon energies below the energy bandgap, E_g (Urbach rule) [16]:

$$\alpha(\hbar\omega; 0) = B \exp\left[\sigma \frac{(\hbar\omega - \hbar\omega_g)}{k_B T}\right], \tag{2.73}$$

where B and σ are adjustable constants. Other prominent features may result as consequences of semiconductor doping: the whole absorption spectrum shifts to higher energies as the free-electron concentration increases (Moss-Burnstein shift), and a more pronounced absorption tail appears [50].

Interband Absorption in Indirect Gap Semiconductors [50]

Direct interband transitions in semiconductors between states at different $\vec{k}$-values are not possible. An indirect transition can be described by a direct transition to a short-lived virtual state in the conduction band with simultaneous absorption or emission of a phonon to scatter the electron to the final conduction band state. In a similar manner, an indirect transition can also be described as a transition from a bandedge state to a virtual state, both in the valence band, followed by a direct transition to the conduction bandedge state. Indirect transitions described in terms of these virtual states can be analyzed by second-order perturbation theory, as their probabilities are much smaller than those for direct transitions [51]. Because of the additional degree of freedom intoduced by the phonon with energy $\hbar\omega_s$ and momentum $\hbar\vec{q_s}$, transitions to many more states are possible, and indirect band-to-band absorption begins at photon energies below the bandgap.

It can be shown that the absorption coefficient in indirect allowed transitions varies as $(\hbar\omega - E_g)^2$, whereas in direct allowed transitions it varies as $\sqrt{\hbar\omega - E_g}$ (this different behaviour allows the absorption processes to be distinguished experimentally). For a direct semiconductor, the absorption coefficient is very strong once the photon energy exceeds the bandgap, whereas for indirect materials the absorption coefficient rises much more gradually. It should be noticed that only phonons were considered in the scattering process. However, other scattering mechanisms, such as alloy scattering, impurity scattering, can also cause optical absorption in semiconductors.

Interband Absorption in Quantum Wells [22][24]

The periodic part of a Bloch function in QWs is relatively unaffected by the presence of the confining potential. However, the nature of the wavefunction, at least for the low lying states confined to the well region, is dependent on the potential; the density of states has now a step-like form. The absorption rates are calculated only for the well region since the barrier material has a higher bandgap and does no participate in the optical process until the photon energies are

much higher.

The absorption coefficient for interband transitions in QWs is given by

$$\alpha(\hbar\omega) = \frac{\pi e^2 \hbar}{n_r c \epsilon_0 m_0^2} \left(\frac{1}{\hbar\omega}\right) |\vec{u} \cdot \vec{p_{cv}}|^2 \frac{N_{2D}(\hbar\omega)}{d_{z_{eff}}} \sum_{n,m} f_{nm}^{eh} H\left(\hbar\omega - E_{nm}^{eh}\right), \tag{2.74}$$

where n and m indicate conduction and valence band states, respectively, f_{nm} is close to unit for $n = m$, and about zero otherwise (f_{nm}^{eh} is related to the overlap integral of the conduction- and valence-band envelope functions), H is the Heaviside function, $N_{2D}(\hbar\omega) = \frac{m_r}{\pi\hbar^2} n$ represents the jointed density of states of the two dimensional structure, and $d_{z_{eff}}$ is the effective well width, $d_{z_{eff}} = d + \sqrt{2m(U-E)/\hbar^2}$.

The momentum matrix element $|\vec{u} \cdot \vec{p_{cv}}|^2$ is isotropic for bulk cubic semiconductors, but is polarization dependent for quantum-well semiconductors. The absorption coefficient in the quantum well structure is quite distinct from the bulk case mainly because of the density of states function. Another difference arises because of the lifting of the heavy-hole and light-hole degeneracy, which makes the absorption coefficient strongly polarization dependent.

Intra-Band Absorption in Bulk Semiconductors [50][58]

A second class of optical transitions takes place between different energy levels in the same band, either in bulk or in quantum well structures. Most of the relevant transitions occur between levels in the conduction band.

Absorption of radiation by electrons in the conduction band and by holes in the valence bands produces an absorption background below the fundamental absorption edge which increases with wavelength.[21] In this process the electric field of the incident radiation accelerates the free carriers, which, in turn, are decelerated by collisions with the lattice: *free carrier absorption* [67]. In general, free carrier absorption involves both intra- and interband transitions. Intraband free-carrier absorption, α_{fc}, can be described in terms of the classical dispersion theory. For acoustic phonon-assisted processes, the absorption coefficient is

$$\alpha_{\mathrm{fc}}(\hbar\omega) \approx \frac{e^3 \lambda^2 N_{\mathrm{C}}}{4\pi^2 c^3 m^2 \mu_{\mathrm{ac}} n_{\mathrm{r}} \epsilon_0}, \tag{2.75}$$

where N_{C} is the free carrier density, n_{r} is the refractive index, and μ_{ac} the mobility due to acoustic phonon scattering, c is the light velocity and λ is the wavelength of light in the vacuum.

The same free carrier effect that leads to an increase in the imaginary part of the dielectric function can also lead to a reduction of its real part. The effect on the refractive index has two origins: free-carrier plasma dispersion (intraband free-carrier absorption), and modification of the fundamental absorption gap (Moss-Burstein shift and/or bandgap shrinkage) [65][66]. In a semiconductor with N_{C} free carriers per unit volume, the permitivity below the bandedge is

[21] Various intraband scattering processes contribute to the background absorption, such as polar optical phonon scattering, acoustic phonon scattering, and ionised-impurity scattering [53][67].

given approximately by

$$\epsilon = \epsilon_f - \epsilon_{fc} = \epsilon_f - \frac{e^2\lambda^2 N_C}{4\pi^2 c^2 m}, \tag{2.76}$$

where ϵ_f is the permitivity of pure material, λ is the wavelength of the optical field in free space, and c is the velocity of light. Since ϵ_{fc} is small compared to ϵ_f, the refractive index is given approximately by

$$n = \sqrt{\frac{\epsilon}{\epsilon_0}} = n_f - \Delta n = n_f - \frac{e^2\lambda^2 N_C}{8\pi^2 \epsilon_0 n_f c^2 m}, \tag{2.77}$$

where n_f is the refractive index of the pure film material at λ.

In n^+-GaAs, n_f=3.5, $m = 0.067m_0$, where m_0 is the mass of a free electron, $\Delta n = 1.8 \times 10^{-21} N_C$, with N_C in cm^{-3} and λ in μm. For λ=1 μm and $N_C = 2 \times 10^{18}$ cm^{-3}, $\Delta n = 0.0036$, which is sufficient to permit guiding if the pure layer is sufficiently thick [65][66].

Intra-Subband Absorption in Quantum Wells [22][28]

The intra-subband levels in a quantum well can be tailored to any arbitrary separation by adjusting the quantum well geometry. The optical properties of superlattices and quantum wells have attracted much attention in recent years due to the large oscillator strength of optical transitions between quantum subbands (levels) [24].

A quantum well which is n-type doped, so that the ground state has a certain electron density and the excited states are unoccupied, will absorb light with a photon energy equal to the ground-excited states separation energy, with the ground state electrons being scattered to upper states. In n-type doped quantum well structures, intersubband absorption is of interest because of its applications to far-infrared photodetectors [3][64][68].

Radiative Recombination [12][53]

It has been assumed so far that the initial electron state is occupied while the final state is empty. In actual devices this is, of course, not true; when electrons and holes are present in the conduction and valence bands, respectively, they can recombine radiatively, emitting photons. Non-radiative recombination effects such as defect trapping, will not be discussed here.[22]

Bulk Semiconductors [50]

Radiative recombination by spontaneous emission of photons occurs typically at a rate of $\sim$ 1 ns^{-1}, provided an electron is present in the state $\vec{k}$ and a hole is present in the same state $\vec{k}$ in the valence band. In reality, however, the recombination rate per unit volume depends upon the occupation probabilities of the electron and hole with the same $\vec{k}$-value. Under certain conditions, injected electrons into the conduction band and holes into the valence band can

[22]The non-radiative recombination competes with radiative recombination, and can have a positive or negative impact depending upon the device [18]: in a laser the non-radiative recombination is not desirable, but it is purposely increased in pn diodes to increase speed [11].

recombine and emit more photons than can be absorbed. The diference between the emission and absorption coefficents is defined as the *gain coefficient*; for simple parabolic bands, the gain $g(\hbar\omega)$ is given by

$$g(\hbar\omega) = \frac{\pi e^2 \hbar}{n_r \epsilon_0 m_0^2 c} \frac{1}{\hbar\omega} \left| \vec{u} \cdot \vec{p_{if}} \right|^2 N_{cv}(\hbar\omega) \left\{ f_e(E_e) - [1 - f_h(E_h)] \right\}. \tag{2.78}$$

A positive value of gain (negative absorption) occurs for a particular energy when $f_e(E_e) \geq 1 - f_h(E_h) \equiv f_e(E_h)$, that is, the number of electrons in the conduction band surmounts the number of electrons in the valence band: *population inversion condition.* Laser action in semiconductors requires $f_e(E_e) \geq f_e(E_h)$ [24].

In real semiconductors, intentionally or not, there are always impurities which produce electronic levels in the bandgap. These impurity levels are due to chemical impurities or native defects such as vacancies or anti-site defects (i.e., in compound semiconductors, an atom in the wrong sublattice-Ga on an As site, for example) [15]. The bandgap levels are states in which the electron is "localized" in a finite space near the defect, unlike the Bloch states which are extended. As the "free" carriers move in the allowed bands, they can be trapped by the defects, which also allow the recombination of an electron and a hole without emitting a photon.

Quantum wells [28][29]

The observation of intersubband radiative recombination in simple superlattices and quantum wells is a rather difficult task because of the overwhelming contribution of nonradiative phonon relaxation. A typical value of the optical phonon emission time τ_0 is on the order of 1 ps, the acoustic phonon relaxation time is $\sim 3 \times 10^{-10}$ s, while the interband radiative time $\sim 10^{-7}$ s [53]. Optical phonon emission is the most important mechanism of electron relaxation in polar semiconductors.

The non radiative recombination process through phonon emission is more efficient than the radiative recombination. In spite of the inefficiency of the radiative recombination, there have been several suggestions to create a population inversion between minibands - a necessary condition for laser action. These structures, when designed for practical realization of population inversion, encounter serious difficulties related to the necessity to simultaneously provide high injection current density and quickly remove electrons from the lower ground state laser level. However, in 1994 J. Faist *et al* [42] have successfully overcome these drawbacks and demonstrated efficient intersubband laser emission at a wavelength of 4.2 μm, the *quantum cascade laser.*

2.3.3 Electro-Optic and Electro-Absorption Effects

The real part of the complex refractive index of a material determines the velocity of light in the medium; the imaginary part determines the attenuation of the light. In principle, any modification in either of these quantities can be exploited to design specific optoelectronic devices,

such as switches and logic gates [13][58]. Optical signals propagating in the medium can be modulated by changing either the material real refractive index and/or its extinction coefficient.

When the real refractive index is modified by an electric field, in a region where the photon energy is such the value of the absorption coefficient can be neglected, the effect on a light beam propagating in the material is called *electro-optic*. On the other hand, if the photon energy falls in an interval where the absorption coefficient is relevant and can be directly changed, the intensity of light emerging from the sample will be altered and the effect is known as *electro-absorption*.[23] Fast and efficient optical intensity modulation based on these effects has been demonstrated [13][15].

The electric field across the semiconductor material is normally created either by reverse biasing *p-i-n* epitaxial structures (with the *i* layer being the active medium, either a bulk-like semiconductor or a MQW layer), or by a Schottky-barrier contact located directly on top of a high purity semiconductor region, which is the optical waveguiding region where the lightwave propagates [13][15][16].

Electro-Optic Effect [13][16][71]

The action of an applied electric field $\vec{\mathcal{E}}$ on a given medium changes the material index ellipsoid, and the medium *index tensor* element K_{ij} becomes $K_{ij} + \Delta K_{ij}$.[24] In the simplest case, the change ΔK_{ij} is proportional to the electric field

$$\epsilon_0 \Delta K_{ij} = \sum_{k=1}^{3} r_{ijk} \mathcal{E}_k. \tag{2.79}$$

The linear electro-optic effect is also known as the *Pockels effect*, and the r_{ijk} coefficients are the Pockels coefficients. (Depending on the symmetry of the crystal, many of the matrix elements r_{ijk} may vanish.) Equation 2.79 can be generalized to include the *quadratic (Kerr) electro-optic effect*, which is usually much smaller than the linear effect. The Pockels and Kerr coefficients constitute third- and fourth-rank tensors, whose components depend on the light wave frequency ω due to material dispersion. The Pockels and Kerr effects do not involve any transfer of charge.

In the electro-optic effect, the applied electric field is essentially used to alter the phase of propagating signals. This effect can be exploited in an interference scheme to alter the polarization or intensity of the light [15][71].

---

[23] The electric field-induced changes (Δn in the real refractive index, also referred as electro-refraction, and $\Delta \kappa$ in the imaginary part, electro-absorption), are simultaneously present, as a consequence of the Kramers-Kronig relations, Eq. 2.55, but this effect is only appreciable near optical resonances [12][16].

[24] The index tensor $\widetilde{K}$ is defined as the inverse of the permitivity tensor $\widetilde{\epsilon}$, $\left(\epsilon^{-1}\right)_{ij}$, and is often represented as $\left(1/n^2\right)_{ij}$ [16][71].

Electro-Absorption Effects [12][16][19]

Optical absorption at frequency ω in a direct semiconductor can occur only if two carrier states can be found with separation energy $\hbar\omega$, and provided that the state functions overlap in space. As a consequence, optical absorption does not occur efficiently at photon energies $\hbar\omega < E_g$, because there are no final states available in the forbidden-gap, i.e., in the absence of any perturbation, the material is transparent below a certain photon energy ($\hbar\omega < E_g$), above which the absorption rises rapidly [50]. However, application of an electric field across a bulk semiconductor causes the bandedges to slope in space, giving rise to states in both valence and conduction bands at all energies, Fig. 2.26(a).

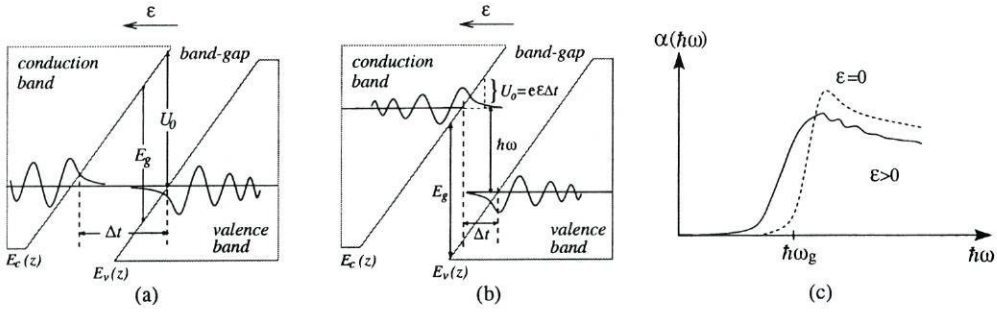


Figure 2.26: Energy band diagram under an electric field, without (a) and with (b) photon absorption. Absorption edge broadening (c) [23].

The valence and conduction band state functions under an applied electric field have evanescent tails in the forbidden gap [23]; there is a triangular potential barrier between any electron state in the valence band and the nearest available conduction band state, through which valence electrons must tunnel to reach the conduction band [23][50].

In the absence of a photon, the tunnelling distance Δt and barrier height U_0, Fig. 2.26(a), for direct transitions, are given by $E_g/e\mathcal{E}$ and $E_g + \hbar k^2/2m_r$, respectively, where m_r is the reduced electron-hole effective mass [50]. With photon absorption, the barrier height and thickness are reduced [50], Fig. 2.26(b), becoming

$$U_0 = E_g + \hbar k^2/2m_r - \hbar\omega \tag{2.80}$$

$$\Delta t = \frac{E_g + \hbar k^2/2m_r - \hbar\omega}{e\mathcal{E}}. \tag{2.81}$$

Absorption of a photon with energy below the bandgap is considerably enhanced in the presence of the electric field, with a corresponding decrease in the absorption of photons with energy higher than the bandgap; the electric field induces absorption edge broadening to lower energies [76][77], Fig. 2.26(c). In bulk semiconductors this phenomenon, which is associated with interband photon-assisted tunnelling, is known as the *Franz-Keldysh effect* [50].

In quantum wells, Fig. 2.27(a), no absorption is possible for $\hbar\omega < E_{g,well}$, and there is a continuous band of absorption for $\hbar\omega > E_{g,barrier}$. For a photon energy $E_{g,well} < \hbar\omega < E_{g,barrier}$

the absorption spectrum shows discrete lines (peaks) because the states in the well are quantized. The strongest transitions occur between states in the valence and conduction bands with the same quantum number, i.e., $n_e = n_h$ [22][79]. The lowest energy at which absorption can occur is $E_{h_1}^{e_1} = E_{e_1} - E_{h_1}$, corresponding to the energy difference between the lowest states of the wells in the conduction and in the valence bands [79]. Under an electric field perpendicular to the quantum well plane, the bands in the well slope, resulting in a decrease of the energy difference $E_{h_1}^{e_1}(\mathcal{E}) = E_{e_1}(\mathcal{E}) - E_{h_1}(\mathcal{E})$ [22][79], Fig. 2.27(b). As a consequence the absorption peak *shifts* to lower energies, and is accompanied with an absorption reduction above the edge, Fig. 2.27(c) [22][79]. This phenomenon is polarization dependent, and is known as the *quantum-confined Stark effect* (QCSE) [79][80][81].

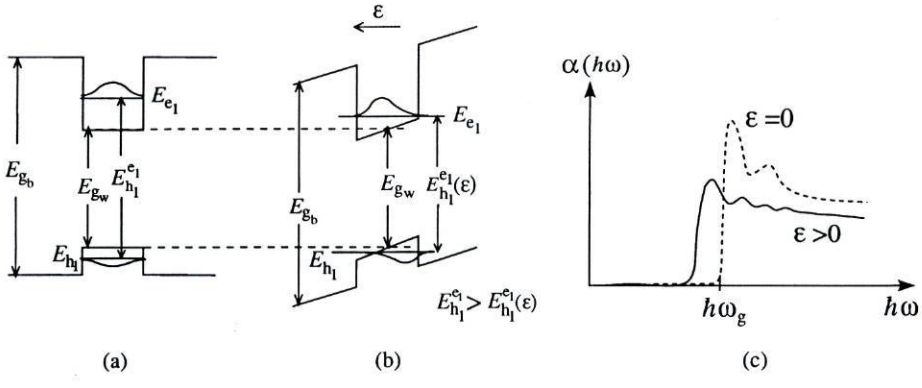


Figure 2.27: Quantum-confined Stark effect (QCSE) in semiconductor quantum wells [23].

2.3.4 The Franz-Keldysh Effect

In 1958, W. Franz and L. V. Keldysh reported theoretical studies of the effect of an electric field on the absorption edge of a semiconductor [76][12]. They predicted that absorption of a photon with energy below the bandgap is considerably enhanced in the presence of the electric field [77]. In bulk semiconductors this phenomenon, which is associated with interband photon-assisted tunnelling, is known as the *Franz-Keldysh effect* [12][50].

The absorption coefficient was derived in section **2.3.2** (Eq. 2.68 [12]) as:

$$\alpha(\hbar\omega) = \frac{\pi e^2}{n_r c \epsilon_0 m_0^2 \omega} \frac{2}{V} \sum_{i,j} \left| \left\langle \psi_f \left| e^{i\vec{q}\cdot\vec{r}} \left(\vec{u} \cdot \vec{p} \right) \right| \psi_i \right\rangle \right|^2 \left[f(E_i) - f(E_f) \right] \delta(E_f - E_i - \hbar\omega) . \tag{2.82}$$

The summation over the initial $|\psi_i\rangle$ and final $|\psi_f\rangle$ states, taking into account the Fermi occupation factor, gives the overal absorption spectrum; the delta function accounts for energy conservation and the matrix element in the formula incorporates momentum conservation automatically, with no interaction being considered between electrons and holes; V represents the interaction volume. The initial state corresponds to all electrons in the valence band, with the final state being the electron-hole pair (resulting from photon absorption).

To describe an electron-hole pair state, the two-particle wave function $\psi(\vec{r_e}, \vec{r_h})$ for an electron at position $\vec{r_e}$ and a hole at position $\vec{r_h}$ can be expressed as a linear combination of the direct product of the single (uncorrelated) electron and hole Bloch functions, $\psi_{c,k_e}(\vec{r_e})$ and $\psi_{v,-k_e}(\vec{r_e})$:

$$\psi(\vec{r_e}, \vec{r_h}) = \sum_{k_e} \sum_{k_h} A(\vec{k_e}, \vec{k_h}) \psi_{c,k_e}(\vec{r_e}) \psi_{v,-k_h}(\vec{r_h}), \tag{2.83}$$

where $A(\vec{k_e}, \vec{k_h})$ represents the amplitude function [12]. In the effective mass approximation for electron and hole pairs, the wavefunction envelope $\Phi(\vec{r_e}, \vec{r_h})$ can be written as the plane-wave expansion of the two-particle wave function [12]

$$\Phi(\vec{r_e}, \vec{r_h}) = \sum_{k_e} \sum_{k_h} A(\vec{k_e}, \vec{k_h}) \frac{e^{i\vec{k_e}\cdot\vec{r_e}}}{\sqrt{V}} \frac{e^{i\vec{k_h}\cdot\vec{r_h}}}{\sqrt{V}} \tag{2.84}$$

$$= \Phi(\vec{R}, \vec{r}) = \frac{e^{i\vec{K}\cdot\vec{R}}}{\sqrt{V}} \sum_{k} a(\vec{k}) \frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}} \tag{2.85}$$

$$= \frac{e^{i\vec{K}\cdot\vec{R}}}{\sqrt{V}} \phi(\vec{r}), \tag{2.86}$$

where $\phi(\vec{r})$ satisfies

$$\left[-\frac{\hbar^2}{2m_r}\nabla^2 + U(\vec{r})\right] \phi(\vec{r}) = \left[E - E_g - \frac{\hbar^2 K^2}{2M}\right] \phi(\vec{r}), \tag{2.87}$$

with $\vec{r} = \vec{r_e} - \vec{r_h}$, $\vec{R} = (m_e\vec{r_e} + m_h\vec{r_h})/M$, $\vec{k} = \left(m_e\vec{k_e} - m_h\vec{k_h}\right)/M$, $\vec{K} = \vec{k_e} + \vec{k_h}$, and $m_r^{-1} = m_e^{-1} + m_h^{-1}$, with $M = m_e + m_h$.

In the long wavelength (or dipole) approximation $\vec{q} \simeq \vec{0}$ (therefore, $\vec{K} = \vec{k_e} + \vec{k_h} = \vec{q} \simeq \vec{0}$, as $\vec{k} = \vec{k_e} \simeq -\vec{k_h}$, the $\vec{k}$ selection rule), the matrix element between the initial state and the final electron-hole state is given by [12]

$$\left\langle \psi_f \left| e^{i\vec{q}\cdot\vec{r}} \left(\vec{u}\cdot\vec{p}\right) \right| \psi_i \right\rangle = \sum_{\vec{k_e}} \sum_{\vec{k_h}} A^*\left(\vec{k_e}, \vec{k_h}\right) \left\langle \psi_{c,\vec{k_e}} \left| e^{i\vec{q}\cdot\vec{r}} \left(\vec{u}\cdot\vec{p}\right) \right| \psi_{v,-\vec{k_h}} \right\rangle$$

$$\cong \sum_{\vec{k_e}} \sum_{\vec{k_h}} A^*\left(\vec{k_e}, \vec{k_h}\right) \vec{u}\cdot\vec{p_{cv}}\left(\vec{k_e}\right) \delta_{\vec{k_e}+\vec{k_h}, \vec{q}} \tag{2.88}$$

$$\simeq \vec{u}\cdot\vec{p_{cv}} \sum_{\vec{k}} A^*\left(\vec{k}, -\vec{k}\right) \tag{2.89}$$

$$= \vec{u}\cdot\vec{p_{cv}} \sum_{\vec{k}} a^*\left(\vec{k}\right) \tag{2.90}$$

$$= \vec{u}\cdot\vec{p_{cv}} \sqrt{V}\, \phi^*(0), \tag{2.91}$$

where $\vec{u}\cdot\vec{p_{cv}}$ is assumed to be independent of $\vec{k}$.

The absorption coefficient is obtained substituting the matrix element into Eq. 2.82:

$$\alpha(\hbar\omega) = \frac{\pi e^2}{c\epsilon_0 n_r m_0^2 \omega} |\vec{u} \cdot \vec{p_{cv}}|^2 \, 2 \sum_n |\phi_n(0)|^2 \, \delta(E_n + E_g - \hbar\omega), \tag{2.92}$$

where n corresponds to the discrete and continuum states of $\phi(r)$ satisfying the effective mass equation in the difference coordinate system, Eq. 2.87, with $\vec{K} = \vec{0}$. The optical absorption spectrum for an interband free electron-hole transition is given by integrating Eq. 2.92 (with $\sum_n = V/(2\pi)^3 \int d^3\vec{k}$)

$$\alpha_0(\hbar\omega) = \frac{\pi e^2}{c\epsilon_0 n_r m_0^2 \omega} |\vec{u} \cdot \vec{p_{cv}}|^2 \frac{1}{2\pi^2} \left(\frac{2m_r}{\hbar^2}\right)^{3/2} \sqrt{\hbar\omega - E_g}, \tag{2.93}$$

which gives the absorption coefficient due to a free electron and hole pair [12]. The momentum-matrix element of a bulk semiconductor is:

$$|\vec{u} \cdot \vec{p_{cv}}|^2 = \frac{m_0}{6} E_p, \tag{2.94}$$

with $E_p \sim$25 eV for GaAs [53].

The interaction potential, $U(\vec{r_e}, \vec{r_h})$, corresponding to the potential energy of a free electron and a free hole in the presence of a uniform electric field $\vec{\mathcal{E}}$, takes the form[25]

$$U(\vec{r_e}, \vec{r_h}) = e\vec{\mathcal{E}} \cdot (\vec{r_e} - \vec{r_h}) = e\vec{\mathcal{E}} \cdot \vec{r}. \tag{2.95}$$

The Schrödinger equation, for the wave function $\phi(\vec{r})$ under an external uniform electric field $\vec{\mathcal{E}}$ applied along z, with $\vec{K} = \vec{0}$, becomes [12]

$$\left[-\frac{\hbar^2}{2m_r}\nabla^2 + e\vec{\mathcal{E}} \cdot \vec{r}\right] \phi(\vec{r}) = [E - E_g] \, \phi(\vec{r}). \tag{2.96}$$

A solution of the form

$$\phi(\vec{r}) = \frac{e^{i(k_x x + k_y y)}}{\sqrt{A}} \phi(z), \tag{2.97}$$

results in

$$\left[-\frac{\hbar^2}{2m_r}\frac{d^2}{dz^2} + e\mathcal{E}z\right] \phi(z) = E_z \phi(z), \tag{2.98}$$

and the total energy E is related to E_z for the z-dependent wave function as

$$E = \frac{\hbar^2}{2m_r}\left(k_x^2 + k_y^2\right) + E_z. \tag{2.99}$$

[25]The interaction potential of the form $U(r_e, r_h) = -\frac{e^2}{4\pi\epsilon|\vec{r_e} - \vec{r_h}|}$ leads to the exciton effect in optical absorption [12]. Excitonic electro-absorption in bulk semiconductors is mostly observed at low temperatures [13][19].

Equation 2.98 can be simplified by a change of variable:

$$Z = \left(\frac{2m_r e\mathcal{E}}{\hbar^2}\right)^{1/3}\left[z - \frac{E_z}{e\mathcal{E}}\right], \tag{2.100}$$

becoming

$$\frac{d^2\phi(Z)}{dZ^2} - Z\phi(Z) = 0. \tag{2.101}$$

The solutions are Airy functions $\mathrm{Ai}(Z)$ or $\mathrm{Bi}(Z)$ [82]. Since the wave function has to decay as $|z|$ approaches $+\infty$ because the potential grows as $e\mathcal{E}z$, the function $\mathrm{Ai}(Z)$ has to be chosen.[26]

In the absence of the field only certain energies are allowed; however, under the applied field the energy spectrum becomes continuous (i.e., all energies from $-\infty$ to $+\infty$ are, in principle, possible for these transitions), since the potential is not bound as $z \longrightarrow -\infty$ [23]. Therefore, the wave function satisfying the normalization condition,

$$\int_{-\infty}^{\infty} \phi_{E_{z_1}}(z)\phi_{E_{z_2}}(z)dz = \delta\left(E_{z_1} - E_{z_2}\right), \tag{2.102}$$

for a continuum spectrum, is [12]

$$\phi_{E_z}(z) = \left(\frac{2m_r}{\hbar^2}\right)^{1/3}\frac{1}{(e\mathcal{E})^{1/6}}\mathrm{Ai}\left[\left(\frac{2m_r e\mathcal{E}}{\hbar^2}\right)^{1/3}\left[z - \frac{E_z}{e\mathcal{E}}\right]\right]. \tag{2.103}$$

The optical absorption spectrum for an interband electron-hole transition under an uniform applied electric field (with $\sum_n = \sum_{k_x}\sum_{k_y}\int dE_z$) becomes [12]

$$\alpha(\hbar\omega;\mathcal{E}) = \frac{\pi e^2}{c\epsilon_0 n_r m_0^2\omega}\left|\vec{u}\cdot\vec{p_{cv}}\right|^2\frac{m_r}{\pi\hbar^2}\int_{E_z=-\infty}^{\hbar\omega-E_g}\left|\phi_{E_z}(z=0)\right|^2 dE_z \tag{2.104}$$

$$= \frac{A_0}{2\pi}\left(\frac{2m_r}{\hbar^2}\right)^{3/2}\sqrt{S_0}\left[-\eta\left|\mathrm{Ai}(\eta)\right|^2 + \left|\mathrm{Ai}'(\eta)\right|^2\right], \tag{2.105}$$

where $A_0 = \frac{\pi e^2}{c\epsilon_0 n_r m_0^2\omega}\left|\vec{u}\cdot\vec{p_{cv}}\right|^2$, $S_0 = \left(\frac{\hbar^2 e^2\mathcal{E}^2}{2m_r}\right)^{1/3}$, $\eta = \frac{E_g-\hbar\omega}{S_0}$; $\mathrm{Ai}'(\eta)$ represents the derivative of the Airy function $\mathrm{Ai}(\eta)$ with respect to η [82].

For $\mathcal{E} = 0$, equation 2.105 goes over into the familiar expression for the absorption due to direct allowed transitions (the absorption spectrum for interband free electron-hole transition given Eq. 2.93):

$$\alpha(\hbar\omega;\mathcal{E}) \to \alpha_0(\hbar\omega) = A_0\frac{1}{2\pi^2}\left(\frac{2m_r}{\hbar^2}\right)^{3/2}\sqrt{\hbar\omega - E_g}. \tag{2.106}$$

The prefactor A_0 depends on the bulk momentum matrix element, which can be determined experimentally by fitting the measured absorption spectrum with the above theoretical functions with $\mathcal{E} = 0$ and $\mathcal{E} \neq 0$ [50][77]. At photon energies below the energy bandgap the absorption

[26] $\mathrm{Ai}(t) = \int_{-\infty}^{\infty} e^{i\left(kt+\frac{k^3}{3}\right)}\frac{dk}{2\pi}$

coefficient at zero bias, $\alpha_0(\hbar\omega; 0)$, can be approximately given by Eq. 2.73 (Urbach rule):

$$\alpha_0(\hbar\omega; 0) = B \exp\left[\sigma \frac{(\hbar\omega - \hbar\omega_g)}{k_B T}\right],$$

where B and σ are determined experimentally by fitting the measured absorption spectrum with Eqs. 2.93 and 2.105 [12][13].

The wave functions, $\phi_{E_z}(z)$, tunnel into the potential for $z > E_z/e\mathcal{E}$, decaying more rapidly as the potential increases. (This corresponds to the tail of the wave function that tunnels into the classically forbidden region, where the kinetic energy is negative.) The wave functions oscillate for $z < E_z/e\mathcal{E}$, undulating more rapidly as $(z - E_z/e\mathcal{E})$ becomes more negative and the kinetic energy increases. As an electron with constant energy travels in the $+z$ direction until it hits the potential barrier at $E_z/e\mathcal{E}$, being reflected and returning along $-z$, interference between the two waves of equal intensity sets up a standing wave, Fig. 2.28(a) [23].

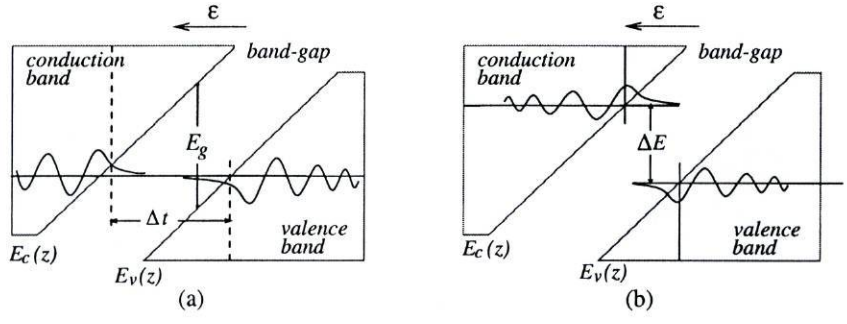


Figure 2.28: Energy band diagram in an electric field (a) without and (b) with photon absorption.

Figure 2.28(b) shows the states in the valence and conduction bands separated by $\Delta E < E_g$, but overlapping because of the tail that tunnels into the bandgap [23][50]. In the presence of photons with energy $\hbar\omega = \Delta E$, the electron tunnels into the conduction band, leading to absorption, in spite of the fact that it has only absorbed a photon with energy $\hbar\omega < E_g$ [50]. The overlap in space between valence and conduction band states depends on the difference in their energies; it is strong if $\Delta E > E_g$, when the oscillating part of the wavefunctions overlap. Only the tails overlap if $\Delta E < E_g$ and this decays rapidly with $E_g - \Delta E$, the absorption edge tail resulting from assisted tunnelling [50]. The electric field spreads out the edge of the band, which introduces absorption at frequencies below the bandedge, which is compensated by a reduction in the absorption above the bandedge.[27]

In conclusion, the Franz-Keldysh absorption spectrum shows decaying behaviour for energies below the bandgap, and oscillation phenomena above the bandgap [23][50]. At a given photon energy, $\hbar\omega$, the absorption change due to the electric field is:

$$\Delta\alpha(\hbar\omega; \mathcal{E}) = \alpha(\hbar\omega; \mathcal{E}) - \alpha_0(\hbar\omega). \tag{2.107}$$

[27] Consequence of the f-sum rule: $\sum_k f_k = N$, where f_k represent the oscillator strengths and N is the total number of electrons [23].

The absorption bandedge broadening can be interpreted as a reduction of the effective bandgap which results in an absorption bandedge "shift" to longer wavelengths. It can be shown [76][79] that the effective change in bandgap energy, ΔE_g, is given by

$$\Delta E_g \simeq -\left[\frac{(e\hbar\mathcal{E})^2}{2m_r}\right]^{\frac{1}{3}}, \tag{2.108}$$

where m_r is the electron-hole system effective reduced mass.

The resonant tunnelling diode electro-absorption modulator (RTD-EAM), consisting of an optical waveguide incorporating a double barrier quantum well resonant tunnelling diode (DBQW-RTD), employs the electric field across the DBQW-RTD depletion region to modulate, via the Franz-Keldysh effect, the transmission properties of the waveguide.

2.4 Optical Modulation

To carry information, the light waves must be modulated either in amplitude, frequency, or phase, the choice depending upon the particular application and technology available [14]. The characteristics of a light wave travelling through a medium can be modified through modulation of the complex refractive index of the medium by an external agent, such as a magnetic field, a mechanical stress or an electric field [12][16].

The most frequently used optical modulators are based on electric field induced changes in the real and imaginary parts of the complex refractive index of the material in which the optical field propagates, resulting on variations in measurable quantities (such as phase, intensity, polarization or direction of propagation) - *electro-refraction* and *electro-absorption* [12][16], respectively. Apart from electro-refraction and electro-absorption, other effects, such as acousto-optic or magneto-optic effects, can be employed for lightwave modulation [16].

Devices can be implemented either in bulk or waveguide form. The waveguide types are most relevant for photonic integrated circuits, in which the input and/or output are to be propagated in a waveguide to another device along the surface of the substrate [13][15]. In future optical transmission systems, external modulators with very high bandwidth will be needed. The most promising candidates in this respect are electro-absorption modulators, whose response time extends to the subpicosecond scale [13]. Provided that parasitics are limited by suitable mounting techniques, the bandwidth of electro-absorption modulators can be increased to values in the hundred GHz range, by reducing device capacitance [13][15][19].

The primary emphasis in this section is on a discussion of electro-absorption modulation through the Franz-Keldysh effect [12]. However, a summary of electro-optic modulators, which can also be used for the desired electro-optic transduction, is also given.

2.4.1 Modulation Schemes and Modulation Implementations

The information to be transmitted is coded into a carrier beam (an optical beam for optical communication) which is then transmitted over an appropriate medium (usually an optical fibre in optical communication). Communication systems may involve coding the information into the intensity, amplitude, frequency, or phase of the signal [14].

Two major lightwave modulation implementations are used: *direct modulation*, where light is directly modulated inside the light source [14][18]; and *external modulation*, where an external device (*modulator*) acts on the continuous light wave output from the light source [13][14].

Modulation Schemes [14][18]

The most common approach used for transmitting signals is by *intensity modulation*, in which the power of an optical source is modulated to send a signal; the signal is then directly detected by the detection module. Such modulation scheme is compatible with photo-detectors. The noise levels produced are quite high and the full bandwidth of the optical system cannot be used. This coding has a relatively high noise level, compared to coding schemes that involve frequency, phase, or amplitude modulation.

In *amplitude modulation* the field of the carrier wave is given by $\mathcal{E}_c = \mathcal{E}_{0c}\sin(\omega_c t)$, and that of the modulation signal by $\mathcal{E}_m = \mathcal{E}_{0m}\sin(\omega_m t)$. In this scheme, the maximum amplitude $\mathcal{E}_{0c}$ of the carrier wave is made proportional to the instantaneous modulation field $\mathcal{E}_m$. The modulation index is defined as $m = \mathcal{E}_{0m}/\mathcal{E}_{0c}$.[28] The amplitude modulated carrier contains three terms: i) the unmodulated carrier term; ii) an upper side band with frequency $\omega_c + \omega_m$; iii) a lower side band with frequency $\omega_c - \omega_m$.

In the *frequency modulation* approach, as the name implies, the carrier frequency is modulated. The unmodulated carrier wave is $\mathcal{E}_c = \mathcal{E}_{0c}\sin(2\pi f_c t + \phi)$. If the modulating signal is $\mathcal{E}_m = \mathcal{E}_{0m}\cos(\omega_m t)$, the frequency of the modulated signal is $f = f_c[1 + \varsigma\mathcal{E}_{0m}\cos(\omega_m t)]$, where ς is a proportionality constant which is dependent upon the details of how the modulation is done. The frequency modulated signal has a number of frequencies present, unlike amplitude modulated signals, and requires a larger spectral width for transmission; the extremes in the frequency of the modulated carrier occur at values $f = f_c[1 \mp \varsigma\mathcal{E}_{0m}]$. The higher bandwidth needed is, however, compensated by the lower noise that is possible in frequency detection systems. Closely related to the frequency modulation scheme is the *phase modulation* approach where the modulating signal varies the phase of the carrier signal.

In the above description, only modulation of analogue signals was considered; however, *digital coding* schemes are being increasingly used. In this case, the amplitude, frequency, or phase is not varied continuously, but assumes only two values corresponding to the 0 and 1 bits. The modulation schemes are then called amplitude, frequency, or phase shift keying, i.e., ASK, FSK, and PSK, respectively.

[28] It must be kept in mind that the optical carrier frequencies are in the range of $10^{14} - 10^{15}$ Hz while the modulating frequencies are $\sim 10^{10}$ Hz.

Direct and External Modulation [14][18]

Direct modulation is most frequently used with laser diodes and LEDs due to its simple implementation; the signal to be transmitted drives the light source directly, as represented schematically in Fig. 2.29. However, this has several undesirable effects, such as optical frequency chirping (due to refractive index modulation)[29] and linewidth broadening, which can introduce distortion in the transmitted signal, limiting system performance specially in long distance and high data rate operation.

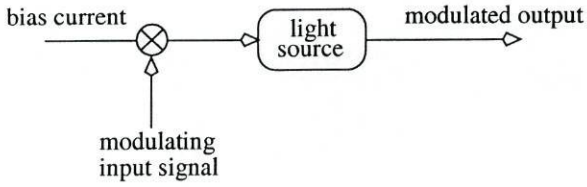


Figure 2.29: Diagramatic represention of direct modulation [14].

The overall modulation response of semiconductor lasers can be limited by extrinsic and intrinsic factors. Relevant intrinsic factors are photon life-time in the optical cavity, electron-hole recombination times, speed at which the carriers can be injected or extracted, and carrier thermalization. The main extrinsic factors are: laser heating, maximum operating power (to avoid catastrophic failure), resistance, capacitance and inductance parasitics. The restriction due to laser heating is specially important when the device is biased at high current to achieve fast response: heat produced causes deterioration of the device parameters, such as gain spectrum and threshold current. Also, there can be output delays and oscillations when the input has large and rapid variations. Fundamental physical processes limit the small signal response of semiconductor optical sources to less than 50 GHz.

Some of these effects can be reduced or eliminated using external modulation, Fig. 2.30. External light modulation and switching were demonstrated in the early 1950s using potassium dihydrogen phosphate (KDP) and $NH_4H_2PO_4$ (ADP) crystals [16]. Since then, other crystals have been employed, including ferro-electric perovskites, such as $LiNbO_3$ and $LiTaO_3$, as well as semiconductors, such as GaAs and InP [16].

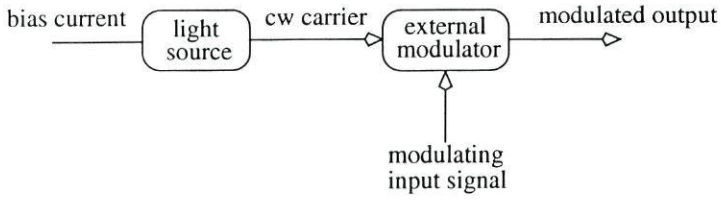


Figure 2.30: Schematic represention of external modulation [14].

The advantages of external modulation are higher attainable modulation frequencies, and

[29]The chirp parameter is defined as the ratio of the changes in the real and imaginary parts of the refractive index, $\Delta n/\Delta\kappa$ [14].

possibility of control over phase content [12][13]. A disadvantage compared to direct modulation is the need to couple the lightwave in and out, which gives rise to loss and packaging difficulties, making this approach more costly [15]. For high speed operation, specially designed contacts have to be used to provide electrical bias as well as the electrical data signal, in order to efficiently encode the optical output [13][15].

2.4.2 Electro-Refraction Modulators

Electro-refraction provides a physical method to modify in a controllable way the phase, direction or polarization of a lightwave [12][14]. Electro-refraction modulators are based mostly on the linear electro-optic effect [13][15]. For GaAs, the phase change induced by an applied electric field of magnitude $\mathcal{E}$ is given by

$$\varphi = \frac{2\pi}{\lambda} L\, n_r^3 \nabla \mathcal{E}, \tag{2.109}$$

where λ is the wavelength of the light, L is the electric field - crystal interaction length, n_r is the GaAs refactive index, and r is the electro-optic coefficient ($\sim 1.2 \times 10^{-12}$ m/V for GaAs) [12].

Intensity modulators can be implemented using electro-refraction through, for example, Mach-Zehnder interferometers and directional coupler switches, Fig. 2.31(a) and (b), respectively, with relatively large effective optical bandwidth [10][13].

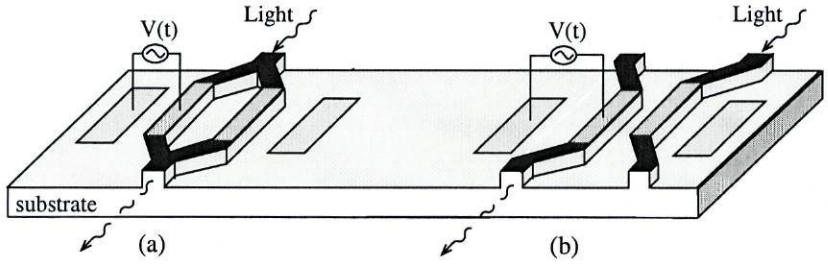


Figure 2.31: (a) Mach-Zehnder modulator. (b) Directional coupler electro-optic switch.

The electrical bandwidth of semiconductor Mach-Zehnder interferometers is generally limited by the RC time constant of the reverse-biased p-i-n structures employed, and bandwidths of ~5-10 GHz tend to result for active lengths of ~1 mm. $LiNbO_3$ based modulators show modulation frequencies up to 60 GHz [70][72]. With travelling-wave designs [10], in which the electrical signal is propagated on a microstrip line along with the optical wave, bandwidths in excess of 100 GHz are possible [70][72][73]. Refractive index modulation is desirable because of the reduced power absorption and associated heating effects. However, due to the small values of the electro-optic coefficients ($\sim 10^{-12}$ m/V), the interaction lengths are long (millimetres or more) and hardly compatible with modern microelectronics [70][72][73]. The need to construct an interferometer or directional coupler switch to provide intensity modulation requires careful design and fabrication control [10].

2.4.3 Electro-Absorption Modulators

There is an intense search for materials and devices with improved physical properties capable of providing more efficient intensity modulation [12][13][16]. Electro-absorption based semiconductor waveguide modulators are becoming widely used for high frequency applications due to their compatibility with the electronics technology and their small size [13][15].

The use of the Franz-Keldysh and the quantum-confined Stark effects for modulation is straightforward, where the usual Schottky barrier or *p-i-n* diode on a heterostructure waveguide are often employed for application of the electric field [58]. Electro-absorption modulators can be designed using either a waveguide configuration, Fig. 2.32(a), or a transverse transmission (or reflection) configuration, Fig. 2.32(b). For operation close to the absorption-edge, absorption can be electrically augmented by several orders of magnitude, typically from background values of several tens cm^{-1} to several thousands cm^{-1}.

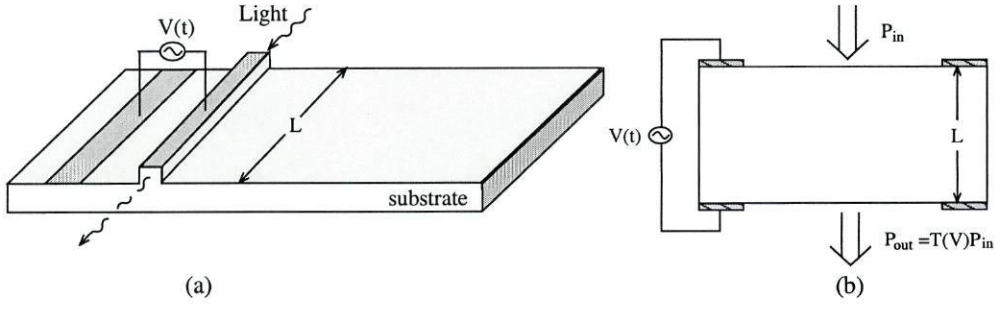


Figure 2.32: Electro-absorption modulator types. a) Waveguide modulator. b) Transverse transmission modulator [58].

Electro-absorption modulators draw photo-current and thus can present a heat dissipation problem (absent in electro-optic modulators) [15]; however, by monitoring the current of the diode when biased into absorption, electro-absorption modulators may also be operated as photodetector waveguides [58].

Electro-Absorption Modulator Characteristics [13][14]

There are five important characteristics to be examined in intensity modulators: on/off ratio, applied voltage for required on/off ratio, 3-dB bandwidth, insertion loss, and chirp. A modulator should provide large modulation bandwidth (or fast response), large modulation depth (large dynamic range of the modulated quantity), high linearity (in the case of analogue communications), small size, low insertion and propagation loss, and low electric power consumption.

On/Off Ratio

A modulator electro-absorption spectrum is represented schematically in Fig. 2.33(a), at two different bias voltages, where $\hbar\omega_0$ is the photon energy of the incident light, which is assumed to

be fixed and is supposedly chosen near the absorption bandedge; $\alpha_a(V_{off})$ and $\alpha_a(V_{on})$ represent the absorption coefficient at $V(t) = V_{off}$ and V_{on}, respectively.

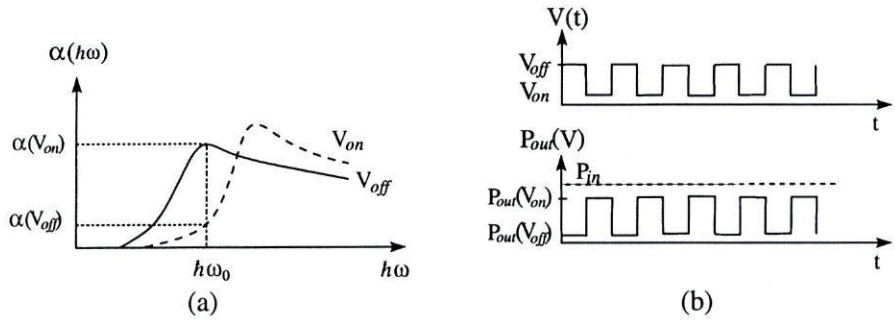


Figure 2.33: a) Absorption spectrum. b) Applied bias and transmitted optical power.

The on/off ratio is defined as the ratio of transmitted light intensity at $V(t) = V_{on}$, $P_{out}\,(V = V_{on})$, to transmitted light intensity when $V(t) = V_{off}$, $P_{out}\,(V = V_{off})$, [Fig. 2.33(b)]:

$$R_{on/off} = \frac{P_{out}\,(V = V_{on})}{P_{out}\,(V = V_{off})} = \frac{T(V_{on})}{T(V_{off})}. \tag{2.110}$$

When a voltage $V = V(t)$ is applied, Fig. 2.33(a), the transmission coefficient is

$$T(V) \propto \exp\left[-\alpha_t(V)L\right], \tag{2.111}$$

where L is the total active length of the guide the waveguide modulator; $\alpha_t(V)$ represents the total loss coefficient in the waveguide, given by $\alpha_t = \gamma_f \alpha_a + \alpha_s$, with α_a being the absorption coefficient due to electro-absorption, and α_s the loss due to all other effects related to the particular waveguide structure, such as absorption by free carriers and scattering loss due to rough waveguide edges. (For the transverse transmission modulator, $\alpha(V)$ is the average absorption coefficient of the active region, and L is its total thickness). γ_f is the optical filling factor, given by the overlap integral of the (normalised) optical mode intensity profile with the absorption region, Eq. 2.52, which in this case corresponds to the fraction of the optical power in the depleted region of the waveguide. It is implicitly assumed in Eq. 2.111 that γ_f is independent of the applied bias. Electro-absorption has also been observed to be dichroic, i.e., it attenuates TM modes more than TE modes [77][80].

Using a logarithmic scale:

$$R_{on/off}(\text{dB}) = 10 \log \left[R_{on/off}\right] \tag{2.112}$$

$$= 4.343 \gamma_f \left[\alpha(V_{off}) - \alpha(V_{on})\right] L, \tag{2.113}$$

that is, the on/off ratio (in dB) is given by the product of the absorption coefficient change and the sample thickness or length for transverse and waveguide mode configurations, respectively.

Therefore, the magnitude of the extinction ratio or the on/off ratio can be made, in principle, as large as possible by increasing the cavity length. Usually, 15 or 20 dB is necessary from the system point of view.

Applied Voltage for Required On/Off Ratio

As discussed previously, the absorption coefficient change, $\Delta\alpha$, strongly depends on the applied voltage and the operating wavelength. From a practical point of view, the smaller the applied voltage, the better for design of driving electronic circuits. Up to now, 2 V peak-to-peak operation is necessary for high speed (more than 10 Gb/s) operation. In traditional electro-absorption modulators, the applied voltage is determined by the total thickness of the intrinsic/low-doped layer.

Bandwidth

The bandwidth is usually determined by the device capacitance, C, when the device is operated in the reverse bias condition. The 3-dB bandwidth is given as follows:

$$f_{3dB} = \frac{1}{\pi RC}, \tag{2.114}$$

where R is the load resistance. The device capacitance is proportional to the sample length L and width W, and is inversely proportional to the low doped region thickness d when the stray capacitance (induced from the bonding pads, etc.) is neglected. It is assumed that the transit time of the light, τ_{lig}, through the sample is much smaller than RC, otherwise the bandwidth will be limited by the the transit time τ_{lig} ($f_{3dB} \approx 1/\pi\tau_{lig} = c/\pi n_g L$).

Insertion Loss

Operation of a modulator based on the Franz-Keldysh effect implies relatively high insertion loss [13]. Insertion loss, α_{ins}, defined as [Fig. 2.33(b)]:

$$\alpha_{ins} = \frac{P_{in} - P_{out}(V_{on})}{P_{in}} = 1 - T(V_{on}) = 1 - \exp\left[-\gamma_f \alpha(V_{on})L\right], \tag{2.115}$$

at the transmission (on) state, consists of transmission loss, reflection loss, and coupling loss ().

The transmission loss results from residual absorption loss of the material, free carrier absorption, and scattering loss. The reflection loss is at most 3 dB and can be eliminated by anti-reflection coating. The coupling loss is due to the mode mismatch between the incident and the guided light. The coupling efficiency, η_{C_m}, and coupling loss (in dB), α_{C_m}, are defined as:

$$\eta_{C_m} = \frac{\text{power coupled into (out of) the } m\text{th order mode}}{\text{total power in optical beam prior to coupling}} \tag{2.116}$$

$$\alpha_{C_m} = -10 \log \eta_{C_m}. \tag{2.117}$$

The coupling efficiency to the waveguide fundamental mode ($m = 1$), η_{C_1}, can be calculated from the overlap integral of the field pattern of the incident beam and the waveguide fundamental mode profile, given by

$$\eta_{C_1} = \frac{\left[\int \mathcal{E}_{in}(z) \cdot \mathcal{E}^*_{g_1}(z)dz\right]^2}{\left[\int \mathcal{E}_{in}(z) \cdot \mathcal{E}^*_{in}(z)dz\right]\left[\int \mathcal{E}_{g_1}(z) \cdot \mathcal{E}^*_{g_1}(z)dz\right]}, \tag{2.118}$$

where $\mathcal{E}_{in}(z)$ is the amplitude distribution of the input beam, and $\mathcal{E}_{g_1}(z)$ is the amplitude distribution of the fundamental mode.

In semiconductor waveguides, the mode spot size is generally very small, specially in the direction perpendicular to the layers due to the large refractive index difference between the guide and cladding layers. III-V semiconductor optical waveguides present high insertion loss when compared with other materials, such as $LiNbO_3$ or glass, for example [58][69][70]. The large insertion loss in semiconductor waveguides is mainly due to the coupling loss, but mode spot size transformers can be used to decrease it.

The modulator insertion loss can be measured as the simple transmission loss of the device or it can be cast in the context of total insertion loss of the device inserted in a single-mode fibre system. The latter is more meaningful for most applications and points to the need to have a modulator optical mode cross section that provides a good overlap with the fibre mode profile. Any modulator optimisation should include good coupling efficiency with standard single-mode optical fibres, which requires the simultaneous consideration of both material and waveguide design optimisations. However, in advanced systems employing highly efficient optical amplifiers, the electro-absorption modulator relatively large insertion loss will not be a crucial issue.

Chirping

As discussed previously, any change of absorption coefficient in a material will be accompanied by a phase shift of the lightwave, since the refractive index is also changed; the real and imaginary parts of dielectric function constitute a Kramers-Kronig transformed pair. This causes frequency chirping in intensity modulation, and intensity fluctuation in phase modulation. The magnitude of chirping is defined as the ratio of the change of the refractive index n to the change of the extinction coefficient κ, $\Delta n/\Delta\kappa$. Chirping decreases as the operating wavelength approaches the absorption bandedge energy, while the "on" state absorption (the insertion loss) increases; therefore, some compromise is necessary.

When chirping is induced, the transmitted optical pulses are broadened along the fibre path due to dispersion, this being a limiting factor of the capacity of long-haul high bit rate optical fibre communication systems. Chirping from electro-absorption external modulators is thought to be much smaller than from directly modulated laser diodes.

Electro-Absorption Waveguide Modulator [13][15][58]

A practical implementation of a GaAs waveguide modulator based on the Franz-Keldysh effect is represented schematically in Fig. 2.34. The left hand-side shows a GaAs/AlGaAs optical waveguide structure, with the epilayer surface being a Schottky contact.

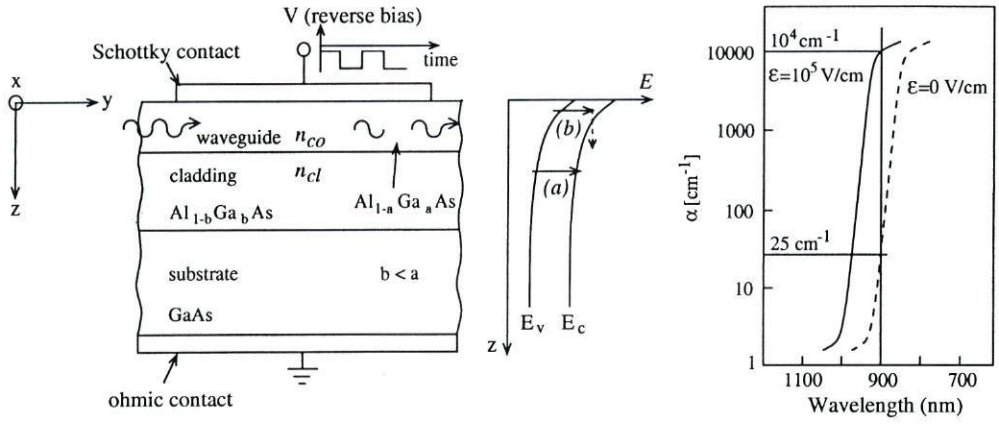


Figure 2.34: Franz-Keldysh effect in n-doped AlGaAs ($n = 3 \times 10^{16}$ cm^{-3}) [58].

Application of a reverse bias voltage to this rectifying junction causes a charge depletion layer to form, resulting in band bending extending to a depth z within the semiconductor. Outside this region the bands are flat, and a photon can be absorbed only if it has enough energy to promote an interband transition, as transition *(a)*. Close to the surface, where the bands have been bent by the electric field, a transition *(b)* can occur in which the photon energy is below the bandgap energy, due to the presence of the electric field which broadens the valence and conduction band states into the gap. Because of the steepness of the absorption edge in a direct bandgap material, such as GaAs, very large changes in absorption can be obtained; in the specific example shown in Fig. 2.34, for light of 900 nm wavelength, the absorption coeficient α increases from 25 cm^{-1} to 10^4 cm^{-1} when a field $\mathcal{E} \simeq 10^5$ V/cm is applied. The length of the modulator and the applied voltage are chosen by using absorption curves, such as those in Fig. 2.34, in order to establish a desired minimum insertion loss and a maximum modulation depth for a given wavelength. The maximum applied voltage in these devices is limited by avalanche breakdown; the maximum allowable field strength, $\mathcal{E}_{max}$, increases with increasing bandgap and doping [15]. In electro-absorption waveguide modulators based on semiconductor structures, the dopant concentration in the guiding region should, ideally, be small enough so that the depletion layer extends all the way through the guiding region.

The Franz-Keldysh and the QCSE effects produce primarily electro-absorption, and can be used to implement useful optoelectronic devices, particularly modulators and switches [19][81]. Recalling the Kramers-Kronig relations, Eq. 2.55, electro-absorption includes not only a change $\Delta\kappa$ in the optical extinction coefficient but also a change Δn in the real refractive index. At wavelengths quite close but below the bandgap, both $\Delta\kappa$ and Δn are strongly wavelength dependent. For wavelengths much longer than the bandgap, the magnitude of absorption and real

refractive index changes decreases [77]. For direct bandgap semiconductors, near the bandedge, the magnitude of the refractive index n is much larger than the magnitude of the extinction coefficients κ; however, their perturbations are of the same order of magnitude [77]. The enhanced absorption gives rise to a change in refractive index which, far below the bandgap energy, manifests itself as a quadratic electro-optic effect: the variation in Δn with applied electric field is quadratic, taking the form $\Delta n = F_{FK}\mathcal{E}^2$, where F_{FK} is a proportionality constant [77]. For GaAs at $\lambda = 1.06$ μm, $F_{FK} = 4.5 \times 10^{-19}$ m^2/V^2, and $F_{FK} = 7.9 \times 10^{-20}$ m^2/V^2 at 1550 nm.

If the medium shows gain, the structure results in an electro-absorption amplifier, where the output can actually be larger than the input in the "on" state [15]. Optical bistability has been also demonstrated employing p-i-n MQW electro-absorption structures (self-electro-optic effect devices or SEEDs) [19][79][80].

For either gain or loss electro-absorption modulators, the optical bandwidth is small because they can only be operated efficiently near the absorption edge of the semiconductor [13][15]; in both cases, the effective optical wavelength bandwidth is ~10 nm, slightly above (gain) or below (loss) the absorption edge for bulk material [15]. For quantum well based modulators, the optical bandwidth is generally even smaller, ~4 nm, although the efficiency can be correspondingly higher [13][15].

Other negative aspects of these modulators include power consumption/heat generation and high chirp, and in the case of amplifiers problems with reflections and crosstalk [14][15]. For gain or absorption modulation chirp is a major problem, and for small-signal modulation the amount of chirping tends to be similar to that from directly modulated lasers [18]. However, for large-signal modulation the wavelength chirping effect tends to be much smaller than with lasers, since relaxation ringing effects associated with the resonant laser cavity are absent [18]. Absorption generates free carriers, which produce a non-uniform field and a reduction in modulation bandwidth if allowed to accumulate, and which also induce a refractive index change [65][66].

The electrical bandwidth of reverse-biased *pn* absorption modulators can be limited by the creation of carriers that must be swept away from the active region, as in photo-detectors [15]. However, this sweep-out time still generally allows bandwidths of ~40-50 GHz, which are usually larger than the *RC*-limited bandwidth [10][13][15]. This *RC* bandwidth also tends to be a little higher than in index modulators, simply because the length of the devices tends to be shorter [13][19]. Of course, travelling-wave geometries are possible here as well [10][72][73].

2.4.4 AlGaAs and InGaAlAs Material Systems for Optical Modulation

high speed optoelectronic devices require a direct bandgap semiconductors with high electron mobility and high saturation velocity, but high thermal conductivity and high breakdown field are also welcome properties.

$Al_xGa_{1-x}As/GaAs$ System [12][17][67]

The $Al_xGa_{1-x}As/GaAs$ material system can be used to implement optoelectronic devices operating in the wavelength range 680-900 nm, such as LEDs, lasers, modulators and detectors; electro-optic modulators operating in the wavelength ranges 1300 nm and 1550 nm can also be implemented using $Al_xGa_{1-x}As/GaAs$ [15].

The electron mobility is high in $Al_xGa_{1-x}As$ for $x < 0.45$ as a result of the small effective mass associated with the Γ-valley conduction band minimum, m =$0.067m_0$ (in pure material, free of impurity scattering, the electron mobility reaches 8500 $cm^2V^{-1}s^{-1}$). The high electron mobility tends to decrease the electron transit time and increases the conductivity of undepleted n-type semiconductor regions. Hole mobility is not particularly high because of the presence of heavy holes (the effective hole mobility in pure GaAs is around 400 $cm^2V^{-1}s^{-1}$). At electric fields higher than 3kV/cm, electrons in GaAs display negative differential mobility under steady-state conditions, and their velocity decreases with increasing electric field until it saturates at around 0.6×10^7 cm/s [11]. When electrons reach high field regions, their velocity considerably exceeds this limit over short times. This velocity overshoot is of high importance in decreasing the transit time of carriers across the depletion regions in devices [11].

As a result of the wide bandgap of GaAs, the intrinsic carrier concentration is low at room temperature, $n_i = 2 \times 10^6$ cm^{-3}. Intrinsic material (frequently described as "semi-insulating", SI) has a resistivity on the order of 5×10^8 Ωcm [11]. GaAs substrates can be manufactured conveniently with resistivities approaching this limit, which helps to simplify the task of device and interconnect isolation, and considerably helps to reduce the capacitance between devices or interconnects and the substrate. This is of major significance in devices operating at microwave and millimetre-wave frequencies. GaAs substrates allow monolithic integration of a variety of electronic and optoelectronic components.

The $Al_{0.33}Ga_{0.67}As/GaAs$ alloy, MBE grown lattice matched to GaAs, was employed to form the waveguide and the DBQW RTD of the short wavelength RTD-EAM studied in this work.

$In_{0.52}(Al_xGa_{1-x})_{0.48}As/InP$ System [12][17][21][25]

For band gap energies between 0.75 eV and 1.439 eV, quaternary alloys lattice matched to InP, which combine In, Ga, Al, and As ($In_{1-x-y}Ga_xAl_yAs$) or In, Ga, As, and P ($In_{1-x-y}Ga_xAs_{1-y}P_y$), can be used in optoelectronics. Most efforts have been concentrated on devices based on InGaAsP; however, phosphorus based heterostructures show unfavourable characteristics: the conduction band discontinuity is much smaller than the valence band discontinuity, preventing a strong localization of electrons in the lower band gap material. These phosphorus based compounds are difficult to grow with conventional MBE systems due to the need to handle solid phosphorus, deal with high phosphorus concentrations at its vapour pressure and because the As/P ratio is difficult to control [57].[30]

[30]Structures incorporating InGaAsP are usually grown by MOCVD [61].

InP substrates have several shortcomings. The current small production scales make the substrate crystals significantly more expensive than GaAs. The material is somewhat weaker and may suffer higher breakage rates during processing; although breakage can be reduced by increasing wafer thickness, this only exacerbates the cost issue. Finally, InP will evaporate as a molecule (e.g., congruently) at temperatures of only 360 ^{0}C, sharply limiting the range of acceptable device processing temperatures. Semi-insulating substrates are also available, usually InP doped with Fe (which produces a deep acceptor level near midgap), with thermal conductivity higher than GaAs (0.7 versus 0.5 $\mathrm{Wcm^{-1}K^{-1}}$).

The quaternary alloy $\mathrm{In_{1-x-y}Ga_xAl_yAs}$ ($x + y \simeq 0.47$), which can be grown lattice matched to InP with bandgap energy varying from $\mathrm{In_{0.52}Al_{0.48}As}$ (1.439 eV) to $\mathrm{In_{0.53}Ga_{0.47}As}$ (0.75 eV) at 300 K, has considerable potential for fabrication of optical emitters, waveguides and detectors operating in the wavelength range 1000 - 1600 nm [61][62]. The conduction bandedge discontinuity of $\mathrm{In_{0.52}(Al_xGa_{1-x})_{0.48}As}$ varies linearly with the Al composition, $\Delta E_c(x) = 0.53x$ eV for $0 \leq x \leq 1$, being proportional to the bandgap difference, ΔE_g: $\Delta E_c(x) = 0.72\Delta E_g(x)$ eV for $0 \leq x \leq 1$. The lower ΔE_v in the InGaAlAs system, compared with the value in InGaAsP, is a major advantage for high frequency operation, since holes can easily tunnel through the barriers. Thus, under applied voltage, their sweeping out by the electric field should not limit the bandwidth of these modulators; for the same reason, the applied electric field will not be screened by hole pile-up in the wells. Therefore, modulation contrast should not decrease at high input optical intensities. In MQW designs there can be a problem with hole trapping in the active region [24]. In AlGaAs/GaAs or InGaAlAs/InP this effect can be dealt with, but for InGaAsP/InP, in which the MQW hole barriers are higher, hole trapping appears to be the dominant speed limitation [15].

Electron mobility in $\mathrm{In_{0.53}Ga_{0.47}As}$ is 11000 $\mathrm{cm^2V^{-1}s^{-1}}$, $\sim$1.6 times higher than in GaAs. The electrons effective mass associated with the Γ-valley conduction band minimum is also smaller, m =0.045m_0. Electron saturation velocity higher than 1.5×10^7 cm/s can be achieved [2][83], and the extent of transient electron velocity overshoot is also greater in InGaAs, InP, and InAlAs than in GaAs. The separation between the Γ-valley conduction band minimum and the satellite valleys is considerably higher than in the case of GaAs, $\Delta E(\Gamma - L) \approx 0.3$ eV and $\Delta E(\Gamma - X) \approx 0.48$ eV for GaAs, against 0.55 eV and 1.0 eV, respectively. The recombination velocity at surfaces of InGaAs is much smaller than with GaAs surfaces (10^3 cm/s versus 10^6 cm/s), so there is less current caused by periphery recombination and scaling to small device dimensions is easier.

Lasers and LEDs, modulators and detectors for operation at 1300 and 1550 nm, based on InGaAs/InAlAs layered structures have been implemented with high speed capabilities [13][18]. The low value of the bandgap energy in InGaAs can lead, however, to low values of the breakdown voltage and to relatively high leakage current [11].

The $\mathrm{In_{0.52}(Al_{0.11}Ga_{0.87})_{0.48}As}$ and $\mathrm{In_{0.52}Al_{0.48}As}$ alloys, MBE grown lattice matched to InP, were employed to form the core and the upper cladding of the waveguide of the long wave-

length RTD-EAM, with the InP substrate being the lower cladding; the RTD consisted of a $In_{0.52}Al_{0.48}As$/ $In_{0.53}Ga_{0.47}As$ DBQW.

2.5 Final Comments

This chapter presented and discussed essential semiconductor physics, emphasizing III-V AlGaAs-GaAs and InGaAs-InAlAs-InP compound properties: band structure, the nature of bandgap discontinuities and charge distribution in heterostructures, carrier transport mechanisms and their influence on device characteristics. The transport processes include non-linear velocity - electric field characteristics (negative differential mobility), transient effects (velocity overshoot and ballistic behaviour) and tunnelling conduction. These effects can give rise to special terminal characteristics, such as high frequency operation and negative differential conductance. The key transport effects, such as thermionic emission, tunnelling and carrier confinement, in *p-n,* metal-semiconductor junctions and sharply defined layered heterostructures resolved on a nanometer scale, were considered. The special case of resonant tunnelling in unipolar double barrier quantum well structures (the active region in the RTD-EAM devices) was described, and will be discussed in greater detail in chapter 3. The optoelectronic properties of III-V direct bandgap compounds, of quantum wells, of multiple barrier and quantum well structures were briefly analysed. Optical absorption and its application to modulation and switching were addressed, together with modulation and switching schemes. The Franz-Keldysh effect and its application to light modulation was discussed in detail.

The next chapter includes a detailed discussion of resonant tunnelling in DBQWs, and presents the RTD-EAM concept, together with an analysis of its high frequency operation. Device operation relies on a depletion region that is formed in the collector region; a small change in the bias voltage, close to the peak-to-valley transition region, can cause a large variation in the electric field distribution across the depletion region which, via the Franz-Keldysh effect, modulates the transmission characteristics of guided light with photon energy slightly smaller than that of the core bandgap.

Chapter 3

RTD Electro-Absorption Modulator

In this chapter, the resonant tunnelling diode electro-absorption modulator (RTD-EAM), consisting of a a semiconductor optical waveguide incorporating a double barrier quantum well (DBQW) resonant tunnelling diode (RTD) embedded in the waveguide core, is presented. The physics of *resonant tunnelling* through double potential barriers is discussed and analysed, aiming at its application in high speed optoelectronic converters (rf-optical and optical-rf), such as light emitters, modulators and detectors. Finally, the RTD-EAM is presented, with a discussion of its operation principle and a proposal for its application.

3.1 Double Barrier Quantum Well RTD

Classical electronic devices, such as FETs and BJTs, even in their heterojunction forms, are limited in high frequency performance by their operating mechanism [11]: the voltage controls the charge in the active region, which flows across it with a velocity that is limited by the drift-diffusion processes of semiconductors; the maximum operating frequency is limited by a combination of geometric and physical constraints.

The transit time limitation is one of the driving forces towards scaling in the direction of ever smaller device geometries. Even if the problems of lithography on increasingly finer scales continue to be solved, this progress will be limited by material constraints, such as the maximum doping densities that are allowed by solid solubilities, and quantum phenomena.

Some of the limitations on high frequency operation of devices based on planar transport can be overcome employing charge transport perpendicular to active epitaxial heterostructures [22][26]. Higher functionality is also made possible by complex association of semiconductor layers [11]. A further advantage of perpendicular transport is that the current is essentially proportional to the area of the device, so that large currents can be handled.

Double barrier quantum well resonant tunnelling diodes (DBQW-RTDs) are based on quantum tunnelling mechanisms, with the carrier transport being perpendicular to the DBQW plane,

and are excellent candidates for nanoelectronic circuit applications due to their wide-bandwidth negative differential conductance (NDC), pronounced non-linear current-voltage (I-V) characteristic, inherent high speed, structural simplicity, relative ease of fabrication, flexible design, and versatile circuit functionality [26][31][32]. Very high speed operation arises from the extremely small size of the RTD structure along the direction of carrier transport and the tunnelling process responsible for carrier flow.

One of the main characteristics of tunnelling phenomena is that the tunnelling time is not governed by the conventional transit time concept ($t = W/v$, where W is the barrier width and v is the carrier velocity), but rather by the quantum transition probability per unit of time (the corresponding tunnelling time being, in general, much smaller than the transit time through the barrier if the tunnelling conditions are fulfilled) [11].

3.1.1 Resonant Tunnelling

As referred in chapter 2, the wave function solution associated with a particle travelling towards a finite potential barrier contains two counter-propagating decaying terms in the potential barrier region, and travelling terms in regions outside the barrier. This means that a particle has *a finite probability to cross* or *tunnel* through the barrier, continueing its motion without being affected [54]. The case of two potential barriers separated only by a few nanometers (the double barrier structure) is of special interest because of the resonant tunnelling effect.

Tunnelling through a double barrier was first solved in the WKB approximation by David Bohm in 1951 [84], who pointed out that resonances in the transmission coefficient occur at certain incident electron energies. Later, in 1964, Iogansen [85] discussed the possibility of resonant transmission of an electron through double barriers formed in semiconductor crystals. In the early 1970's, Tsu, Esaki, and Chang [27] computed the two terminal current-voltage (I-V) characteristic of a finite superlattice, and predicted that resonances could be observed not only in the transmission coefficient but also in the I-V characteristic [1]. Advances in the MBE technique led to observation of negative differential conductance (NDC) at terahertz frequencies, as reported by Sollner *et al* [7], in the early 1980s. This triggered a considerable research effort to study tunnelling through multi-barrier structures.

Two important transport phenomena have emerged from the study of double-barrier structures, namely *resonant tunnelling* and *single-electron tunnelling* [2][30]. Resonant tunnelling arises from the wave nature of electrons, which gives rise to *energy quantization* in confined structures, while single-electron tunnelling arises from the *quantized charge* nature of electrons [30]. Resonant tunnelling is not observed if the distance between the barriers is long enough, so that the spacing between allowed energy levels is negligible compared to k_BT. But single-electron tunnelling can still be observed, as long as the capacitance (C) is small enough so that the electrostatic energy of a single electron (e^2/C) exceeds k_BT [30]. (This effect would be absent if the charge was not quantized, that is, if e was equal to zero.) The following discussion will concentrate only on the resonant tunnelling phenomenon.

Energy Resonances in a Double Barrier Structure [2][11]

As discussed in the previous chapter, the electronic states in a quantum-well with very thick barriers are bound states, and an electron cannot leave the well without an external excitation, because there are no equivalent energy states outside the well to where the electron can move. However, if the barriers are made thin enough, the electron remains in the well for a certain time, the *dwell time*, tunnelling then through the barriers. The structure is represented schematically in Fig. 3.1, and forms what is known as a *double barrier quantum well*, which is a close analogue of an optical Fabry-Pérot étalon [71]. The electron states are *now quasi-bound* or *resonant* states, rather than true bound states; as a consequence of the finite barriers, the energy of a resonant state is spread into a range $\hbar/\tau$, where τ (the dwell time state) is the lifetime of the carriers in that state.

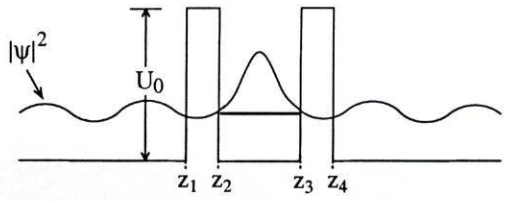


Figure 3.1: Schematic structure of a double-barrier. The electron occupation probability is also represented.

The distinctive feature of the tunneling process in the DBQW structure is that the transmission rises to much higher values than the product of the two individual barrier transmissions at energies around the values corresponding to the resonant levels: *resonant tunnelling.*[1] In a structure with identical barriers there is perfect transmission at the resonance energy, however small the transmission through the individual barriers.

Quantitative knowledge of the transmission coefficient is necessary to estimate the current-voltage characteristic of a DBQW-RTD. The basic structural parameters of a symmetric DBQW are the barriers height U_0 and width L_b, and the quantum well width, L_w. In Fig. 3.1, the four potential transitions are specified by the coordinates (z_1, z_2, z_3, z_4). Assuming that the potential function $U(\overrightarrow{r})$ in the effective mass Schrödinger equation (Eq. 2.2) can be written as the sum of a transverse (cross-sectional) confining potential $U_{XY}(x,y)$ and a longitudinal potential $U_Z(z)$, the problem can be separated into two parts because the effective mass will depend only on the z-coordinate. The transverse potential determines the transverse state energies,

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)+U_{XY}(x,y)\right]\psi_{XY}(x,y)=E_{XY}\psi_{XY}(x,y), \tag{3.1}$$

while the longitudinal potential defines the scattering problem associated with the double-barrier

[1] Resonant tunneling through single layer heterostructures has also been reported and discussed [86].

structure,

$$\left[-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\left(\frac{1}{m}\frac{\partial}{\partial z}\right)+U_Z(z)\right]\psi_Z(z)=E_Z\psi_Z(z), \tag{3.2}$$

where $E_Z = E_{XYZ} - E_{XY}$ (E_{XYZ} is the electron energy in the structure).

In a first approximation, the transverse confining potential $U_{XY}(x,y)$ does not change along the z direction in DBQW structures, and consequently it does not induce any electron scattering from one state to another. Therefore, knowledge of the potential $U_Z(z)$ is sufficient to obtain the transmission probability $T(E)$ for an incident electron (in any transverse state xy, with energy E_{XY}, with longitudinal energy $E \equiv E_Z = E_{XYZ} - E_{XY}$) [54]. The following discussion neglects all scattering processes not determined by the potential barriers.

The longitudinal effective mass Schrödinger equation in any of the five regions (I to V) represented in Fig. 3.1, assuming identical effective masses across them ($m \equiv m_i$, where subscript i indicates the ith region), can then be written as

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dz^2}+U_{Z_i}(z)\right]\psi_{Z_i}(z)=E\psi_{Z_i}(z). \tag{3.3}$$

The wave function $\psi_{Z_i}(z)$ in the ith region can be expressed as

$$\psi_{Z_i}(z) = A_i \exp(jk_i z) + B_i \exp(-jk_i z), \tag{3.4}$$

where $k_i = \sqrt{2m\left(E-U_{Z_i}\right)/\hbar^2}$. A_i and B_i are constants to be determined from the boundary conditions: the wave functions and their first derivatives (i.e., $d\psi_{Z_i}/dz$) at each potential discontinuity must be continuous [54]. The boundary conditions give A_5/A_1 as

$$\begin{aligned}\frac{A_5}{A_1} &= e^{-i\beta(2L_b+L_w)}\left\{\left[\cosh(k_bL_b)-i\frac{2E-U_0}{2\sqrt{E\left(U_0-E\right)}}\sinh(k_bL_b)\right]^2 e^{-i\beta L_w}\right.\\ &\left.+\frac{4U_0^2}{E\left(U_0-E\right)}\sinh^2(k_bL_b)\,e^{i\beta L_w}\right\}^{-1},\end{aligned} \tag{3.5}$$

and then the transmission probability through a double-barrier, $T(E)$, is

$$T_{DB}(E)=\left(\frac{A_5}{A_1}\right)^2=\frac{1}{1+U_0^2\left[\sinh^2(k_bL_b)\right]H^2/\left[4E^2\left(U_0-E\right)^2\right]}, \tag{3.6}$$

with

$$H \equiv 2\sqrt{\left[E\left(U_0-E\right)\right]}\cosh(k_bL_b)\cos(\beta L_w)-(2E-U_0)\sinh(k_bL_b)\sin(\beta L_w), \tag{3.7}$$

where U_0 is the barrier height, $k_b = \sqrt{2m(U_0-E)/\hbar^2}$ and $\beta = \sqrt{2mE/\hbar^2}$ [87]. The resonance condition occurs at incident energies $E = E_n$, so that $H(E_n) = 0$ [87], and thus $T = 1$. The electron momentum in the DBQW is quantized in the direction perpendicular to the barriers

(growth direction), giving rise to a set of n discrete energy levels.

The resonant energy levels, E_n, are solutions of the transcendental equation $H(E_n) = 0$:

$$\frac{2\sqrt{E\,(U_0 - E)}}{2E - U_0} = \tan\left(\beta L_w\right) \tanh\left(k_b L_b\right). \tag{3.8}$$

As a first-order estimate, the results of a quantum well with infinite barrier heights, Eq. 2.47, can be used:

$$E_n \approx \left[\frac{\pi^2 \hbar^2}{2 m_w L_w^2}\right] n^2. \tag{3.9}$$

For a double-barrier structure, the energy levels, E_n, will be lower, but have a similar dependence on the effective mass and well width; that is, E_n increases with decreasing m_w or L_w, and is essentially independent of L_b [88].

The transmission coefficient $T(E)$ for single and symmetrical double barrier structures is shown schematically in Fig. 3.2. The lobes of the transmission coefficient for DBQWs broaden with an increase in energy, because the barriers become more transparent.

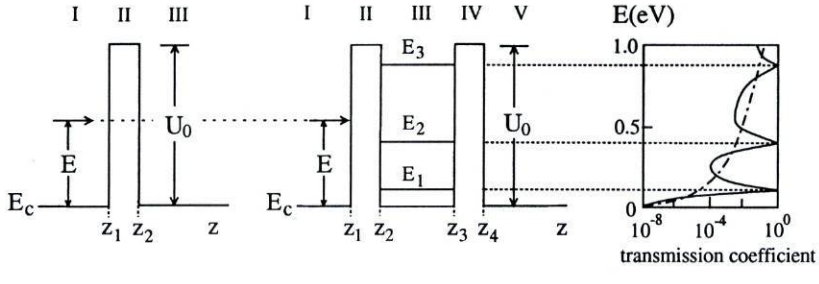


Figure 3.2: Single and DBQW transmission coefficients as functions of energy.

In the Wentzel-Kramers-Brillouin (WKB) approximation [54], the width of the nth resonance, ΔE_n (i.e., the full-width at half-maximum points, where $T = 0.5$), is

$$\Delta E_n \approx E_n \exp\left[-2L_b\sqrt{\frac{2m_b\,(U_0 - E_n)}{\hbar^2}}\right], \tag{3.10}$$

where m_b is the effective mass in the barrier[2] of width L_b. The transmission coefficient can be approximated by

$$T(E) \approx \frac{4T_{L_b}T_{R_b}}{\left(T_{L_b} + T_{R_b}\right)^2}\left[1 + \left(\frac{E - E_n}{\frac{1}{2}\Delta E_n}\right)^2\right]^{-1}, \tag{3.11}$$

where $T_{L_b}(T_{R_b})$ is the transmission coefficient through the left (right) barrier, Eq. 2.42.

The electron flow across nth resonant state is given approximately by

$$J_{RT} = e n_\perp(E_n) v_{E_n} T(E_n) \Delta E_n, \tag{3.12}$$

[2]The effective mass in the barrier is not a well defined quantity, since the conventional definition for the effective mass holds only in the allowed energy bands, not in the forbidden regions [11].

where v_{E_n} is the perpendicular component of the velocity of the electrons with incident energy $E = E_n$, and $n_\perp(E_n)$ is the incident electron distribution with energy around $E_n \pm \Delta E_n$.

The lifetime τ_n of an electron in the resonant state E_n is given approximately by $\tau_n \approx \hbar/\Delta E_n$ (Heisenberg's uncertainty principle), and is the characteristic time of DBQWs, which is important in high speed DBQW applications. Numerical calculations of the time evolution of a wave packet have been performed in order to evaluate the dwell time [89][90]. An experimental study of dwell time has also been performed by Tsuchiya *et al* [91] using time-resolved picosecond photoluminescence (PL) techniques. There is an overall agreement between these results in the regime L_b<4 nm (where the tunneling escape process dominates the radiative recombination process, to be discussed later).

3.1.2 Effect of an Applied Bias

The basic RTD configuration consists of a DBQW sandwiched between thick layers of highly doped lower bandgap material (the emitter and the collector contacts). The material forming the barriers must have a positive conduction-band offset with respect to the smaller bandgap material (type I heterostructure). Figure 3.3 shows schematically a GaAs/AlAs RTD, together with the corresponding Γ-valley conduction band profile, with the respective first resonant state, in equilibrium and under applied bias. The RTD structure may have many variations in terms of the potential energy profile, which is determined primarily by the specific material layers used.

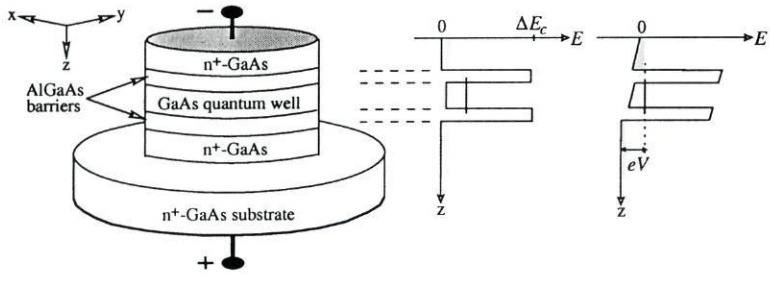


Figure 3.3: Schematic structure of a GaAs/AlAs RTD, and the conduction Γ-valley profile without bias. The band diagram drawn is a greatly simplified version of the real band structure.

A DBQW surrounded by a Fermi sea of electrons is considered, as shown in Fig. 3.4. The double barrier structure acts as a filter, whose action can be exploited applying a voltage across the structure to control the number of electrons which can take part in the conduction through resonant tunnelling. Under applied bias, the overall electron flow through a well designed double barrier structure is qualitatively different from that of a single barrier diode [23]. Peaks in electron transmission of a semiconductor double-barrier quantum well structure, a few nanometers thick, will correspond to peaks in its current-voltage characteristic [2].

Under an applied bias V, the electron energy on the positive side (collector) relative to that on the negative side (emitter) is lowered by $-eV$ (the electron energies are measured with respect to the left side contact, the emitter contact). The first resonant state is pulled down, approximately

by about $\frac{1}{2}V$ for symmetric structures, so that $E_1(V) \approx E_1(0) - \frac{1}{2}eV$ (assuming almost all the applied voltage to be dropped across the DBQW). Assuming operation at low temperature, the resonance occurs when $0 \leq E_1(V) \leq E_{F_e}$. The effect of the applied bias on the Γ-valley profile and on the first resonant level of a DBQW is shown schematically in Fig. 3.4.

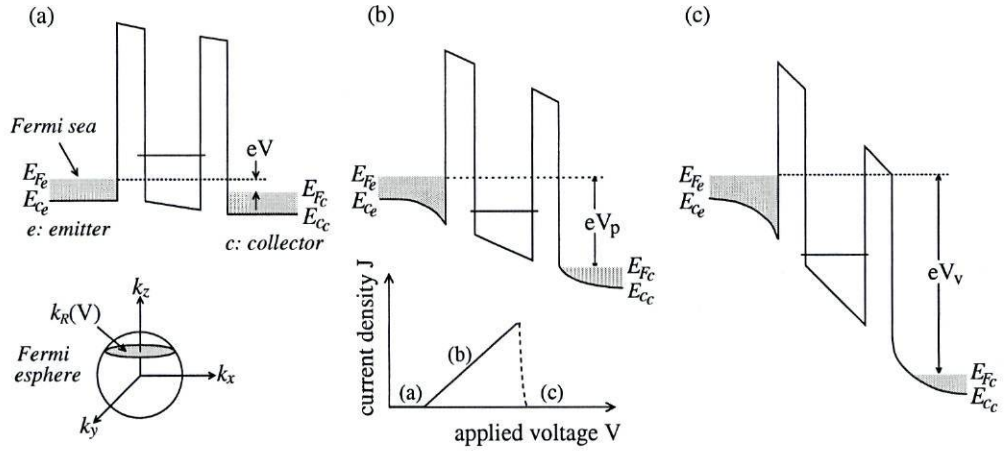


Figure 3.4: Effect of bias on the DBQW Γ-valley profile and on the first resonance [23].

In equilibrium, the distribution of electrons in the emitter and the collector is given by the Fermi-Dirac function, which is no longer valid under bias as the potential profile is perturbed due to non-uniform charge distribution acrosss the structure. Figure 3.4 shows that the bias has major effects, altering the profile and transmission properties of the barriers:

- the band profiles close to the barriers are curved (due to electron accumulation at the emitter and creation of a depletion region at the collector) - the applied voltage necessary to line up the electrons in the emitter Fermi sea with the resonance creates an asymmetry in the potential profile which significantly reduces the transmission coefficient [92][93][94];

- emitter and collector have now their own quasi-Fermi energy levels, E_{F_e} and E_{F_c}, respectively, which differ by $E_{F_e} - E_{F_c} = eV$, where V is the applied voltage.

When the applied bias is small, i.e., $V \ll V_p$, the band structure is not much affected, remaining almost flat, as indicated in Fig. 3.4(a); the first resonant energy stays well above the sea of incoming electrons, and little current flows. As voltage increases, the energy of the first resonant level is moved downwards to the Fermi sea of electrons in the emitter, Fig. 3.4(b); the current increases almost linearly with the voltage, reaching a local maximum I_p at $V_p \simeq 2E_1(0)/e$, when the overlap between the spectrum of the incident electrons and the first resonant level spectral function reaches a maximum. A further increase in the applied voltage pulls the energy of the first resonant level towards the bottom of the Γ-valley and into the forbidden gap, where there are no longer electrons available to efficiently cross the DBQW, Fig. 3.4(c); the current decreases sharply: the *negative differential conductance* region of the current-voltage characteristic. An additional increase of the bias will further lift up the electron distribution in the emitter, and tunnelling through higher resonant levels or through the top regions of the barriers will lead to a current increase.

3.1.3 Negative Differential Conductance

Negative differential conductance was first demonstrated experimentally by Chang, Esaki, and Tsu in 1974 [1], who reported oscillatory behaviour of the vertical differential conductance in a structure consisting of a thin GaAs layer between AlGaAs barriers sandwiched by heavily doped layers of GaAs grown by MBE.

The negative differential conductance in DBQW diodes can be explained from momentum and energy conservation considerations [95][31]. In the process of resonant tunnelling, electrons are incident from a 3D region (where the electrons have a plane-wave-like behaviour) into the 2D quantum well (where the electrons have a plane-wave-like motion only in the x and y directions). The electron energy in the 3D regions can be written as

$$E_{3D}\left(\overrightarrow{k}\right) = E_z(k_z) + \frac{\hbar^2}{2m}\left|\overrightarrow{k_{xy}}\right|^2, \tag{3.13}$$

where $E_z \equiv E_z(k_z) = E_c + \frac{\hbar^2}{2m}k_z^2$, E_c is the conduction bandedge energy, and $\overrightarrow{k_{xy}} = (k_x, k_y)$. The electron energy in the quantum well is

$$E_{2D}\left(\overrightarrow{k}\right) = E_n + \frac{\hbar^2}{2m}\left|\overrightarrow{k_{xy}}\right|^2, \tag{3.14}$$

where E_n is the energy of an available nth subband in the quantum well, corresponding to the zeros of Eq. 3.8 (the subband bandedge).

A reduction of dimensionality occurs, but as the tunneling electrons are not subject to any potential change in the transverse directions, their transverse momentum, $\overrightarrow{k_{xy}}$, is conserved. The electron longitudinal momentum k_z, however, generally changes with distance across the structure, except in the flat band regions. As $E_1(n = 1)$ is the first accessible (unoccupied) energy state in the quantum well, energy conservation implies that resonant tunnelling through the first energy level is possible only for electrons with a longitudinal momentum [95]

$$k_z \simeq k_R \equiv \sqrt{2m\left(E_1(V) - E_{C_e}\right)/\hbar^2}, \tag{3.15}$$

where V represents the applied voltage (Figure 3.4). As the collector-edge is pulled down under bias (relative to the emitter-edge), the resonant level samples the emitter Fermi sea and the number of electrons that can tunnel is enhanced according to the Fermi-Dirac distribution (first PDC region). When $E_1(V) \simeq E_{C_e}$, which corresponds to $k_z \simeq 0$, the number of tunnelling electrons per unit area reaches a maximum, giving high current flow. As E_{C_e} rises above $E_1(V)$, the number of electrons in the emitter that can tunnel into the quantum well, while conserving their transverse momentum, drops sharply (the NDC region). The current is now mainly due to thermionic emission, tunneling through higher resonant levels and through the barriers (the beginning of the second PDC region) [2].

3.1.4 Current-Voltage Characteristic

Ideally, the DBQW-RTD valley current will be very small and the ratio I_p/I_v will be very high [2][30]. However, at room temperature, there is always a significant background valley current due to resonant tunnelling through other conduction band valleys than the Γ valley, $\Gamma - X$ intervalley scattering, phonon scattering, impurity scattering, etc. [2]. Figure 3.5 presents the experimental I-V characteristic of a AlGaAs/GaAs DBQW-RTD at two different temperatures, showing *negative differential conductance* ($G_n = dI/dV < 0$) for a range of applied voltages.

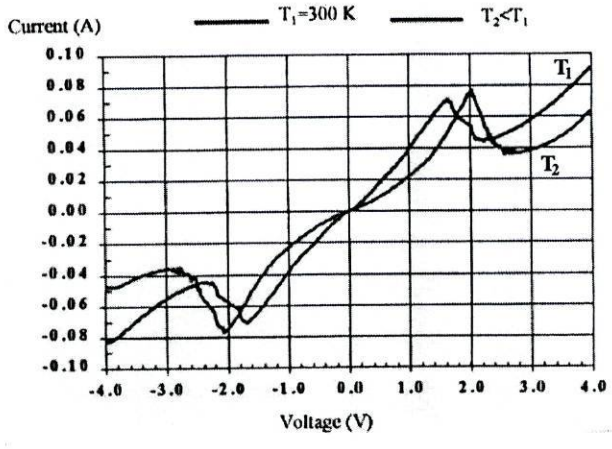


Figure 3.5: Experimental current-voltage characteristic of a GaAs/AlAs DBQW-RTD at two different temperatures.

The RTD current-voltage characteristic is anti-symmetric around the zero bias point, displaying negative differential conductance (NDC) for both polarities, because the DBQW-RTD structure is symmetric and the device is unipolar. For a given polarity of the applied bias, the I-V curve of the basic RTD can be divided into three branches. There are two *positive differential conductance* (PDC) zones, one before the peak point and one beyond the valley point, with the region of negative differential conductivity (NDC) in between.[3]

As the applied voltage across the RTD terminals is increased from zero, the current increases, mainly due to tunnelling, until it reaches a local maximum, I_p (the peak current), when the voltage reaches V_p (the peak voltage). Although there is always a background current due to different transport mechanisms, resonant tunnelling dominates the current flow in the first PDC branch of a well-designed device. Continueing to increase the voltage beyond V_p, the current drops to a local minimum, due to a reduction in resonant tunnelling, until the voltage reaches V_v (the valley voltage): the NDC branch. The current corresponding to V_v is the valley current I_v. Beyond V_v, the current starts to increase again, approximately in accordance with conventional diode behaviour.

The voltage across the RTD when it is operating in the first PDC region and the current equals I_v is called the *first voltage* V_f. Similarly, the voltage across the RTD when it is operating

[3]Dc negative differential resistance of double barrier diodes at small dc bias voltages, as a result of an applied ac voltage, has been also demonstrated [96].

in the second PDC region and the current equals I_p is called the *second voltage* V_s. A *voltage swing* (V_{I_p}) may also be defined as the voltage difference between the peak on the first PDC branch and the point at which the current reaches I_p on the second PDC branch. For a current in the interval $[I_v, I_p]$, there are two possible stable voltages: $V < V_p$ or $V > V_v$.

In most applications, a too large value of the peak current I_p is certainly undesirable from the point of view of power dissipation, specially for future large-scale circuit integration [32], because it implies exceptionally low contact resistance. In addition, it takes time to reach a large peak current point and then decrease the current to the valley point, which is undesirable for high speed operation. In general, a low valley current density is desired, being necessary for switching and advantageous for low power consumption when large-scale RTD circuits are integrated [32]. For microwave applications, a peak current density on the order of a few tens of kA/cm^2 is adequate [32].

Other important RTD I-V parameters used in this work are defined as follows [31]:

$$G_d = \frac{1}{R_d} = \left| \frac{I_v - I_p}{V_v - V_p} \right| \tag{3.16}$$

$$PVCR = \frac{I_p}{I_v} \tag{3.17}$$

$$PVVR = \frac{V_p}{V_v}. \tag{3.18}$$

Here, G_d (R_d) is the average magnitude of the negative differential conductance (resistance), PVCR is the peak-to-valley current ratio, and PVVR is the peak-to-valley voltage ratio.

From an application standpoint, the negative differential conductance can provide the gain necessary to sustain oscillation, and small intrinsic delay and RTD capacitance enable the diode to self-oscillate at high frequencies [2]. There have been reports of oscillators at frequencies never reached by other semiconductor devices [5][6][7].

3.2 Physical Mechanisms in DBQW-RTDs

The current in ideal DBQW structures would be the result of resonant tunnelling through the Γ conduction band valley. However, other transport processes take place in real structures, such as resonant tunnelling through other conduction band valleys rather than the Γ valley, phonon scattering, impurity scattering, $\Gamma - X$ intervalley scattering, and dimensional changes across the interfaces [2]. The electrons may suffer from incoherent scattering processes which, in general, degrade the PVCR through momentum and energy relaxation, allowing transitions of electrons from resonant to non-resonant states. These incoherent mechanisms of electron transport contribute mainly to the valley current, although the resonant tunnelling process must be dominant in well-designed RTD devices [2].

3.2.1 Transport Processes

A simplified picture of the major physical processes pertaining to resonant tunnelling and current transport mechanisms in DBQW-RTDs is given in Fig. 3.6; only mechanisms involving electron transport in the Γ-conduction valley, under applied bias, are represented. The resonant tunnelling process is indicated in Fig.3.6 by (a) and (b), for electrons with energies coincident with the two resonant levels represented. Under applied bias, resonant tunnelling will occur at certain bias voltage values mainly through only one resonant level, in general through the first one.

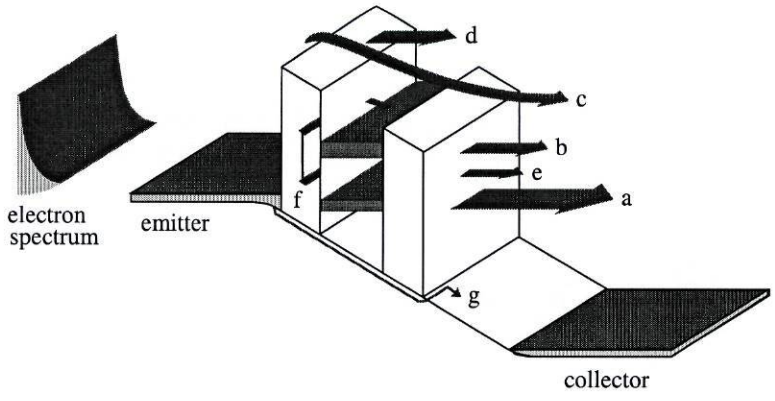


Figure 3.6: Physical processes and current transport components in DBQW-RTDs.

In addition to transport through resonant levels, other processes are pictured:

- *thermionic emission,* (c): as a result of thermal excitation, a fraction of the electrons in the tail of the Fermi-Dirac distribution will have sufficiently high energy to overcome the barriers; *field assisted tunnelling,* (d), and *tunnelling through evanescent states,* (e): incident electrons have a finite, but small, probability to tunnel through the non-resonant energy ranges that lie between the resonances; electrons with sufficiently high energies can flow non-resonantly or through higher lying quantum levels [97]. The current due these processes is the main contribution to the valley current, and increases monotonically with increasing voltage in a similar manner to that of a tunnel diode. It is strongly temperature dependent and appears to overshadow the NDC at high temperatures. To reduce this non-resonant current the barriers should be made as high as possible, and the RTD should be designed to operate at relatively low bias voltages;

- *inelastic tunnelling,* (f): elastic scattering in 3-dimensional space (such as impurity scattering) is, in general, inelastic scattering in one-dimensional space, because the energy component in each axis is not conserved. The resulting incoherent electrons redistribute themselves in the emitter and well regions and subsequently tunnel through the barriers incoherently [2];

- *leakage current,* (g): current flowing along the periphery of the device generated by a surface effect, or a component resulting from carrier multiplication across high field regions, or current due to generation-recombination effects.

In a more realistic picture, the physical processes involved in RTD operation are actually much more complex than the preceding simple description. Electrons exchange energy and momentum with lattice vibrations, impurities, interfaces, as well as in interactions among themselves

[2][98]. Electrons in the double-barrier structure may emit or absorb longitudinal optical (LO) phonons during resonant tunnelling; as this is an inelastic scattering process [99], the electrons involved do not conserve the components of momentum, neither parallel or perpendicular to the barriers, affecting the transmission probability [23]. The LO phonon-assisted resonant tunnelling is particularly pronounced in RTDs with lowly doped layers between contacts and barriers, in which the emitter states present are quantized [2]. The valley current, due to non-resonant or inelastic tunnelling current, is also composed of phonon-coupled and impurity-assisted tunnelling components [2].

The energy dissipation processes of tunnelling electrons, such as LO-phonon emission, lead to a non-equilibrium electron distribution in the quantum well. Thus a simple Fermi-Dirac distribution function does not necessarily describe the accumulated electrons in the well and a complex non-equilibrium transport theory would in general be required [2]. The non-uniform electron charge distribution across the structure will generate an electric field, $\mathcal{E}(z)$, which modifies the DBQW profile,[4] affecting the transmission coefficient T and the quantized energy levels, E_n. An exact quantitative analysis of the resonant tunnelling behaviour requires knowledge of $\mathcal{E}(z)$ as a function of the applied voltage V [93]. The modified profile will affect the total charge in the well, which in turn changes the electric field distribution. This feedback mechanism is believed to be responsible for the hysteresis and shoulder-like behaviour observed in the I-V curve of RTDs [100][101]. Determination of the charge distribution in RTDs under non-equilibrium conditions is one of the crucial issues under lively discussion, as it is closely related to the energy dissipation processes in the system [102].

Up to now only electrons in the Γ-valley were considered, without paying attention to the contributions from those in the X- and L-valleys [103][104][105]. This is reasonable as long as the minima of these upper valleys are at much higher energies than the Γ-valley. However, the X-valley forms the conduction band minimum for $Al_xGa_{1-x}As$ when $x > 0.45$, and the $\Gamma_{GaAs}-X_{AlAs}$ energy difference, $\Delta E_{C\Gamma X}$, is much lower (0.27 eV for $x = 1$) than the $\Gamma_{GaAs}-\Gamma_{AlAs}$ difference, $\Delta E_{C\Gamma\Gamma}$ (1.12 eV). Therefore, it is expected that the smaller $\Delta E_{C\Gamma X}$ barrier height will play a significant role in carrier transport [104][105].

As a consequence of these mechanisms, development of an accurate physical model of realistic RTDs becomes a formidable task.

3.2.2 Models of DBQW-RTDs

Significant effort has been devoted to development of RTD models, with varying degrees of sophistication and success [31]. Formal treatment of RTD modeling entails advanced quantum theory, possibly including rather intricate formulations of multi-band effective mass theory, quantum statistical theory based on the density matrix, Wigner and Green's functions [2][102]. Interest in achieving an accurate RTD model lies in the fundamental motivation to gain insight into the

[4]The voltage dropped across a typical DBQW (10 nm thick) is around 0.1 V, therefore a typical electric field on the order of 10^5 V/cm exists in the well and in the barriers.

quantum mechanical aspects of electron transport in mesoscopic structures, but also on the need of guidance for device design.

The RTD, as a nanoelectronic device, is an open quantum system in which electron transport is three dimensional, time dependent, time irreversible, dissipative, and many-body interactive, with both particle and energy exchanges with the device environment [102]. The device system is therefore very different from a simple isolated quantum system, where a conservative Hamiltonian and the boundary conditions for the Schrödinger equation may be readily formulated [22][54].

Any accurate RTD modeling effort requires energy-band calculations, such as the band offsets at the heterointerfaces, multiband effects, carrier effective masses, etc.. The transport formalism must be able to model the major processes in the RTD, such as those represented in Fig. 3.6, and must include an adequate treatment of the particular boundary conditions, space charge distribution effects, and the various scattering mechanisms (beside the simplifications, approximations and numerical discretization typically employed to simplify the calculations of specific device properties) [31].

Coherent Models [2][30][31]

Tsu, Esaki and Chang described the resonant tunneling current-voltage characteristic as the result of the transmission and reflection of coherent electron waves through the structure [27]. They assumed that the distribution of incoming electrons on each side is given by a Fermi-Dirac distribution with an appropriate value of the Fermi energy, E_F (as electrons are assumed in thermal equilibrium, due to interactions among them and with their surroundings), and that electrons are not reflected after crossing the double barrier, otherwise the incoming distribution would be perturbed (the contacts are taken as regions where the potentials have reached their plateaus, so the wave functions are plane waves).[5]

Assuming the transverse dimensions of the RTD structure to be much larger than the electrons de Broglie wavelength, the system can be considered translationally invariant in the xy-plane. Since the transmission probability through the structure only depends on the motion perpendicular to the interfaces, it is possible to group together the states that have the same energy for perpendicular motion but different kinetic energy for parallel motion, at a given applied potential V, as

$$S(E_z; E_F; V) = \frac{mk_BT}{\pi\hbar^2} \ln\left[1 + \exp\left(\frac{E_F - E_z - eV}{k_BT}\right)\right]. \tag{3.19}$$

This function is designated as the *supply function*, and corresponds to the density of electrons in a subband $E_z(k_z)$. The spectrum of electron eigenstates in the emitter is driven by the voltage applied across the RTD, as shown in Fig. 3.6. The current density (the current components due to electrons impinging on the DBQW from the left and due to electrons arriving from the right) is obtained by summing the current density of each state, over occupied states, multiplied by

[5]The contacts are typically heavily doped to provide low ohmic contact resistance and high current density.

their transmission probability:

$$I_{tun}(V) = \frac{e}{2\pi\hbar}\int_{E_{Cem}}^{\infty}\left[T_{left}(E_z;V)S_e\left(E_z;E_{F_e};0\right) - T_{right}(E_z;V)S_c\left(E_z;E_{F_c};V\right)\right]dE_z, \quad (3.20)$$

where E_z represents the "longitudinal component" of the total energy, $E_z = E_{c_{e,c}} + \hbar^2 k_z^2/2m$. This simple picture is called *global coherent tunnelling model* since the phase-coherence of an electron wave is mantained throughout the whole structure.

The coherent tunnelling approximation assumes that the phase-coherence of the electron waves is conserved. However, the electrons involved in resonant tunnelling stay in the DBQW for a dwell time ranging from a few tens of femtoseconds, to a few hundred nanoseconds depending on structural parameters.[6] While in the DBQW, the electrons may suffer from scattering processes which lead to momentum and energy relaxation, with breaking of phase-coherence (caused by electron-electron interactions, electron-phonon interactions, residual impurity scattering, disorder scattering, and so on). A *sequential tunnelling model* was proposed by S. Luryi [111] as an alternative explanation of the resonant tunneling process. This model describes resonant tunnelling as two continuous tunnelling processes: tunnelling from the emitter into the quantum well followed by tunnelling into the collector. Between these two processes electrons suffer from scattering processes in the DBQW and are relaxed into locally quasi-equilibrium states. Weil and Vinter [112] demonstrated theoretically the equivalence between resonant tunnelling and sequential tunnelling.

Space charge effects, which reflect the density of electrons, are due to the impurity doping profile and the mobile electronic charge that exists in the quantum well, in the emitter notch or accumulation region, and give rise to non-uniform potential profiles [108]. A simple Fermi-Dirac distribution function does not necessarily describe the accumulated electrons in the well and emitter regions. Since this charge density responds to the same electrostatic potential it generates, the potential and charge distributions have to be solved in a self-consistent form. An important assumption in such methods resides in which locations the injected carriers are assumed to be thermalized, as this plays a critical role in the quantitative results for the quantum well charge and the current density. These aspects depend on the specific device structure being modeled and the operating conditions [108].

In the framework of the global coherent tunnelling model, the electron density in the RTD can be expressed as

$$\begin{aligned} n(z) &= \frac{k_B T}{2\pi^2\hbar^2}\left\{\int_0^{\infty}|\Psi_{k_z}(z)|^2\, m\, \ln\left[1+\exp\left(\frac{E_{F_{em}}-E_z}{k_B T}\right)\right]dk_z\right. \\ &+ \left.\int_{-\infty}^{0}|\Psi_{k_z}(z)|^2\, m\, \ln\left[1+\exp\left(\frac{E_{F_{col}}-E_z}{k_B T}\right)\right]dk_z\right\}. \end{aligned} \quad (3.21)$$

The distortion of the potential due to the presence of electrons requires a self-consistent

[6] The typical mean free time of electrons in bulk GaAs material is of the order of subpicoseconds, at room temperature [17].

solution of Schrödinger's and Poisson's equation. The space charge distribution determined by Eq. 3.21 results in *the self-consistent Hartree potential*, $U_{sc}(z)$, which is obtained through the following Poisson equation:

$$\frac{d}{dz}\left[\epsilon(z)\frac{dU_{sc}(z)}{dz}\right] = -e\left[N_d^+(z) - n(z)\right]. \tag{3.22}$$

The self-consistent potential is then added to $U_Z(z)$ in the Schrödinger equation, Eq. 3.2. These equations can be solved numerically in an iterative way until self-consistent solutions for $\Psi_{k_z}(z)$ and $U_{sc}(z)$ are finally obtained [108].

The self-consistent calculations take into account effects such as band bending, changes in effective barrier height, the effective voltage drop across the barrier, and shifting of the resonant state energies in the quantum well. A consequence of bandbending is the formation of a 2-D accumulation layer in the RTD emitter region, in which the injected electrons can scatter into the quantized states and then tunnel through the resonant energy levels with or without further scattering [31]. Current transport through the emitter accumulation layer is important since it can change the device current significantly [109].

The coherent tunnelling approximation is a single-band model, which does not include the effects of band nonparabolicity (which is due to the $\vec{k}.\vec{p}$ repulsion of the light hole and conduction bands) [94]. Nonparabolicity in the conduction band causes the effective mass to rise as energy increases above the band minimum, which lowers the resonance energies. This is specially important for RTDs made of small bandgap materials, such as InGaAs lattice-matched to InP. Also important is the decay length of the wave function in the barriers, which is also affected by nonparabolicity, being significantly longer in the two-band model [31][94]. The single-band models can easily underestimate the current density by a factor of two or more. Several formulations incorporating optical phonon scattering mechanisms have been attempted with varying degrees of success [109][110].

Recent descriptions using envelope functions have incorporated additional important physical aspects, such as space charge effects, the 2-D accumulation layer in the emitter region, multiband effects, and phonon scattering [2][94][31]. These models cannot in principle treat the time-evolution of a system with internal force terms, such as inelastic phonon scattering [102].

Kinetic Models [2][30]

A more realistic formalism must include scattering processes due to longitudinal-optical phonons and ionised impurities, which destroy the phase-coherence of the electron waves. These dissipative tunnelling processes result in broadening of the transmission peaks at resonance and a degradation of the peak-to-valley current ratio. The scattering processes may also influence the electron distribution in the quantum well.

A proper description of the complex transport properties and of the heavily doped contact regions (and, if used, undoped spacer regions as well), beyond the coherence tunnelling formalism,

is given in terms of *non-equlibrium quantum transport models* adopting a dissipative quantum transport theory, such as the density matrix, Wigner function or non-equilibrium Green's function theory. Phonon scattering is treated semiclassically in these models, within the relaxation time approximation.

In principle, the fundamental difference between the coherent (or dynamic) and kinetic models is that the former are based on the single-particle picture using the pure state description, while the latter are based on the many-particle picture with mixed state description (due to the statistical nature of quantum theory, a definite, conservative Hamiltonian for the system cannot be specified).[7] These formalisms have not been used for device modelling because of their difficult implementation. Major problems, such as boundary conditions and correct initial conditions for the density matrix and Wigner function, are still under discussion [93][95][102].

Other Models

RTD quantum models based on kinetic theories are too complicated and not sufficiently accurate to be directly called from a device simulator. However, there is a need of a relatively simple analytic model that closely relates essential RTD physical parameters with measured *I-V* characteristics. Analytical approximations for the I-V characteristics based on transmission calculations with adjustable empirical parameters, linked as closely as possible to real physical quantities, are usually implemented [114]. Recently, attempts have begun to incorporate the NDC into commercial circuit simulators based on SPICE-type computer design tools [113][114][115].

There are several important issues to be noted when comparing theoretical and experimental results, including uncertainties in material and doping profiles, possible errors in growth data, material parameters employed such as effective masses and band offsets, and effects of imperfectly known parasitic impedances on the measured I-V curve. These factors make it very difficult to obtain an accurate prediction or reproduction of experimental I-V characteristics. Finally, inclusion of series and contact resistances, which modify the device I-V characteristics, is conceptually simple but very difficult to be realized in practice. Usually, a rough estimate or measurement is made to fit the measured I-V curve, which in general is not satisfactory if the RTDs being used have large fluctuations in their resistance values and, therefore, in their peak and valley parameters. A recently proposed technique to determine the RTD I-V characteristics experimentally, based on measurement of microwave reflection coefficients, may be useful in this regard [116].

At present these models cannot accurately predict the I-V characteristics for the whole range of RTD design parameters.

[7] In kinetic quantum transport models, a quantum system is considered to be a particle ensemble, characterized by the density matrix, the Wigner distribution function, or Green's functions [30][102].

3.2.3 Material Systems

From the electronics point of view, the desirable material system would be silicon based, compatible with current silicon IC technology. Attempts at building RTDs in the Si/SiGe system have not yet been encouraging, with PCVRs not much larger than one, mainly because the conduction-band offset between Si and SiGe alloys has been limited to a few tenths of an electron volt [117]. This is not enough to produce the necessary quantum confinement for a strong resonant state in the DBQW structure. Other possibilities using materials with different bandgaps that are lattice matched to silicon, for example SiGeC materials, have been proposed [118], but a real breakthrough is yet to come.

A proper treatment of boundary conditions, space charge distribution effects and scattering mechanisms requires the knowledge, with sufficient accuracy, of structural and material parameters such as geometric dimensions, doping profile and other material characteristics. The relevant material parameters in DBQW-RTDs are the energy-band offset, particularly the Γ valley conduction band offset, $\Gamma - X$ and $\Gamma - L$ valleys energy difference and the materials effective masses. Most of the work on RTDs is based on III-V compound semiconductor materials due to the relatively large energy-band offset between different lattice matched materials. RTDs made from these semiconductors can achieve very high peak current density, sufficient large PVCR, and very high speed [2][11].

GaAs/AlAs RTDs [20][67]

The original material system, $GaAs/Al_xGa_{1-x}As$ ($x \leq 0.45$) [1][119], has been replaced by GaAs/AlAs [120]. AlAs is used in the DBQW barriers because it presents a higher value of the barrier potential for Γ-valley electrons, yielding a larger PVCR when compared with $Al_xGa_{1-x}As$ with $x < 0.4$ [107][119]. The barrier height determines the width of the transmission resonance lobes. The peak current density J_p is approximately proportional to the width ΔE_n of the resonant level through which a great part of the resonant current flows (Eq. 3.12). Moreover, the AlAs barriers are free from various complexities associated with AlGaAs barriers, such as statistical fluctuations of the interfaces, barrier heights, and thicknesses.

The profiles of the three lowest conduction band minima of GaAs/AlAs heterojunctions are shown in Fig. 3.7. For the GaAs/AlAs heterojunction, the $\Gamma_{GaAs} - \Gamma_{AlAs}$-points offset, $\Delta E_{C_{\Gamma\Gamma}}$, is 1.12 eV, the $\Gamma_{GaAs} - L_{AlAs}$-points offset, $\Delta E_{C_{\Gamma L}}$ is 0.45 eV, and the $\Gamma_{GaAs} - X_{AlAs}$-points offset, $\Delta E_{C_{\Gamma X}}$, is 0.27 eV, which correspond to the Γ, L, and X barrier heights, respectively.[8] Also shown are the $\Gamma - X$ and $\Gamma - L$ valley separations in GaAs, which are approximately 0.48 eV and 0.28 eV, respectively. Because the $\Gamma_{GaAs} - X_{AlAs}$ barrier height is much smaller than the $\Gamma_{GaAs} - \Gamma_{AlAs}$ and the $\Gamma_{GaAs} - L_{AlAs}$ barrier heights, the $\Gamma_{GaAs} - X_{AlAs}$ barrier plays a significant role in carrier transport.

Tunnelling through the X-valley was first reported by Bonnefoi *et al* and Mendez *et al* [104][105]. It is now well established that tunnelling through this $\Delta E_{C_{\Gamma X}}$ barrier is a dominant

[8] For the GaAs/AlAs heterojunction, the $\Gamma_{GaAs} - \Gamma_{AlAs}$ valence band offset, ΔE_V, is 0.474 eV [23].

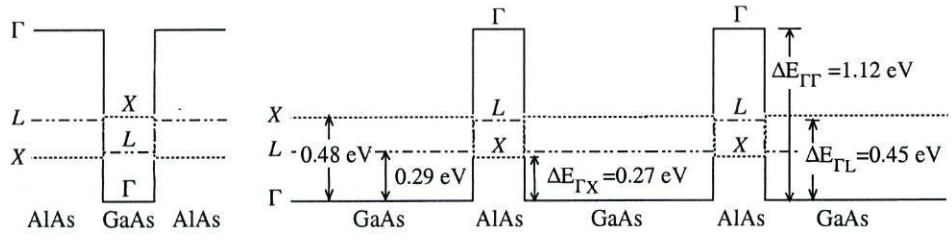


Figure 3.7: AlAs/GaAs/AlAs and GaAs/AlAs/GaAs/AlAs/GaAs $\Gamma-$, L- and X-conduction band valleys energy profiles.

component of the valley current in a GaAs/AlAs DBQW; this has been confirmed through hydrostatic pressure experiments [105].[9] These studies have also shown that the resonant tunneling current, which is the dominant component of the current on the first PDC region, is determined by the Γ-band profile, and is essentially independent of the applied pressure [104]. The saturation velocity of electrons in n-GaAs layers is less than 1×10^7 cm/s [20].

Although impressive results have been obtained with GaAs/AlGaAs based RTDs, higher performance can be achieved with InGaAs/AlAs DBQW-RTDs [121].

InGaAs/AlAs RTDs

The discussion above, which illustrates the importance of X-valley related transport in GaAs/AlAs DBQWs, suggests that higher performance (higher PVCR without sacrificing the peak current density, J_p) would be achieved if the $\Delta E_{C\Gamma X}$ barrier height could be increased. This is possible if the lower bandgap material (GaAs) is substituted by $In_{0.53}Ga_{0.47}As$ [121][122]. The conduction bandedge for the $In_{0.53}Ga_{0.47}As$/AlAs interface is shown in Fig. 3.8, where the $\Delta E_{C\Gamma\Gamma}$ and $\Delta E_{C\Gamma X}$ barrier heights are now 1.2 eV and 0.65 eV, respectively [12][121].

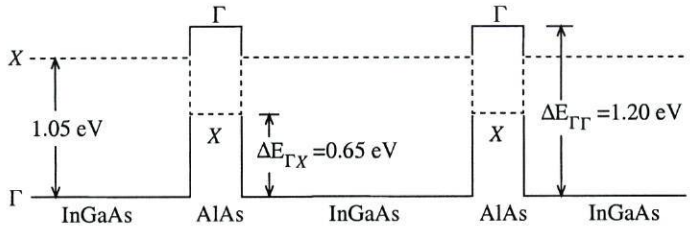


Figure 3.8: $\Gamma-$ and X-conduction band valleys energy profiles for $In_{0.53}Ga_{0.47}As$/AlAs heterojunctions.

The $\Gamma - X$ valley separation in $In_{0.53}Ga_{0.47}As$ is approximately 1.05 eV, which is higher than in GaAs. Comparison between the two conduction band profiles, Figs. 3.7 and 3.8, while showing similar $\Delta E_{C\Gamma\Gamma}$ barrier height, indicates that the $\Delta E_{C\Gamma X}$ barrier height is considerably higher in the $In_{0.53}Ga_{0.47}As$/AlAs case. Therefore, this material system is superior to GaAs/AlAs since the parasitic $\Gamma - X$ mediated transport will be reduced. Indeed, PVCRs as high as 30 at room

[9]These experiments rely on the fact that the $\Delta E_{\Gamma X}$ barrier height varies strongly with applied pressure, decreasing at a rate of 11 meV/kbar, while the $\Delta E_{\Gamma\Gamma}$ barrier height is essentially independent of the applied pressure [105].

temperature (with J_p about 5 kA/cm^2) [122], and peak current densities as high as 460 kA/cm^2 (with a PVCR of 4, resulting in a ΔJ of 345 kA/cm^2 - the highest ΔJ published) have been reported [121]. Another advantage results from the lower effective mass of conduction electrons in the $In_{0.53}Ga_{0.47}As$ material. The lower electron effective mass, the higher the tunnelling probability - higher J_p - and, possibly, the lower the scattering probability of tunnelling electrons resulting in reduced J_v. Indeed, $In_{0.53}Ga_{0.47}As$/AlAs pseudomorphic DBQW-RTDs exhibit both a remarkably high PVCR and large peak current density, as demonstrated by the I-V curve in Fig. 3.9 of a 200 μm^2 RTD grown by MBE, with 2 nm thick barriers and a 6 nm wide well.

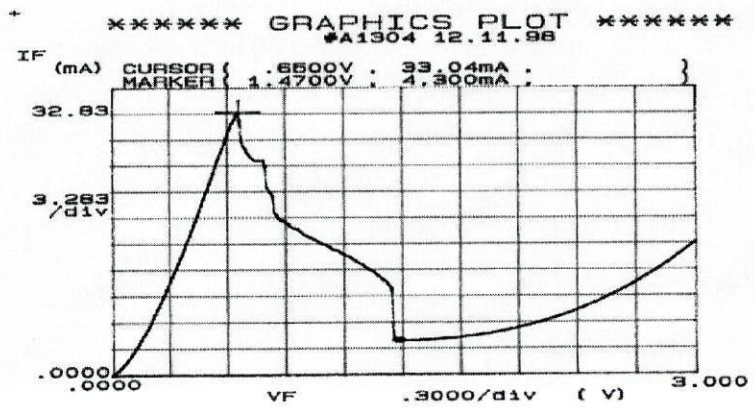


Figure 3.9: Experimental $In_{0.53}Ga_{0.47}As$/AlAs DBQW-RTD I-V characteristic.

Because $In_{0.53}Ga_{0.47}As$ has a Schottky barrier height of 0.2 eV (lower than for GaAs, which is approximately 0.4 eV) and can be very heavily doped, non-alloyed ohmic contacts with resistivity as low as 3×10^{-8} Ωcm^2 [11][21] (about 10^{-6} Ωcm^2 for GaAs [5]) are achieved. This is an important point since the extrinsic voltage drop at the ohmic contact could be serious (with r_c of $10^{-6} - 10^{-5}$ Ωcm^2) when the current density is as high as 10 kA/cm^2. Saturation velocities of electrons in n-$In_{0.53}Ga_{0.47}As$ layers are as high as 1.5×10^7 cm/s [11][21], a significant improvement compared to the GaAs/AlAs system. These two properties make clear that the InGaAs/AlAs pseudomorphic material system is the preferred are for high speed oscillator applications.

$In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ [123] has been also employed because of lattice matching and the $In_{0.52}Al_{0.48}As$ light electron effective mass (which causes high tunneling probability, and hence high J_p) [123]. However, AlAs instead of $In_{0.52}Al_{0.48}As$ is used as the barrier material because it presents a higher barrier height for the Γ-valley electrons, which is effective in minimising the non-resonant tunnelling current through higher lying quasi-continuum sub-bands, which constitutes a substantial portion of the non-resonant current. AlAs acts also as an efficient thermal barrier, although complications may arise from the crossover of the various conduction band minima.

InAs/AlSb RTDs

The InAs/AlSb system has been suggested as an alternative to the GaAs/AlAs and InGaAs/AlAs systems due to the very high conduction band offsets available: 1.8 eV for $\Gamma - \Gamma$ barrier height and 1.35 eV for the $\Gamma - X$ barrier (InAs Γ-point to AlSb X-point). This is expected to reduce X-point mediated tunnelling currents and, consequently, increase the PVCR. InAs exhibits a negative Schottky barrier, which implies exceptionally low contact resistance (5×10^{-9} Ωcm^2) leading to calculated maximum operating frequency values exceeding 1 THz for a 4 μm^2 RTD [3]. The highest published [6] operating frequency for a RTD, 712 GHz, was achieved in the InAs/AlSb material system taking advantage of a variety of subtle effects which favour tunnelling through the barrier in this staggered-bandgap material [143]. However, from a device application perspective, InAs/AlSb DBQW-RTDs may suffer from impact ionization events occurring in the very narrow bandgap InAs material (E_g =0.35 eV) [12][53][143]. This is particularly true for high current density RTDs, where electric fields on the order of several hundred kV/cm are encountered [2].

3.2.4 Structural Parameters

The DBQW-RTD characteristics can be adjusted up to a certain point, to its specific application by judicious choices of the size and properties of the various semiconductor layers within the device. For a given material system, to optimise the performance of DBQW-RTDs, one can vary both structural (well and barriers widths, L_w and L_b respectively; barrier height U_0, and impurity concentration N_d in the electrode regions outside the DBQW) and growth parameters (such as substrate temperature). Resonant tunnelling diodes are often fabricated with low-doped/undoped spacer layers. Relatively thick spacer layers produce major effects on device performance: under applied bias a significant portion of the spacer layer may be depleted, resulting in a transit time delay which can dominate the device frequency response. Well, barriers and spacer layers characteristics affect both the resonant and the non-resonant current components and are, therefore, critical parameters in DBQW-RTD design.

Quantum-Well Width

To decrease the unwanted current through higher resonant states, and thus to increase the PVCR, the well region should be thin. When the well width L_w is decreased, the resonant tunnelling current is predicted to rise because it leads to lower effective barrier height $U_0 - E_n$ for incident electrons with energy around E_n; this decrease of the barrier effect gives rise to broadening of the transmission coefficient around the resonance peaks, ΔE_n, which results in larger resonant tunnelling current through the dominant resonant level (in general, the first resonance), Eq. 3.12.

However, both peak and valley current densities increase as the well width is reduced [125]. As the thickness of the well region is decreased, the voltage at which the peak resonance occurs will increase due to a rise in the lowest resonant state energy with respect to the bottom of the

well, E_{C_w}, which also increases the current due to thermionic effects.[10] *Tsuchiya et al* observed that the valley current, which is a measure of the leakage current, depends superlinearly on the bias voltage V [125].

Barriers Thickness

Considering Eq. 3.10, J_p varies exponentially with barrier width, L_b, a fact that has been observed by Broekaert *et al* and Chow *et al* [126]. Modeling this exponential dependence of the peak current density on L_b as

$$J_p \propto \exp\left(-\frac{L_b}{\lambda_{c_b}}\right), \tag{3.23}$$

where L_b is in monolayers, ML, and λ_{c_b} is a characteristic length, it results that λ_{c_b}=1.3 ML, which indicates that a variation in barrier thickness on the order of one monolayer (~0.283nm) will cause the peak current density to change by a factor of two, approximately.

The PVCR is influenced mainly by the barriers width. The published results indicate that elastic tunnelling dominates for very thin (<5 nm) barriers, while thermionic emission is the primary transport mechanism in thick (>10 nm) barriers [119][120]. The current in the intermediate thickness range, however, appears to consist of both elastic tunnelling through the $\Delta E_{C\Gamma\Gamma}$ barrier and inelastic tunnelling through the $\Delta E_{C\Gamma X}$ barrier [104][119][120]. To decrease non-resonant tunnelling, the RTD should operate at relatively low bias voltage, which diminishes electron accumulation on the emitter side and increases the effective barrier height, as previously discussed.

The non-resonant current components (consisting of inelastic tunnelling current, non-resonant tunnelling current through higher lying levels, etc.) are strongly dependent on both L_b and barrier height U_0 [119][120].[11] The valley current increases exponentially with a decrease of the barrier thickness; however, the expected increase of the peak current with a decreased L_b is overshadowed by a larger increase of the valley current. Therefore, the overall effect is a decrease of the PVCR with a reduce of L_b [126][127]. The negative correlation between J_p and PVCR can be interpreted as follows: higher current density results in a higher concentration of accumulated electrons, which may increase electron scattering and increase J_v.

Spacer Layers

Under applied bias, the potential close to the emitter barrier in basic RTD structures tends to form a deep triangular quantum well, giving rise to quantized states [128]. Under certain conditions, approximately half of the applied bias is dropped in the cathode well. At the valley point the barriers are pulled down in energy, thus reducing the barrier height to a ballistic electron injected from the cathode, enhancing the valley current and degrading the PVCR. Inclusion of

[10] Care must be taken because the quantum well width is related to the amplitude of the electric field across the DBQW, which cannot be too strong to avoid impact ionisation [11].

[11] The thermionic current over the barriers is nearly independent of L_b, being primarily determined by the barrier height U_0, and increasing with temperature. [119]

low/undoped space layers between the highly doped contacts and the barriers leads to much smaller downward bending of the conduction band as bias is applied, and the charge distribution does not deplete as sharply as without the spacer layer [128][129]. Use of low doped spacer layers further increases the effective U_0 value by reducing the Fermi energy of the electrons close to the emitter barrier [129][130]; simultaneously, it reduces ballistic injection over the barriers, so that the PVCR is improved. The potential drop occurs mainly across the barriers.

The undoped spacer layer also helps to suppress diffusion of Si-impurities from n-GaAs or n-InGaAs into the AlAs barriers [131]; also the electron mean free path is longer where ionized donors are absent [11]. Finally, the bistability of the NDC region of the I-V curve is found to be much sharper [100][101].

Growth Conditions

The growth rate depends on the substrate temperature and, to a large extent, on the thickness of the layer to be grown. Typical substrate temperatures are around 560 to 640 ^{0}C, to suppress rapid diffusion of Al and Si, yielding sharper interfaces and cleaner barrier material [131].

Interruption of growth between deposition of layers can improve the morphological quality of the interfaces by allowing the surface kinetic processes to relax [56][57]. Intra-layer thickness fluctuations caused by the density of mono-layer terraces in the plane of the hetero-interfaces can have a detrimental effect on resonant tunnelling by weakening the coherence of the interfering electron waves reflected by the two barriers; in addition, they cause changes of the energy levels of the wells [131][132]. The method of interrupted growth is recommended for all interfaces of the RTD, including spacer layers.

3.3 Electronic Applications of RTDs

Most of the solid-state devices and circuits operating up to 300 GHz are, in general, advanced developments of conventional electronics where reduced dimensions or sophisticated layer structures are utilized. These devices may be broadly classified as *transit time devices* and include familiar examples as bipolar junction transistors (BJTs), heterojunction bipolar transistors (HBTs), field effect transistors (FETs), high electron mobility transistors (HEMTs), and transferred electron (Gunn) devices [10][11][49]. In all of these, the time taken for carriers (usually electrons) to move a characteristic distance determines the maximum frequency of operation. An interesting case is the DBQW-RTD, which currently holds the record as "the fastest purely electronic device".[12]

RTDs are *voltage-controlled negative differential conductance* devices, which can be designed to have very high cut-off frequencies due to the short distance through which carriers have to travel. Because of their non-linear I-V characteristic, various applications, such as frequency multipliers, signal generators, oscillators, self-oscillating mixers with gain, have been proposed

[12] Although the DBQW-RTD ultimate frequency limit is set by the width of the barrier transmission resonance, the device can also be limited by transit time effects in the depletion region [11].

and demonstrated in both passive and active operating modes [2][11][142]. Their N shaped current-voltage characteristic makes possible to operate RTDs as so-called *functional devices*, enabling the implementation of functions such as multi-valued memory cells in signal processing circuits (reducing significantly the number of elements per circuit). Functional applications using RTDs are highly promising since their simple structure and small size favours their integration with conventional devices, such as transistors.

This section describes briefly RTD applications in millimetre/submillimetre-wave generation, high speed switching, storage and signal processing. The small-signal equivalent circuit of an RTD and its ability to self-oscillate are first addressed. The problem of RTD dc stability in the negative differential conductance region (which is highly relevant to applications) is also discussed. Before introducing the RTD small-signal equivalent circuit, the *tunnel diode* is briefly analyzed.

3.3.1 Tunnel Diode

In 1958, while studying the internal field emission in a heavily doped p^+-n^+ junction, L. Esaki observed an "anomalous" current-voltage characteristic in the forward bias direction corresponding to a negative differential conductance (NDC) region over part of the forward characteristic: *the tunnel diode* [49]. He explained the negative differential conductance region as a consequence of electron conduction through tunnelling (from the conduction band of the degenerate n^+ doped region into empty states in the valence band of the degenerate p^+ doped region of the semiconductor).

Tunnel Diode Current-Voltage Characteristic [11][49]

Figure 3.10 shows a schematic energy diagram of a tunnel diode in thermal equilibrium, and a representative current-voltage characteristic. In the reverse bias direction (p-side negative with respect to n-side), the current increases monotonically; in the forward direction, however, the current first increases to a maximum value I_p (peak current) at a voltage V_p, then decreases to a minimum value I_v (valley current) at a voltage V_v. For voltages larger than V_v, the current increases exponentially with the voltage.

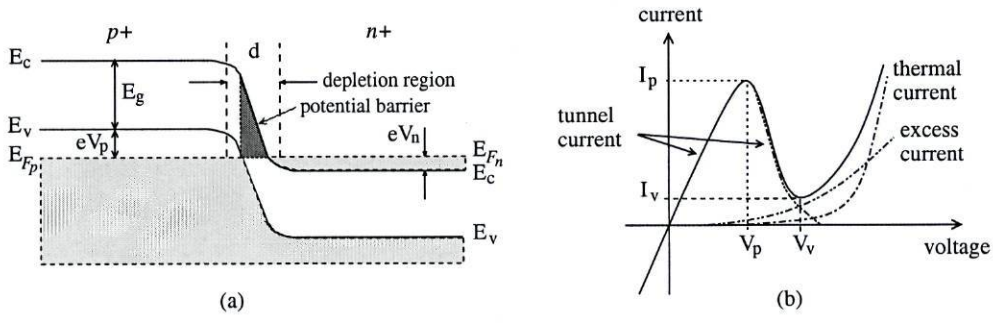


Figure 3.10: Band diagram of a $p^+ - n^+$ junction under thermal equilibrium; (b) static current-voltage characteristic of a a $p^+ - n^+$ *junction* [49].

To understand the $p^+ - n^+$ junction I-V characteristic, tunnelling through the triangular barrier formed in the depletion region, represented in Fig. 3.10(a), has to be considered. Because of high doping, the junction depletion region thickness, d, is on the order of a few nanometres, and the p and n quasi-Fermi energy levels, $E_{F_{(p,n)}}$, are located within the valence and conduction bands, respectively; the degeneracies, eV_p and eV_n, amount typically to about 50 to 100 meV. (For simplicity, it is assumed that the device is operated at low temperatures, so there are no filled states above the Fermi level and no empty states below the Fermi level on either side of the junction.)

As a forward voltage, V_f, is applied, Fig. 3.11(a), the electrons energy in the n-side is risen relative to that on the p-side, and a range of energies exists in which filled states on the n-side correspond to unoccupied states on the p-side, that are available to accept electrons. They can thus tunnel from the n-side to the p-side, giving a significant direct current: the tunnelling current, shown in Fig. 3.10(b). Continueing to increase the forward voltage, the current reaches a local maximum (peak current) corresponding to maximum crossing of electrons from filled states on the n-side to unoccupied states on the p-side. This corresponds to the first positive differential conductance (PDC) region of the junction I-V characteristic ($V_f < V_p$), Fig. 3.10(b).

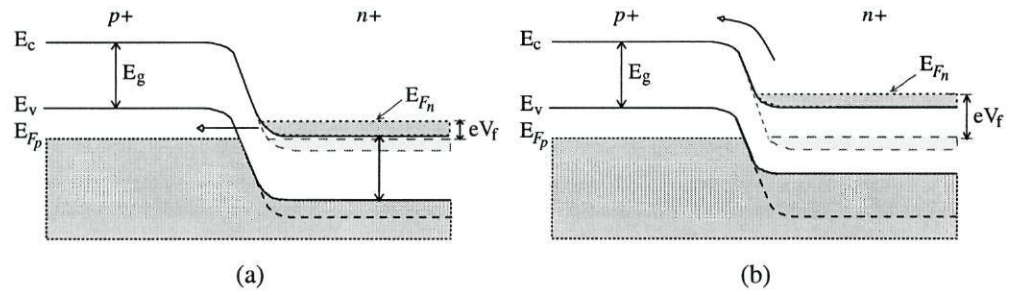


Figure 3.11: Forward biased p^+-n^+ junction band diagram [49].

Continueing to increase the forward voltage, the number of available unoccupied states on the p-side starts to diminish, decreasing the tunnelling current: negative differential conductance (NDC) region ($V_p < V_f < V_v$), Fig. 3.10(b). When the applied bias is such that the bands are "uncrossed", i.e., the bottom of the conduction band is exactly levelled to the top of the valence band, Fig. 3.11(b), there are no available energy states that are opposite to filled states (thus, the tunnelling current can no longer flow). With a further increase of the voltage, the normal diode current will flow and will increase exponentially with the applied voltage: second PDC ($V_f > V_v$), Fig. 3.10(b). When the diode is reverse biased, the n-side energy is lowered with respect to the p-side, and electrons can tunnel from the valence band into the conduction band. The reverse current will increase monotonically with the reverse bias, because more occupied states on the p-side and more unoccupied states on the n-side become available as the reverse voltage increases.

The tunnelling current density is related to the tunnelling probability, T_t, by Eq. 2.28. A large tunnelling current (high tunnelling probability) can be obtained for high doping concen-

tration, low carrier effective mass, and low bandgap energy. The tunnelling current is essentially independent of the temperature, since the tunnelling probability varies only slightly with temperature through E_g. In real devices, the forward I-V characteristic is the result of three current components [Fig. 3.10(b)]: the tunnelling current, the excess current, and the thermal current.

Tunnel diode figures of merit are peak-to-valley current and voltage ratios, $PVCR = I_p/I_v$ and $PVVR = V_p/V_v$, respectively, which are related with the amount of NDC, maximum output power and dc-to-ac conversion efficiency. In tunnel diodes there is little control over these parameters [49], the opposite case to RTDs where the peak current, peak-to-valley current ratio, voltage at peak current, and device capacitance can all be partly controlled.

Tunnel Diode Small-Signal Equivalent Circuit [11][49]

The small-signal equivalent circuit of a tunnel diode is shown in Fig. 3.12, and consists of the series resistance R_s, the junction capacitance C_d, and the diode negative differential resistance $-R_d$ (which is given approximately by $-2V_p/I_p$ [49]).

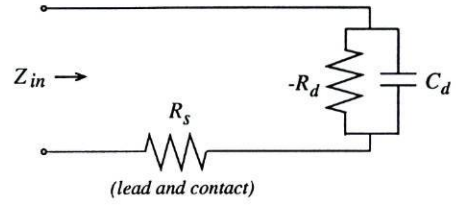


Figure 3.12: Small-signal equivalent circuit of the tunnel diode [49].

The real part of the input impedance of the equivalent circuit, $Z_{in}(f)$, as a function of frequency, f, is given by

$$R_{in}(f) = R_s + \frac{-R_d}{1 + (2\pi f R_d C_d)^2}. \tag{3.24}$$

When a current I flows into the diode, the electrical power dissipated is given by $R_{in}(f)I^2$. Therefore, under certain conditions ($R_s < R_d$), there is a range of frequencies for which $R_{in}(f)$ is negative, and the diode can supply electrical power to the external circuit, thus exhibiting gain. The maximum frequency of oscillation of the diode, f_{max}, is defined as the frequency at which $R_{in}(f)$ becomes zero.[13] For the equivalent circuit given in Fig. 3.12, f_{max} is:

$$f_{max} \equiv \frac{1}{2\pi R_d C}\sqrt{\frac{R_d}{R_s} - 1}. \tag{3.25}$$

The maximum output power for a realistic I-V characteristic is given by

$$P_{max} = \frac{3}{16}(\Delta I \cdot \Delta V) = \frac{3}{16} I_p V_p \left(1 - \frac{I_v}{I_p}\right)\left(\frac{V_v}{V_p} - 1\right); \tag{3.26}$$

[13] f_{max} corresponds to the frequency at which the power gain drops below one.

and the maximum dc-to-ac conversion η is

$$\eta = \frac{P_{max}}{P_{dc}}, \tag{3.27}$$

where $P_{dc} = (I_p - \Delta I/2)(V_p + \Delta V/2)$.

To achieve both maximum power output and conversion efficiency, the ratios I_p/I_v and V_v/V_p must be maximized. Experimental oscillation frequencies up to a few hundred GHz, with a power output of a few tenths of a micro-watt, have been observed; higher power outputs have been obtained at lower frequencies: about 1 mW at 10 GHz and 0.2 mW at 50 GHz [11]. The tunnel diode is used in special low-power microwave applications, such as a local oscillator for satellite communication and high frequency signal sampling.

3.3.2 RTD Small-Signal Equivalent Circuit

The small-signal equivalent circuit of a RTD, represented in Fig. 3.13(a), includes the following elements [139][140]: a series resistance, R_s, arising from ohmic the contacts, the emitter and collector regions, and current spreading effects [5]; the parallel capacitance, C_d, resulting from charging and discharging of electrons at the DBQW and depletion regions; the negative differential resistance, $-R_d$ (representing the corresponding negative differential conductance, $-G_d = -R_d^{-1}$); and an inductance element, L_{qw}, in series with $-R_d$ (associated with the delay of the RTD current with respect to the voltage, which arises from the time necessary to build up the charge in the quantum well in order to change the diode current). L_{qw} is chosen so that the product of L_{qw} and G_d is equal to the electron dwell time in the DBQW, τ_{dw} [5]. Also shown is the tunnel diode small-signal equivalent circuit, Fig. 3.13(b).

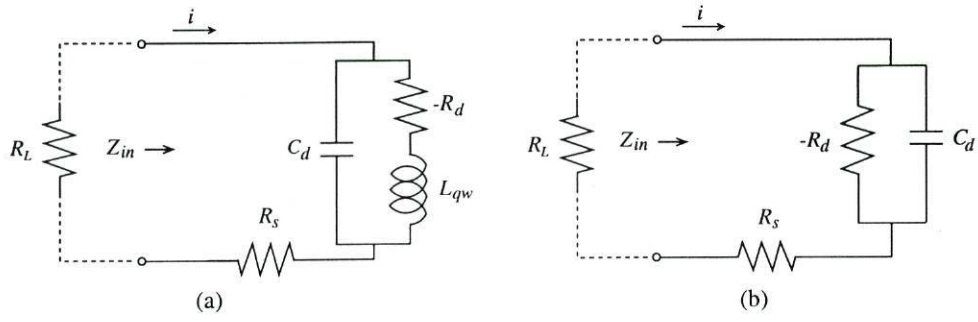


Figure 3.13: Small signal equivalent circuit of a RTD (a) and of a tunnel diode (b) [5].

When a current $i(f)$ flows into the RTD, the electrical power spent in the diode is given by $R_{in}(f) \cdot i^2$, where $R_{in}(f)$ is the real part of the diode impedance, Z_{in}, at frequency f. Therefore, if $R_{in}(f)$ is negative, the RTD supplies electrical power to the external circuit at frequency f, thus exhibiting gain. To sustain oscillation, the real part of the input impedance of the RTD small-signal equivalent circuit must be negative. The maximum frequency of oscillation of a RTD, f_{max}, defined as the frequency at which $R_{in}(f) = 0$, and for the equivalent circuit represented

in Fig. 3.13(a), is given by [5][139]

$$f_{max} = \frac{1}{2\pi}\left\{\left[\frac{1}{L_{qw}C_d}\left(1-\frac{C_d}{2L_{qw}G_d^2}\right)\right]\times\left(1-\sqrt{1-\frac{\left(1+R_sG_d\right)/R_sG_d}{\left(C_d/2L_{qw}G_d^2-1\right)^2}}\right)\right\}^{\frac{1}{2}}. \tag{3.28}$$

The value of L_{qw}, or of τ_{dw}, is mainly determined by the barrier thickness. A reduction in L_{qw} (using thin barriers) is effective to increase f_{max}, enabling oscillation at very high frequencies, up to 712 GHz [6].

Three time response figures for RTDs are commonly discussed in the literature [11]: the *resonant state life time* τ_{dw}, the *depletion layer transit time* τ_d, and the *device RC time* τ_{RC}. Based on the uncertainty relations, the intrinsic time of the electron in the well appears to be on the order of hundreds of femtoseconds. For depletion lengths less than 100 nm, the depletion layer transit times are also sub-picosecond. Typical switching times of best devices published to date are dominated by the effects of current densities and capacitances, i.e., by the RC time constants.

Neglecting the resonant tunnelling time delay, i.e., assuming L_{qw} to be negligibly small, the RTD small-signal equivalent circuit degenerates in the small-signal equivalent circuit of the junction tunnel diode: the NDC element in parallel with the diode capacitance (the energy storage element), as shown in Fig. 3.13(b) [49]. In this case, the RTD f_{max} becomes (Eq. 3.25):

$$f_{max} \equiv \frac{1}{2\pi R_dC_d}\sqrt{\frac{R_d}{R_s+R_L}-1}. \tag{3.29}$$

The highest frequency, f_{max}, (and of course zero power output, i.e., load resistance $R_L = 0$) is obtained when $R_d = 2R_s$, being

$$f_{max} = \frac{1}{4\pi R_sC_d}. \tag{3.30}$$

In tunnel diodes the capacitance is relatively large since, in order to create NDC, the doping level must be sufficiently high for the semiconductor to become degenerate, $\approx 10^{19}$ cm^{-3}, which leads to a very thin depletion region. In contrast, the RTD capacitance, C_d, can be designed by choosing both the structural parameters of the DBQW and the surrounding doping profile, and thus the depletion thicknesses. The main contribution to the RTD capacitance is from the depletion region; since doping is, in general, much lower ($10^{16} - 10^{18}$ cm^{-3}), the depletion capacitance is considerably smaller, making a much higher oscillation frequency possible. The lowly doped and relatively thick collector spacer layer, normally introduced between the barrier and the highly doped collector contact, yields small C_d [132] (as the total device thickness increases, the transit time of the electrons across the whole structure becomes significant [5]); the emitter spacer layer will contribute mainly to the total parasitic series resistance [129].

The high frequency performance of RTDs is also improved by increasing G_d, or decreasing R_s. To obtain a large value of G_d it is necessary to achieve a high peak current density, J_p, and a low valley current density, J_v. High J_p can be expected for RTDs consisting of double-

barrier structures with a high transmission probability and thus a shorter electron dwell time (low L_{qw}). RTDs with J_p as high as 450 kA/cm^2 and PVCR=4 have been reported [126], for which an oscillation frequency of 1 THz can be expected, provided that other parasitics are negligible. To a great extent, R_s (the parasitic series resistance due to any undepleted spacer regions, highly doped contact regions, and ohmic contacts) is also determined by the material used for the emitter and the collector: lower specific contact resistivities significantly improve the RTD performance [5].

Recent work [5][134][137] has demonstrated that the RTD cannot be described by the simple equivalent circuits of Fig. 3.13, because a relatively thick depletion region is formed "downstream" from the DBQW. A substantial part of the terminal voltage may be dropped across it, as shown in Fig. 3.14. An accumulation layer is also located close to the opposite side (cathode) of the DBQW. The width of the depletion layer can be varied by changing doping and thickness of the spacer layer on the collector side.

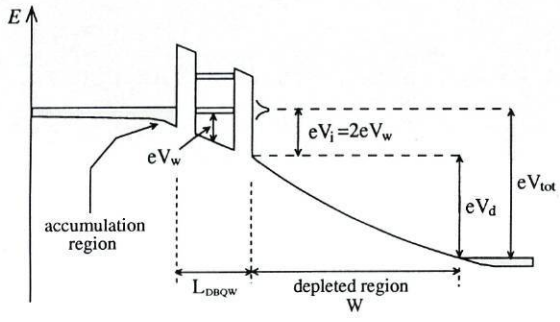


Figure 3.14: $\Gamma-$conduction band energy diagram of a RTD with applied bias.

Because the DBQW is much thinner than the depleted region thickness W, the depletion region transit time, $\tau_d = W/v$ (where v is the carrier velocity), is much greater than the transit time through the DBQW; transport through the DBQW can be considered instantaneous in this approximation. If the transit time, τ_d, is the dominant delay time of an RTD, it can be shown [134] that f_{max} is given by

$$f_{max} = \frac{1}{2\pi}\frac{v_{sat}}{2W}\left[\pi + \sqrt{\pi^2 - \frac{8|\sigma|W}{\epsilon v_{sat}}}\right], \tag{3.31}$$

where σ represents the injection conductance ($\sigma = \partial J/\partial \mathcal{E}$), and v_{sat} is the electron saturation velocity. It is clear that, for a given RTD, W should be small in order to achieve a high f_{max}. There is an optimum value for W which comes from a compromise between the reduction of C_d and the reduction of transit time delay, τ_d.[14] The frequency f_{max} can also be increased by employing emitter and collector materials with higher electron saturation velocity, v_{sat}.

[14]Since the double-barrier structure and the collector spacer layer are either undoped or low doped regions sandwiched between two moderately or heavily doped sections, the RTD capacitance can be approximated by a parallel-plate capacitance model. (A more accurate value of the device capacitance may be found by using a self-consistent quantum mechanical simulation to calculate the width of the depletion region [155].)

The negative differential conductance is an intrinsic property of the double barrier quantum well structure and can only be partially controlled by device design; it remains constant up to a characteristic frequency, then rolls-off with the frequency [133]-[138] (the device series resistance is also frequency dependent [5]; assuming the contact resistance and the epilayer resistance to be constant, it increases with frequency, mainly due to the behaviour of the skin depth effect [5]). Such roll-off will, however, lower the device resistive cut-off frequency and also limit the power-generating capabilities of RTDs at very high frequencies [133][137]. In a real setup, the maximum attainable frequency and power are mainly determined by parasitic access resistance, lead inductance, and impedance matching between the RTD and the external circuitry.

3.3.3 RTD Stability Analysis

The RTD is a voltage-controlled negative differential conductance device; for most practical applications [158][159], a voltage-controlled NDC in combination with the bias circuit should be stable at low frequencies, when biased in the NDC portion of its I-V characteristic; Fig. 3.15 shows experimental low frequency bias circuit oscillation of a RTD biased in the NDC region.

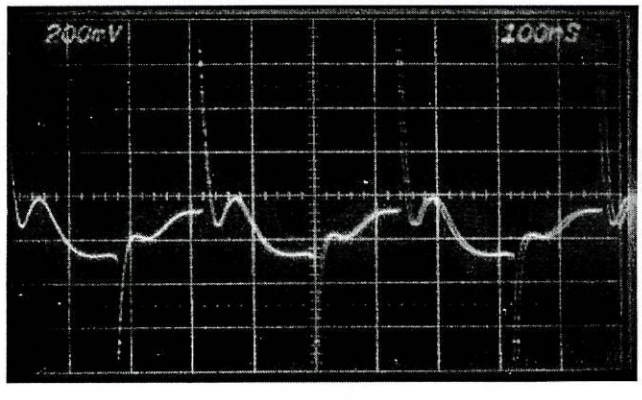


Figure 3.15: Experimental RTD low-frequency bias oscillation at around 4 MHz.

In oscillator operation at high frequency, low frequency bias circuit instabilities introduce unwanted up-converted signals [158][159].

Self-Oscillation

To demonstrate how NDC can lead to self-oscillation, a simplified small-signal equivalent circuit of a RTD mounted in a resonating circuit cavity is schematically represented in Fig. 3.16: the circuit consists of a negative differential resistance (NDR), $-R_d$, connected to a load, R_L, via a resonating LC circuit.

The Laplace-transformed voltage response to a current fluctuation is given by [22]

$$v(s) = \left(\frac{1}{R_L} - \frac{1}{R_d} + \frac{1}{sL} + sC\right)^{-1} i(s), \tag{3.32}$$

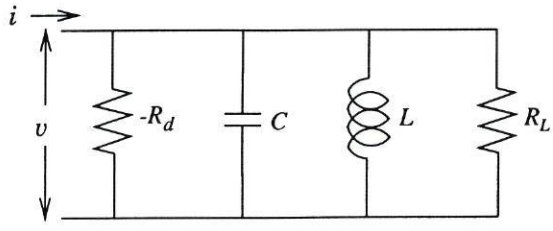


Figure 3.16: Small-signal equivalent circuit of a NDR connected to a load via a resonating LC circuit [22].

which may be written as

$$v(s) = \frac{s}{2i\omega_0 C}\left(\frac{1}{s-\gamma-i\omega_0} - \frac{1}{s-\gamma+i\omega_0}\right) i(s), \tag{3.33}$$

with $\gamma = (1/R_d - 1/R_L)/2C$ and $\omega_0 = 1/LC - \gamma^2$. In the time domain the response to a current pulse $i(t) = I_0\delta(t)$ is, therefore,

$$v(s) = I_0 \frac{e^{\gamma t}}{C}\left(\frac{\gamma}{\omega_0}\sin\omega_0 t + \cos\omega_0 t\right), \tag{3.34}$$

which shows that if $R_d < R_L$, then $\gamma > 0$ and any finite current fluctuation will cause an oscillation to start with exponentially increasing amplitude. (Of course, the amplitude will stabilize when nonlinearities increase the average dynamic resistance of the circuitry.)

Stability Analysis [141][158][159]

In the presence of low-frequency oscillations it is not possible to bias the device at a stable point in the NDC region. Fig. 3.17 presents the small-signal equivalent circuit of an *assembly* containing a RTD, including parasitic elements.

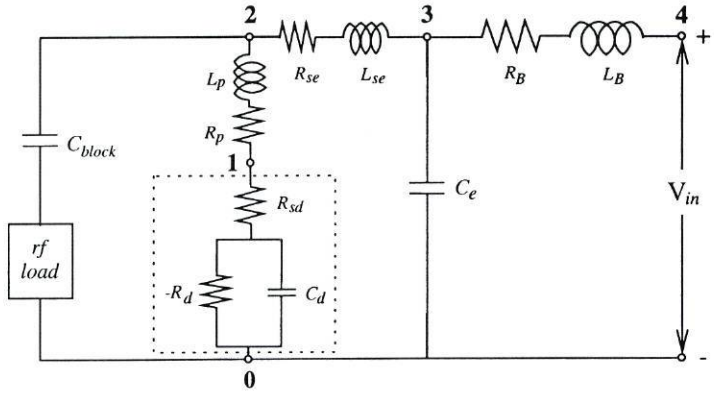


Figure 3.17: Small-signal equivalent circuit of an assembly containing a RTD [141].

The circuit section between nodes **1** and **0** represents the intrinsic device behaviour: $-R_d$ is the NDR of the device, C_d is the device capacitance, and R_{sd} is the positive parasitic resistance

of the device.[15] The circuit section between nodes **2** and **1** represents connecting circuit wires, with L_p and R_p being their inductance and resistance, respectively. The rf signal is isolated from the bias circuit by a low-pass filter (LPF), such as a bias tee (at low frequency, the LPF can be modelled as a series inductance), represented by the elements between nodes **2** and **3**. The bias signal is isolated from the *rf* load by a blocking capacitor C_{block}. Circuit elements R_{se}, L_{se}, and C_e can be used in an attempt to improve the low-frequency stability without affecting the *rf* impedance seen by the device. Finally, R_B and L_B represent the resistance and inductance of the power supply.

The impedance of the RTD coupled to the cavity (across nodes **2-0**), is given by

$$Z_{20} = R_{20} + \frac{-G_d}{G_d^2 + \omega^2 C_d^2} + j\omega \left(L_p - \frac{C_d}{G_d^2 + \omega^2 C_d^2} \right), \tag{3.35}$$

where $R_{20} = R_{sd} + R_p$. If the magnitude of R_{20} is less than R_d, then the real part of Z_{20} is negative at low frequencies and is a monotonically increasing function of ω. At frequency ω_r, given by

$$\omega_r = \frac{1}{R_d C_d} \sqrt{\frac{R_d}{R_{20}} - 1}, \tag{3.36}$$

the real part of the impedance becomes zero; above this frequency the device can no longer supply power to the circuit. The imaginary part of the impedance Z_{20} becomes zero at

$$\omega_i = \sqrt{\frac{1}{L_{20} C_d} - \frac{1}{R_d^2 C_d^2}}, \tag{3.37}$$

which corresponds to the RTD oscillation frequency.

At sufficiently low frequencies, the device is *isolated* from the rf load. Assuming the voltage source is connected to the diode through the LPF, the circuit in Fig. 3.17 reduces to that in Fig. 3.18, where now $L_s = L_p + L_{se} + L_B$ and $R_s = R_{sd} + R_p + R_{se} + R_B$.

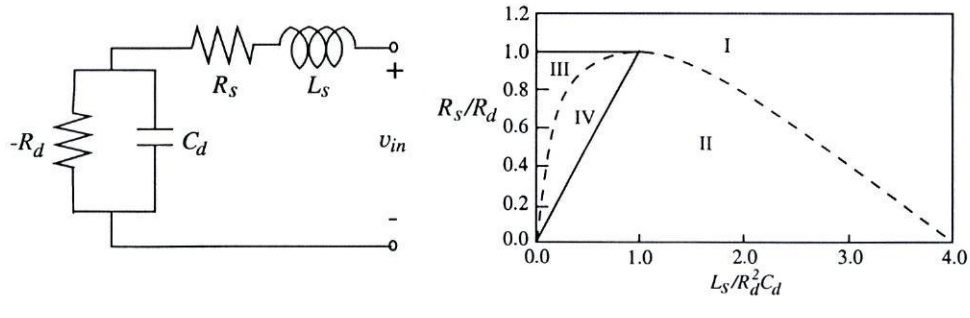


Figure 3.18: Equivalent circuit without a shunt capacitor, and respective stability diagram [141].

[15] In typical RTDs, the impedance associated to the dwell inductance, L_{qw}, can be negleted.

This circuit is described by the differential equation

$$L_s C_d \frac{d^2 v}{dt^2} + \left(R_s C_d - \frac{L_s}{R_d} \right) \frac{dv}{dt} + V \left(1 - \frac{R_s}{R_d} \right) = v_{in}, \tag{3.38}$$

with $v_{in} = 0$ for the short-circuit case. To this differential equation correspond four distinct possible solutions, in correspondence with the roots of its characteristic equation:

Case I: *Growing exponential* occurs if the solutions are real,

$$\frac{R_s}{R_d} > 2\sqrt{\frac{L_s}{R_d^2 C_d}} - \frac{L_s}{R_d^2 C_d}, \tag{3.39}$$

and at least one solution is positive, either

$$\frac{R_s}{R_d} > 1 \quad or \quad \frac{R_s}{R_d} < \frac{L_s}{R_d^2 C_d}. \tag{3.40}$$

Case II: *Growing sinusoidal* results if the solutions are complex,

$$\frac{R_s}{R_d} < 2\sqrt{\frac{L_s}{R_d^2 C_d}} - \frac{L_s}{R_d^2 C_d}, \tag{3.41}$$

with a positive real part,

$$\frac{R_s}{R_d} < \frac{L_s}{R_{dd}^2 C_d}. \tag{3.42}$$

Case III: *Decaying exponential* occurs if the solutions are real,

$$\frac{R_s}{R_d} > 2\sqrt{\frac{L_s}{R_d^2 C_d}} - \frac{L_s}{R_d^2 C_d}, \tag{3.43}$$

and both negative,

$$\frac{R_s}{R_d} < 1 \quad and \quad \frac{R_s}{R_d} > \frac{L_s}{R_d^2 C_d}. \tag{3.44}$$

Case IV: *Decaying sinusoidal* exists if the solutions are complex,

$$\frac{R_s}{R_d} < 2\sqrt{\frac{L_s}{R_d^2 C_d}} - \frac{L_s}{R_d^2 C_d}, \tag{3.45}$$

with a negative real part,

$$\frac{R_s}{R_d} > \frac{L_s}{R_d^2 C_d}. \tag{3.46}$$

This leads to the stability diagram of Fig. 3.18.

The circuit is stable if the solutions are of types III or IV, whose combination gives the

stability criteria for the circuit of Fig. 3.18:

$$\frac{L_s}{R_d^2 C_d} < \frac{R_s}{R_d} < 1. \tag{3.47}$$

Algebraic manipulation shows that the first inequality of Eq. 3.47 implies $\omega_r < \omega_i$ as a stability criterion, with the frequencies defined as before. Assuming that the designer has control over R_s (not all of R_s has to be part of the rf circuit), the stability criterion becomes

$$L_s < R_d^2 C_d. \tag{3.48}$$

For RTDs with peak currents in the mA range, typical values of R_d are tens of ohms or lower, and tens of picofarads for C_d. This constrains L_s to nanohenrys or even tenths of a nanohenry; even with an ideal bias supply, the bonding wires inductance is in this range.[16]

Stability Conditions [158][159]

If the oscillation frequency is calculated when L_s is just small enough to give stability ($L_s = R_s R_d C_d$, the first inequality of Eq. 3.47), the result is

$$\omega = \omega_i = \omega_r = \frac{1}{R_d C_d}\sqrt{\frac{R_d}{R_s} - 1}. \tag{3.49}$$

For a given device, R_d and C_d are fixed; assuming R_s to be smaller than R_d, the circuit inductance L_s controls the stability. The physical origin of L_s is the inductance due to the leads that connect the device to the measuring apparatus or the device package. The idea is to decrease L_s, which in turn increases the oscillation frequency, ω_i, until it is above the resistive cut-off frequency of the device, ω_r.

From the preceding analysis it is seen that instability would be overcome if the dc source could be inserted physically near the RTD, minimising the lead inductance. For practical applications, the device can be contacted either with bonding wires or through whisker contacts; at low frequencies the devices are usually bonded in microwave packages, and the associated inductance is around 1 nH [20][21].

The stability criterion demonstrates clearly that even a small amount of inductance associated with probe contacts or bonding wires will often be sufficient to produce an unstable circuit. A common method for stabilising a tunnel diode or RTD is to place a shunt capacitor across the terminals of the device. Insertion of a shunt capacitor, C_e, near the RTD, as shown in Fig. 3.19, can improve the stability, but only if the inductance between the RTD and the capacitor is minimised.

The large shunt capacitor, C_e, will appear to the bias circuit oscillations as a dc voltage source. For large C_e, on the order of a tenth of microfarad for the circuit under consideration,

[16] For the same R_d, a tunnel diode will be easier to stabilise due to its larger capacitance.

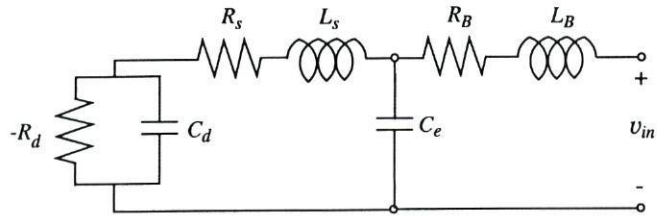


Figure 3.19: Equivalent circuit with a shunt capacitor [141].

Fig. 3.19, the stability criterion for the circuit is

$$\frac{L_s}{R_d^2 C_d} < \frac{R_s}{R_d} < 1, \tag{3.50}$$

and

$$\frac{L_B}{(R_d - R_s)^2 C_e} < \frac{R_B}{(R_d - R_s)} < 1. \tag{3.51}$$

For the circuit shown in Fig. 3.19, $L_s = L_p + L_{se}$ should satisfy the stability condition $L_s < R_d^2 C_d$, and the circuit shunted by C_e can be made stable.

As C_d grows linearly with the device area, the inequality for a stable NDC device can be written as

$$C_d^* > A_d L_s / R_d^{*2}, \tag{3.52}$$

where A_d is the device area, C_d^* and R_d^* are the capacitance and the conductance of the device per unit area, respectively. Therefore, to obtain stable devices the device area must be decreased.

3.3.4 High Frequency and Functional Applications

At present, millimetre-wave radiation (30-300 GHz) is used in high resolution radar, radio astronomy, remote sensing, atmospheric and environmental monitoring, with emerging applications in imaging, surveillance, and telecommunications because of the capacity to carry large amounts of information, either directly by free space transmission or via an optical carrier [3]. Limited atmospheric transmission at these frequencies can be turned to advantage to provide security (in, for example, office local area networks, LANs) and minimize interference between neighbouring networks [3].

One of the key issues in the exploitation of millimetre-wave frequencies is the development of compact, efficient solid-state sources for free space transmission and to act as local oscillators (LO) for heterodyne detection [3]. The specification of power levels required from such a source depends, of course, on many factors such as the noise properties of the associated detection circuitry and the precise nature of the application; additional characteristics required are that they should operate at room temperature and be amenable to integration with a transmission line system, such as waveguide structures and printed lines.

Signal Generation [2]-[6]

A single-port device that has a negative differential conductance in a portion of its operating range may be used as the basis of a bistable or multistable circuit. Such a device may also be used to form astable circuits (relaxation oscillators), monostable circuits (single-pulse generators), and sine-wave generators. Devices manifesting negative differential conductance, $G_d = R_d^{-1}$, have an impedance, $Z_d(f) = R_d(f) + jX_d(f)$, whose real part is negative, $R_d(f) < 0$, $f < f_{max}$. Any signal of frequency f, travelling towards these devices along a transmission line of impedance Z_0, will present a reflection coefficient, S_{11},[17] given by [10]

$$S_{11} = \frac{Z_d - Z_0}{Z_d + Z_0} = \frac{-Z_0 - |R_d| + jX_d}{Z_0 - |R_d| + jX_d}, \tag{3.53}$$

which has a magnitude greater than unity, meaning that the incident signal is amplified; the highest amplification is achieved when $|R_d| = Z_0$. At this point, this amplifier is, in general, unstable and can oscillate spontaneously, depending on existing parasitic reactances. The RTD can be operated as an oscillator when coupled to a resonant tank circuit or a resonant cavity that provides frequency stability (the coupling location in the cavity can serve to partially match its impedance to that of the RTD.) An amplifier of this type is referred to as a *reflection-type amplifier*, with the input and output signals being separated using a circulator.[18]

The power obtainable from a RTD is proportional to the product $\Delta J \times \Delta V$, where $\Delta J = J_p - J_v$ and $\Delta V = V_v - V_p$ are the current density and voltage differences in the range of the NDC region, respectively. Driven by this rule, an effort has been made to increase the peak current density and decrease the device capacitance while maintaining a reasonable PVCR. Pure RTD oscillations are intrinsically low power (<1 μW at 712 GHz [6]) due to low dc bias, small area, and impedance limitations associated with the parasitic elements [2]-[5]. The general shape of the power versus frequency curve for a RTD follows the $1/f^2$ rule, which comes from considerations of dynamics of carriers in a transit-time device and the maximum electric field that can be sustained before breakdown [3]. For a given geometry, RTD microwave/millimeter-wave output power, which requires large current and voltage swing in the active part of the I-V characteristic, is material dependent. From the previous discussion on materials, it is clear that by adopting the InGaAlAs/InP material system higher output powers could be achieved due to the larger ΔJ available; ΔV, which depends mainly on the device structure and doping concentrations, may be slightly improved for a given device by insertion of a moderately doped collector spacer layer. (A larger difference between peak and valley currents, ΔJ, can also be achieved through device design.)

Device parasitics (capacitances, inductances and resistances) - from parts of the device which

[17] Representing the amplitude of an incident signal by a_1 and the amplitude of the reflected signal by b_1, the S_{11} parameter (more commonly known as the reflection coefficient, r) is defined by $S_{11} = b_1/a_1$, with $|S_{11}|^2 = \left|\frac{b_1}{a_1}\right|^2$ being the input power reflection (the ratio of the reflected power to the incident power).

[18] A circulator is a three-port nonreciprocal device. Power incident on port 1 goes to port 2, power incident on port 2 goes to port 3, and power incident on port 3 goes to port 1; $S_{21} = S_{32} = S_{31} = 0$ [10].

do not give a functional contribution to its operation (substrate, contact layers and access contacts, for example) - act in series or in parallel with the device. Their overall effect is to reduce the bias voltage applied to the active regions.[19] It is also desirable to decrease the series resistance, R_s, and the circuitry resistance, R_c, as much as possible because the fraction of the generated power absorbed these resistances is proportional to $(R_s + R_c)$. The importance of these parasitics can be decreased by reducing the device area, with a limit imposed by application (since power scales down with device area), relevant fabrication technology and packaging constraints [20][21].[20] RTD high frequency operation requires small capacitance, small absolute value of the differential resistance R_d (i.e., high peak current density and large PVCR), and an even smaller series impedance R_s [2]-[11]. However, the low impedance of small area RTDs is very difficult to adapt to a useful load, like a waveguide or antenna (it requires elements of very small cross section and very small waveguide resonators connected to ordinary waveguides via gradually widening transformer waveguides) [3][178][179].

The use of a RTD to drive a superconducting mixer based radiometer at 200 GHz has been described [142]. Furthermore, a quasi-optic resonator containing a single RTD has been constructed with excellent stability, linewidth and power-output levels, for use up to 200 GHz [149]. It may yet prove its practical value in high speed logic and certain niche analogue applications, where its low power consumption and versatility can be exploited.

Relaxation Oscillation [143][158]

In practice, bias oscillations occur very easily and impose a serious power loss. Analysis of this parasitic effect indicates that there is incompatibility in the choice of almost all of the design parameters (e.g., peak current, device capacitance) to reach stability or achieve high power output. An entirely fresh approach to this problem has been taken by Brown and Parker [157], who pointed out that this apparent incompatibility may be resolved by adopting a *relaxation-oscillator* approach [158]. The argument is that a more viable alternative to the use of an RTD in the quasi-sinusoidal mode may be to operate it in the relaxation oscillator form, in view of the rich harmonic content of such an oscillator. Further development of self-oscillating multipliers (akin to mode-locked lasers) may also be a realistic possibility.

A relaxation oscillator is a system designed to have two unstable states, and operate by sequential transitions between these states. Typically, the time spent in each state is much longer that the transition time between states. Unlike the sinusoidal oscillator, the relaxation oscillator energy storage system is mainly capacitive or inductive. Typically, a voltage controlled NDC relaxation oscillator uses inductive storage. The useful properties of relaxation oscillators include their ability to generate non-sinusoidal waves with high harmonic content, susceptible to tuning and modulation by means of a direct voltage; in addition, the frequency of oscillation can be synchronized to a control signal with frequency approximating a multiple or submultiple

[19] Although careful layout and use of appropriate material systems may reduce such effects [23][30], they can never be eliminated.

[20] There will be a minimum value of C_d below which the circuit will become unstable.

of the natural frequency of oscillation of the circuit.

In RTD relaxation oscillators, the circuit values are chosen so that the load line intersects the current-voltage characteristic only in the NDC range. The ac load condition is simply to have an energy storage element that can deliver enough energy back to the RTD to induce repetitive switching. Figure 3.20 presents the implementation scheme of the resonant tunnelling diode transmission line relaxation oscillator [157]. The RTD dwells in either of its positive differential conductance regions between switching events, with the repetition rate determined by the electrical delay time associated to the transmission line [157]. The power-output spectrum is in general not sinusoidal for this mode of operation; initial design studies actuallty revealed a very rich harmonic spectrum, with useful harmonic power up to at least 500 GHz [157].

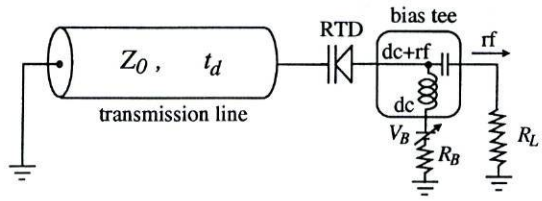


Figure 3.20: Resonant tunnelling transmission line relaxation oscillator schematic [157].

Figure 3.21 shows experimental relaxation oscillations of a RTD mounted in a package incorporating a coplanar waveguide transmission line (CPW) and connected to a coaxial transmission line, for two lengths of coaxial cable (which basically corresponds to two different inductances L_{se}; see small-signal equivalent circuit of Fig. 3.17). The oscillation was monitored using a digital oscilloscope with probes placed across the external coaxial cable connection to the bias supply (nodes **3-0**, Fig. 3.17). The device is a AlAs/GaAs resonant tunnelling diode with AlAs barriers 1.4 nm thick (5 ML), a GaAs quantum well 7 nm wide (25 ML), and GaAs spacer layers on each side 500 nm wide and doped to 2×10^{16} cm^{-3}; the contacts regions were doped to 2×10^{18} cm^{-3}. A typical I-V characteristic of this device was shown in Fig. 3.5.

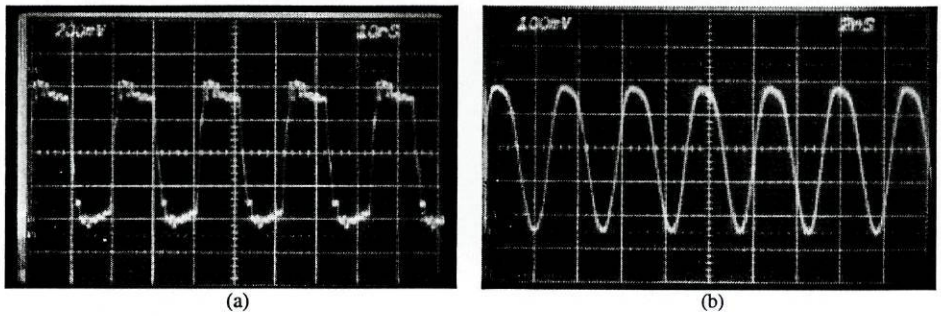


Figure 3.21: Experimental RTD relaxation oscillations for two different inductance values L_{se}: (a) for large and (b) smaller, respectively.

For the longer coaxial cable the circuit is in region I of Fig. 3.18; the voltage swing is shown in Fig. 3.21(a). Oscillation of devices in region I should have a growing exponential waveform, but in reality growth is limited by the extent of the NDC region. For shorter coaxial cable

lengths, the operating point moves to the left on the stability diagram of Fig. 3.18, into region II. The stability diagram predicts growing sinusoidal oscillations, which are also limited by the extent of the NDC region. If R_s is almost equal to R_d, the total resistance of the circuit will be negative only over a small region about the bias point, and so the voltage swing will be further constrained. This results in a nearly sinusoidal steady-state oscillation, as measured and shown in Fig. 3.21(b). In the growing sinusoidal region, decreasing the coaxial cable length (decreasing L_{se}) causes the frequency of oscillation ω_i, Eq. 3.37, to increase, as expected [141].

Frequency Multiplication [2][11]

There is another way to create high frequency signals using the non-linear characteristics of RTDs: *resistive multiplication.* The principle is similar to that of the Schottky diode multiplier: a large amplitude sinusoidal voltage signal is applied to the device and, due to the non-linear I-V characteristics, the resulting current through the device contains higher-order harmonics. Because RTDs can be designed to have symmetric characteristics, only odd harmonics can be generated, giving rise to an improvement in the conversion efficiency [147].

One of the advantages of the multiplier approach over a fundamental sinusoidal oscillator is that the output power of the multiplier can be larger than that for an oscillator at high frequencies because resistive multiplication does not necessarily require gain from the device, and so the frequency limit of f_{max} does not apply [141]. Therefore, supposing that a high power signal source is available and that the undesirable harmonics can be terminated successfully, a resistive multiplication circuit may serve as a more powerful signal source than a fundamental oscillator. Besides, in the NDC region of the I-V characteristic, the restriction to $1/n^2$ no longer applies for efficiency; the shape of the characteristic encourages higher order multiplication [147].

Switching [2][152][156]

Because of their NDC characteristic and extremely fast response, RTDs can also be used in high speed switching circuits. Figure 3.22 illustrates an equivalent circuit of a switching scheme using a RTD. The RTD equivalent circuit for large signals is used instead of the small-signal equivalent circuit ($-R_d$ is replaced by a current source $I_d(V_d)$, controlled by the voltage across the RTD), because the differential conductance of the RTD changes sign during switching; L_{qw} is not taken into account.

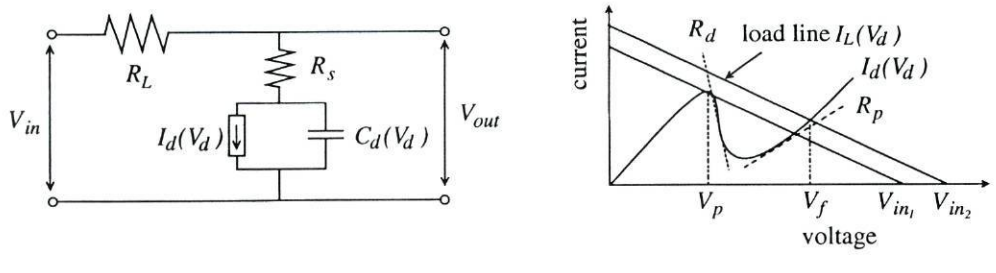


Figure 3.22: A switching circuit comprising a RTD and a load [152].

The basic circuit current equation is

$$C_d(V_d)\frac{dV_d}{dt} = I_L(V_d) - I_d(V_d). \tag{3.54}$$

The steady-state solutions are given by the intersection of the load line, $I_L(V_d)$, with the RTD line, $I_d(V_d)$. When the input voltage is $V_{in} = V_{in_1}$, the two lines intersect at the current peak of the RTD, giving two solutions: one at $V_d = V_p$ and the other at a higher voltage. When $V_{in} = V_{in_2}$ (a slightly higher input voltage) there is only one solution at $V_d = V_f$. Supposing the initial state is $V_{in} = V_{in_1}$ with $V_d = V_p$, and the input voltage is suddenly changed to V_{in_2}, the RTD voltage switches from V_p to V_f (the negative slope of the NDC region, $1/R_d$, must be steeper than the slope of $I = I_L(V_d)$, i.e., $1/(R_L + R_s)$, to allow switching).

The transient behaviour of a switching circuit using a RTD can be expressed in terms of the 10-90% rise time, T_{rise}, given by (Fig. 3.22)

$$\begin{aligned} T_{rise} &= \int_{V_p+0.1\Delta V}^{V_f-0.1\Delta V} \frac{C_d(V)}{I_L(V) - I_d(V)} dV \\ &= |R_d|\, C_d \left\{ \left(\frac{x}{x-1}\right) \ln\left[\frac{10(x+y)}{x(x-y)}\right] + \left(\frac{xy}{x-y}\right) \ln\left[\frac{10y(x+y)}{x(x-y)}\right] \right\}, \end{aligned} \tag{3.55}$$

where $x \equiv (R_L + R_s)/|R_d|$, $y \equiv R_p/|R_d|$, and $\Delta V = V_f - V_p$. R_d and C_d dominate the switching speed, and as for RTD oscillators, a reduction in diode capacitance C_d, an increase in NDC ($1/R_d$), and a reduction in R_s are needed for response time improvement. The smallest attainable transition time of a switching circuit employing a single RTD is shown to be $\sim 4R_dC_d$.

There have been some reports on measurement of the switching time of RTDs, with values as short as 1.7 ps reported, using photoconductive switch excitation and the electro-optic sampling technique [9][153]. RTD switches have been employed in trigger circuits [8]; a rise time of 6 ps and triggering frequency of 60 GHz have been achieved using GaAs/AlAs RTDs in discrete trigger circuits [9]. Further improvement has been achieved using monolithic integration of RTDs, resistors and coplanar transmission lines [8]. Two RTDs were employed in the circuit to adjust the trigger timing to the maximum slope of the incoming signal, reducing jitter (when compared to a single RTD), enabling operation at 100 GHz [8].

Functional Applications [2][11]

Various circuits using RTDs and related devices have been proposed to perform new functions, including binary inverters and memories, multi-valued logic elements and memories, signal processing and logic circuits.

Static Memories [26]

Using tunnel diodes or RTDs it is possible to construct an inverter circuit with hysteresis, which can be used as a static memory element with a greatly reduced number of components compared to conventional circuits [31][32]. The simplest configuration consists of a RTD and a resistor,

constituting a memory element as shown in Fig. 3.23.

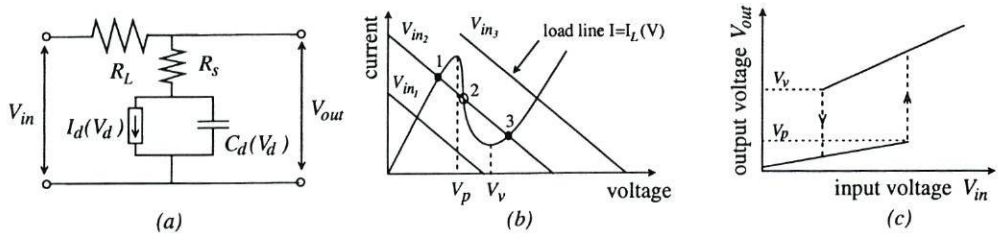


Figure 3.23: Memory circuit comprising a RTD and a resistor [2].

The circuit equation can be written as

$$C_d(V)\frac{dV}{dt} = \frac{V_{in} - V}{R_L} - \left(\frac{V}{R_i} - I_i\right), \tag{3.56}$$

where R_i is the differential resistance of the RTD at the intersection i ($i = 1, 2, 3$) and I_i is the corresponding current intersect. A steady-state solution is available at the intersection of the RTD I-V characteristic $I = I_d(V)$ and the load line $I = I_L(V)$.

Neglecting the voltage dependence of the RTD capacitance, the solution of Eq. 3.56 can be expressed as

$$V_{out} \propto \exp\left[-\frac{1}{C_d}\left(\frac{1}{R_L} + \frac{1}{R_i}\right)t\right] + \frac{R_L R_i}{R_l + R_i}\left(\frac{V_{in}}{R_L} - I_i\right). \tag{3.57}$$

The stability of V_{out} is determined by the sign of the exponent $(1/R_L + 1/R_i)$. In the case of Fig. 3.23(b), 1 and 3 are stable solutions; solution 2 is unstable since $(1/R_L + 1/R_i)$ must be negative for the two lines to have three intersection points.

Changing the states between the two stable points, 1 (high) and 3 (low), provides the static memory function, with operation as follows [Fig. 3.23(b)]: if V_{in} increases from zero, the RTD voltage changes along the first PDC part of $I_d(V)$, from zero through point 1 until it reaches V_p. Then, with a further increase in V_{in}, the RTD voltage switches to the second PDC part of $I_d(V)$. Decreasing V_{in} after the voltage jump, the RTD voltage decreases, passing through point 3, until it reaches V_v; a further decrease in V_{in} causes another step change in the RTD voltage, to the first PDC portion of the $I_d(V)$. Plotting the RTD voltage, or V_{out}, as a function of the input voltage V_{in}, results in the characteristic shown in Fig. 3.23(c), which shows hysteresis whose V_{in} width determines the noise margin of the memory.

Multi-State Memories [26][32]

Multi-state memories can be realized employing RTDs with multiple peak I-V characteristics, which can be obtained by horizontal (parallel) or vertical (series) integration of RTDs [11]. Figure 3.24 shows the experimental I-V characteristic of two RTDs in series, at room temperature. When integrated with an appropriate circuit, the two RTDs can be used as a memory element in 3-state logic systems. The diagram inset in Fig. 3.24 gives an example of a possible three-state memory element employing two vertically integrated RTDs: a suitable load line drawn on the

I-V characteristic of k (=2) RTDs with k peaks will intersect the RTDs I-V curve at $k+1$(=3) stable points in the PDC regions.

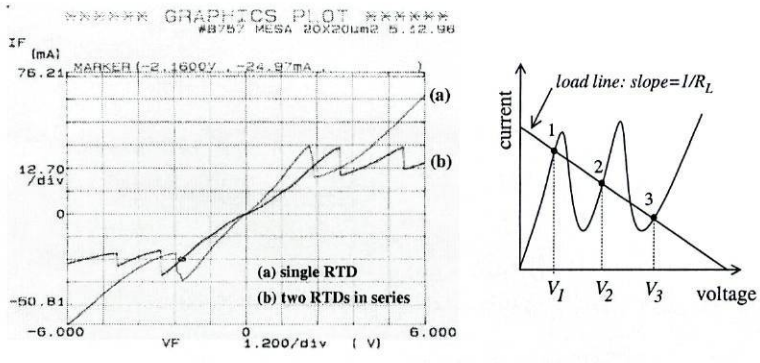


Figure 3.24: Experimetal current-voltage characteristic of two vertically integrated RTDs.

The memory elements discussed are simple and suitable for integration with other active devices, such as FETs, in memory integrated circuits; incorporation of RTDs leads to a reduction on the number of devices needed, and to a decrease in the occupied area.

3.4 Optoelectronic Applications of RTD Structures

Several prominent optoelectronic applications based on integration of resonant tunnelling structures with optical devices, where there is direct optical interaction with tunnelling layers, have also been proposed. Optoelectronic applications of resonant tunnelling devices can be divided in two areas, optical and infrared, depending on the nature of the transitions: the optical applications utilize interband transitions (bandgap transitions), whereas the infrared region employs intersubband transitions (transitions within the same band). Optical applications include photodetection [37], light emission [161][164], optical switching [43][44], whereas infrared applications comprise intraband and intersubband photodetection [19] and infrared emission [28][29]. In this section, these optoelectronic applications of resonant tunnelling devices will be briefly discussed.

3.4.1 Sources and Photo-Detectors

Recent work has demonstrated the appealing application of double barrier resonant tunnelling structures as electroluminescence devices, either by impact ionization in unipolar n-type structures [160] or by injection electroluminescence in resonant tunnelling structures grown in the intrinsic region of a p-i-n diode [161].

Resonant Tunnelling LEDs [161][162]

Resonant tunnelling of both electrons and holes can be achieved in forward biased p-i-$DBQW$-i-n heterostructures, giving rise to pronounced electron- and hole-related resonant peaks in the

current-voltage characteristics. These *p-i-DBQW-i-n* diodes also show strong electroluminescence emission, due to electron-hole recombination in the quantum well and, in addition, recombination of electrons (holes) which tunnel through both barriers and recombine, as minority carriers, in the *p*-type (*n*-type) contact layer. Its intensity is strongly dependent on bias, and peaks at voltages corresponding to electron and hole tunnelling [163].

Recently, *p-i-DBQW-i-n* light emitting diodes have shown attractive features when high speed and narrow linewidth are required, but coherent light is not needed [14][15]. High speed LEDs based on resonant tunnelling (RT-LEDs) take advantage of the fast tunnelling escape time from the quantum well, with a switch-off time not limited by radiative recombination. Moreover, RT-LEDs negative differential conductance allows their use as optical oscillators (e.g., as optical clock). Laser transistors incorporating a resonant tunnelling structure have also been reported, with carrier injection or extraction controlled via the resonant tunnelling structure [164]. The light output is expected to be efficiently controlled by the collector voltage, achieving higher speed than with conventional semiconductor lasers; bistability operation is also foreseen [164].

Resonant Tunnelling Photo-Detectors [37][39]

A light pulse incident upon a resonant tunnelling diode produces photocharges that reduce the series resistance, leading to a shift of the peak and valley voltages which can induce RTD switching and give rise to changes in the current flow [37][38]. Assuming the RTD is dc biased slightly below the peak, interband absorption of photons in the depletion region adjacent to the double barriers generates a voltage with opposite polarity to the applied voltage, thus reducing slightly the peak and valley voltages. This can induce the operation point to switch to the second PDC regions of the I-V curve [38]. The NDC and hence the current flow are then controlled by the incident optical power, making photodetection possible. By taking advantage of these effects and of the NDC characteristic intrinsic gain, an optically switched resonant tunnelling diode (ORTD) photodetector, Fig. 3.25, has been demonstrated [39]. The ORTD exhibits responsivity and gain-bandwidth-efficiency product in excess of 10 A/W and 10 GHz·A/W,[21] respectively, low switching energy (30 fJ), with bit-error rates of less than 10^{-9} at 2 Gb/s. Phase locking of an oscillating GaAs/AlGaAs RTD to a train of light pulses achieved by direct illumination, was also reported [169]. These devices could be easily integrated with conventional electronic elements.

Two mechanisms for RTD switching induced by an optical pulse have been proposed [165][166]: one is a photocurrent mechanism, and the other is a charge accumulation mechanism. When the RTD is illuminated, the overall I-V curve is shifted to lower voltages, because the resistance is reduced due to the photocurrent and also because of the potential profile due to charge accumulation in the depletion layer, which causes band bending. Photogenerated electrons and holes contribute to the reduction of series resistance by increasing the photocurrent, and modify the

[21] To date, the best gain-bandwidth-efficiency product values, ≈100 GHz·A/W, have been obtained using advanced avalanche photodiode designs [168].

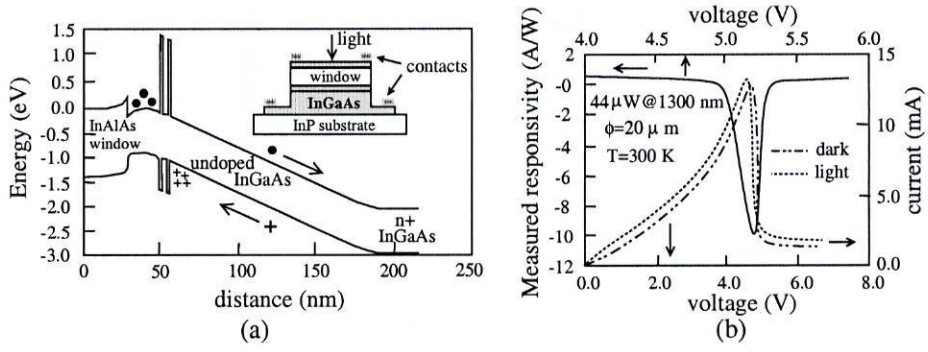


Figure 3.25: (a) Conduction band diagram for a typical ORTD. Schematic cross-section of the device is shown in the inset. (b) Measured dark and illuminated I-V characteristic and measured responsivity (I_{light}-I_{dark}) as function of the bias [39].

electric field around the DBQW structure by accumulating charges near the barriers (screening mechanism) [167]. The holes photogenerated in the depletion region accumulate around the collector barrier and cause a local enhancement of the electric field, which means that a lower voltage is required to switch the device (electrons are less important in this regard since they are rapidly swept out of the active region) [165]. The current density due to hole tunnelling is mainly due to light holes because of their lower mass with respect to heavy holes. It is believed that the photogenerated hole accumulation effect in the depletion region, which results in the hole tunnelling current, is the dominant cause for the shift in the peak voltage, with increasing optical power [165].

Recently, optical control over the resonant tunnelling characteristics of a RTD monolithically integrated with a *pin* photodiode was demonstrated at 1 GHz [170]; optical switching of the bistable resonant tunnelling state and optical injection locking of the resonant tunnelling oscillator were both controlled. An ultra-fast optoelectronic circuit, using resonant tunnelling diodes and a uni-travelling-carrier photodiode (UTC-PD), that can demultiplex an ultra-fast optical data signal into an electrical data signal with lower bit rate and low power consumption [171], has been also demonstrated; this monolithic device demultiplexed a 80 Gbit/s optical signal into a 40 Gbit/s electrical signal.

3.4.2 Optical Bistability

Optical bistable semiconductor devices are very promising as optical logic elements for optical computing and processing systems. The requirements for these applications include low switching energy, well-defined output logic levels, large on/off ratio, high fan-out, and fast response [15][19]; the potential of III-V quantum well based devices for these applications is well recognized.

Bistable semiconductor lasers have the important advantage of possessing inherent large optical gains and steep light/voltage characteristics, or high quantum efficiencies, thus allowing a high ratio between the output intensities of two logic states [18][24]. They use non-linear optical effects such as saturable absorption [13], or saturation-induced refractive index change

[13][24] in conjunction with the laser cavity to induce a bistable mode of operation. The use of the quantum confined Stark effect to realize self-electro-optic effect devices (SEED) has been also exploited to implement bistable devices [13][19].

MQW based RTDs have a NDC with a wide valley region, which is very effective for getting large voltage switching and high on/off ratio current switching [34][172]. When a RTD operation point switches, the optical absorption of the depleted semiconductor changes. This suggests that RTDs can also be used as bistable electro-optic switches; furthermore, it is possible to obtain light-by-light switching with RTDs [34][37]. Optical bistable MQW lasers in which the MQW layer gives rise to NDC have been demonstrated [172]. Optical bistability in a QW laser integrated with a DBQW-RTD, and in a RTD with a MQW modulator/detector based on the p-i(MQW)-n configuration, operating at room temperature, were also reported [34][35][36][173]. However, the properties of this kind of bistable devices are not yet satisfactory.

Clear negative differential conductance and bistability, with high contrast and high sensitivity, in a resonant tunnelling triangular barrier optoelectronic switch (R-TOPS), which consists of a double barrier resonant tunnelling diode and a triangular barrier phototransistor, has been demonstrated [174]. Bistability in the light output of bipolar RT-LEDs has been also reported, showing that this device is capable of ultrafast optical switching and high speed optical oscillation [175]. By achieving electrostatic feedback due to hole accumulation in an asymmetric triple-barrier resonant tunneling light emitting diode, inverted bistability was demonstrated in the current-voltage characteristic [175].

The bistable characteristic using the RT effect is essentially determined by the electrical properties, contrary to the "ordinary" optical bistability which relies on optical non-linearities [13][19]. Because the NDC characteristic is controlled by the parameters of the quantum well structures (such as barrier and well widths), very predictable bistable operation is possible [173]. From the standpoint of an electro-optic device, using a RTD as an optical switching element has some attractive features which distinguish it from, for example, the SEED. In the case of RTDs, there is a connection between the electrical and optical responses (in the case of SEEDs, the optically active region is a *p-i(MQW)-n* junction which is electrically inert in the absence of light) [13][19]. The intrinsic bistability of the RTD implies that the device will remain in the switched or unswitched configuration while electrical power is applied (the RTD acts as a static memory element, while the SEED is a dynamic memory cell - needs refreshing [19]).

3.4.3 Intersubband Infrared Detectors

The optical properties of superlattices and quantum wells have attracted much attention in recent years due to the large oscillator strength for optical transitions between quantum well subbands (levels) [12][22][24].

In a superlattice, the barriers thickness becomes comparable to the carrier de Broglie wavelength, so the wavefunctions of the individual wells tend to overlap due to tunnelling, and energy minibands of width $2\Delta E_n$ are formed [28], Fig. 2.20. The width $2\Delta E_n$ is proportional to the tun-

nelling probability through the barriers and, for rectangular barriers and not too strong coupling between wells, can be approximated by (Eq. 3.10)

$$\Delta E_n \approx E_n \exp\left[-2L_b\sqrt{\frac{2m_b\left(U_0 - E_n\right)}{\hbar^2}}\right]. \tag{3.58}$$

Under certain bias conditions, the ground state in the kth well becomes degenerate with the first excited state in the $(k+1)$th well, and electron transport across the structure results mainly from sequential resonant tunnelling: an electron from the ground state at the kth well tunnels to the vacant excited state at the $(k+1)$th well, followed by nonradiative relaxation to the ground state at the $(k+1)$th well. Because of the resonant nature of this process, its probability is high. This gives rise to a peak in the current at fields corresponding to the above degeneracy, as described by Kazarinov and Suris [28]. When appropriately biased, Fig. 3.26, these superlattices can be employed as infrared detectors with high wavelength selectivity [11][12][53][64].

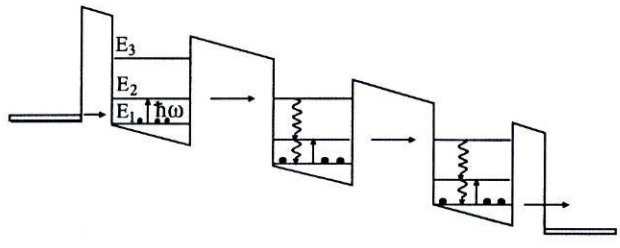


Figure 3.26: Band-diagram of a sequential resonant tunnelling intersubband detector [28].

The bias across the device should be such that the second level in the kth well is degenerate with the third level in the $(k+1)$th well. Infrared radiation incident on the device, having a component of the electric field perpendicular to the plane of the well, will be absorbed strongly only if the photon energy is equal or close to $(E_2 - E_1)$, because of momentum conservation considerations. Once electrons have made the optical transition to the first excited state, they can resonantly tunnel to the bottom of the second excited state of the adjacent well. From here, the most likely route is intersubband scattering to the $n = 2$ level, followed by either intraband relaxation and resonant tunnelling to the adjacent well, or scattering to the $n = 1$ subband with final relaxation to the bottom of the well. The net effect of the absorption, tunnelling, and relaxation processes is a photocurrent.

3.4.4 Intersubband Infrared Lasers

Conventional semiconductor lasers rely on radiative recombination of electrons from the conduction band with holes in the valence band, injected into the active layer through a forward biased *pn* junction; the material bandgap essentially determines the emission wavelength, and population inversion is created by injecting electrons in the conduction band and holes in the valence band [18]. There have been several suggestions to create population inversion between

minibands - a necessary condition for laser action [28][29].

Kazarinov-Suris Intersubband Infrared Lasers [12][28]

As discussed before, under certain bias conditions, the bottom of the first excited quantum subband in the $(k+1)$th well of the superlattices lies below the bottom of the ground state subband in the kth well, Fig. 3.27(a). Amplification (or laser action) of radiation with frequency $v = [e\mathcal{E}d - (E_2 - E_1)]/h$ (where d is the superlattice period), whose electric field is parallel to the external static field $\mathcal{E}$, is then possible. This occurs via a photon-assisted tunnelling transition whereby an electron from the ground state in the kth well tunnels to an excited state in the $(k+1)$th well with simultaneous emission of a photon of frequency v. The amplified frequency (or laser emission frequency) can be varied over a wide range by varying the electric field. Kazarinov and Suris proposed this new type of laser in 1972.

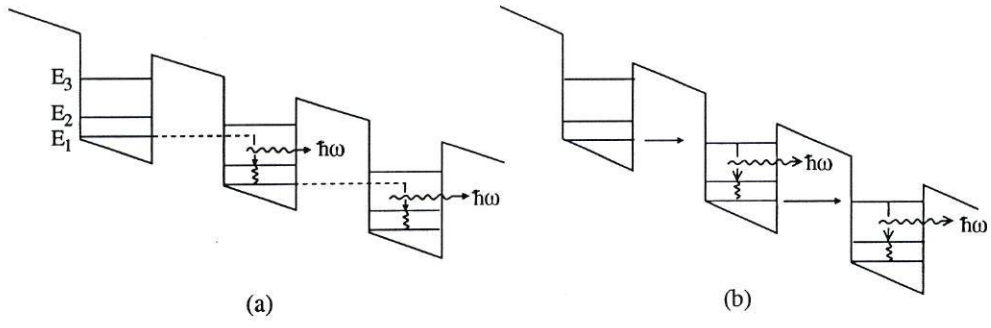


Figure 3.27: Far infrared laser band-diagram based on sequential resonant tunnelling [28]. (a) Interwell photon assisted tunnelling transition. (b) Photon emission during an intrawell transition, followed by resonant tunnelling between wells.

Another possibility of radiation amplification (or laser action) occurs when the ground level of the kth well is in resonance with the second excited state of the $(k+1)$th well, Fig. 3.27(b). The second excited state is populated by tunnelling, and therefore stimulated emission is now a direct intrawell transition from the second to the first excited state of the wells. To achieve laser amplification at frequency $(E_3 - E_2)/h$, there must be population inversion between states 3 and 2. This requires the lifetime of the electron at the bottom of the third subband to be larger than the electron lifetime at the bottom of the second subband. These lifetimes are primarily controlled by intersubband scattering by optical phonons if the intersubband separation is greater than the optical phonon energy.[22] Population inversion between states 3 and 2 is not achieved unless the separation between the two subbands is chosen smaller than the optical phonon energy; the overwhelming contribution of nonradiative phonon relaxation in these structures makes efficient infrared light amplification impossible.

[22]The typical value of the optical phonon emission time τ_0 is on the order of 1 ps (the most important mechanism of electron relaxation); the acoustic phonon relaxation time is $\sim 3\times10^{-10}$ s, while the interband radiative time is $\sim 10^{-7}$ s.

Quantum Cascade Lasers [29][42]

In all the cases mentioned above, sequential resonant tunnelling is the basic mechanism of electron transport in the superlattice. This process involves electron relaxation between the two levels in the QW with subsequent resonant tunnelling into the excited level of the adjacent QW. These structures, however, when designed for a practical realization of population inversion, encounter serious difficulties related to the necessity to provide both high injection current density and quickly remove electrons from the lower level of the QW (within a time shorter than the optical phonon emission time).

In spite of such drawback, F. Capasso [29] proposed unipolar mid infrared LEDs and lasers which utilize intersubband transitions in semiconductor quantum wells; these were first demonstrated by J. Faist *et al* [42] in 1994. The device structure employed InGaAs/InAlAs heterojunctions lattice matched to the InP substrate [29]. The laser transition occurs between the first two excited levels of the first well in a triple barrier structure, Fig. 3.28, with an energy difference $E_3 - E_2 \approx 295$ meV. Relaxation between the first excited state and the ground state of the first QW is very fast because their energy separation is equal or close to an optical phonon energy ($E_1 - E_0$=30 meV) so that scattering occurs with essentially zero momentum transfer. Tunnelling out of the ground state of the second well is made very fast (less than 0.5 ps) by employing a thin barrier ($\approx$3 nm). Because of rapid non-radiative scattering between subbands by LO phonon emission, the quantum efficiency is low, so many different active region blocks are "cascaded" together to "recycle" the injected electrons. The successive active regions are linked by doped "graded-gap" superlattices, which allow electrons to thermalize so they are injected "cold" into the next active region, and also increase confinement of the upper lasing electron state, which enhances population inversion. The structure doping profile is chosen to ensure overall charge neutrality within each period, preventing formation of high field domains and keeping a uniform potential distribution throughout the structure, and also give good transport properties without excessive free carrier absorption (very important at mid infrared wavelengths and beyond). Slightly different structures have been designed to operate at different wavelengths [29].

Quantum cascade LEDs and lasers use band structure engineering and tunnelling, with only one type of carrier, usually electrons, being responsible for the emission (they are unipolar devices). The joint density of states of these transitions and the corresponding gain spectrum are narrower than in conventional interband transition devices, and are essentially symmetric. The emission wavelength of such lasers is defined mainly by the triple barrier QW structure; for a given heterostructure material, it can be tailored over a wide spectral range, from the mid-infrared ($\sim$4 μm) to the submillimetre wave region ($\sim$100 μm). A vast range of materials can be used, including relatively wide bandgap ones, such as in GaAs- and InP-based heterostructures. (Similar considerations apply to infra-red photodetectors employing quantum well inter-subband transitions [11][12][53][64].)

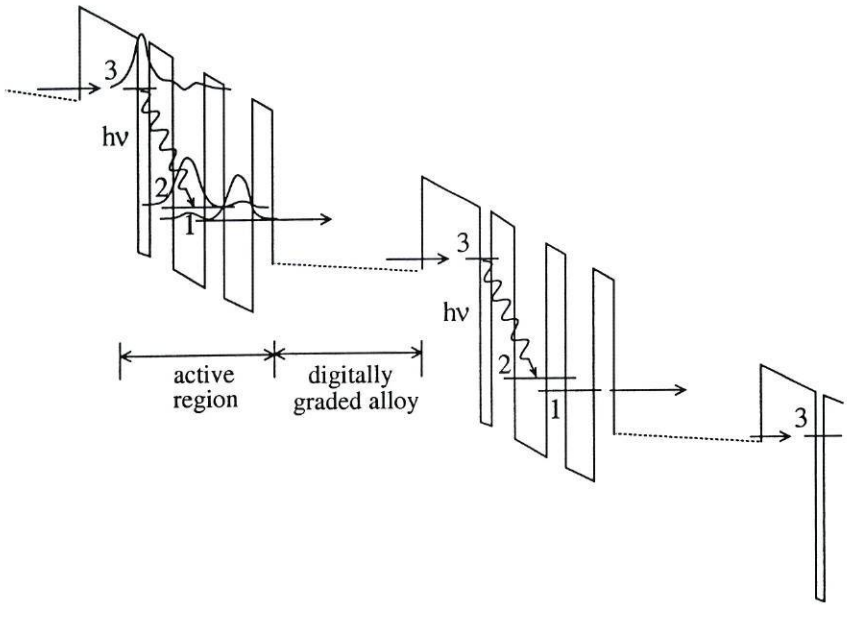


Figure 3.28: QC laser: photon assisted tunnelling "diagonal" transition [29].

3.5 RTD Electro-Absorption Modulator

Novel information and communication technologies relying on microwave/millimetre-wave - lightwave interactions are fundamental to the development of fibre-optic wireless communication networks, mobile communications, cable television signal distribution, radio local area networks, and phased array antennas [75].

In this section, a new electrically active, high speed, highly efficient and low cost microwave electro-absorption modulator concept, based on the integration of a resonant tunnelling diode within an optical waveguide, is proposed and discussed.

3.5.1 RTD-Optical Waveguide Integration Concept

Interband absorption in the depletion region adjacent to the collector barrier of a DBQW-RTD, at photon energies close to but smaller than the collector bandgap energy, can be induced by the electric field across the RTD through the Franz-Keldysh effect. In a typical RTD structure the light is injected perpendicularly to the tunnelling plane, which gives a very small interaction (absorption) length, in general less than 100 nm. This limitation can be overcome by coupling the optical radiation in a direction parallel to the tunnelling plane and efficiently confining the light around the depletion region, therefore increasing the interaction length . That can be easily achieved through embedding of the DBQW-RTD into an optical semiconductor waveguide core [41], the resonant tunnelling diode optical waveguide (RTD-OW), as represented schematically in Fig. 3.29. The RTD-OW is to be implemented in a ridged channel unipolar waveguide configuration (lying on top of the substrate, Fig. 2.22).

The RTD-OW wafer structure, Fig. 3.29, apart from the waveguide confining layers (the

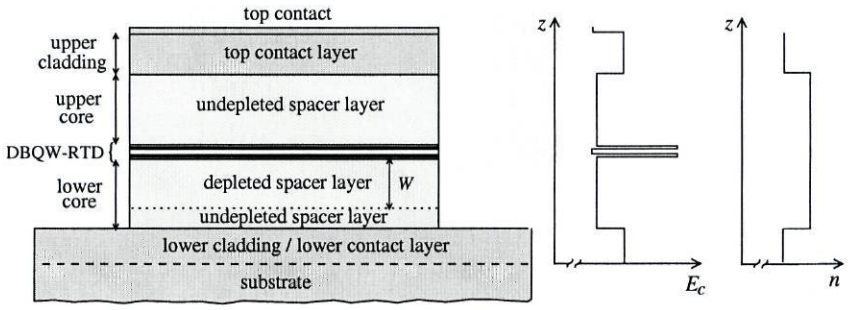


Figure 3.29: Diagram of the RTD-OW wafer structure, Γ-conduction bandedge and refractive index profiles.

lower refractive index regions - upper and lower cladding layers), corresponds to a DBQW-RTD with very thick low doped emitter and collector spacer layers (the core region). The presence of the DBQW within the core changes the linear current-voltage characteristic of the uniploar waveguide: the strong non-linearity of the DBQW-RTD I-V curve leads to a non-linear electric field distribution across the waveguide core (strongly dependent on the bias voltage), due to electron accumulation close to the emitter barrier and creation of a depletion region on the collector spacer layer. As a consequence, a small increase in the peak voltage can give rise to a large change in the electric field across the depletion region, resulting, through the Franz-Keldysh effect, in changes of the waveguide transmission spectrum, as discussed in Chapter 2. This effect can be used to implement either electro-absorption (intensity) or electro-refraction (phase) waveguide modulators. The RTD-OW structures were employed mainly to study the electro-absorption effect in the waveguide, induced by the DBQW-RTD switching.

The DBQW-RTD-OW is designed such that its I-V characteristic shows considerable negative differential conductance, with a significant portion of the waveguide core being depleted when the device is biased at a voltage higher than the peak voltage. This imposes structural limitations both on the DBQW and on the waveguide structures (Fig. 3.29): the core thickness must be sufficiently large, allowing at least one transverse optical mode to be guided, and its doping level must be low to induce small optical losses; however, doping across the structure, specially in the cladding layers, has to be kept as high as possible to minimise the device series resistance. A compromise between the required NDC and the optical properties is needed.

The structures investigated in this research were grown by molecular beam epitaxy in a Varian Gen II system, using either the GaAs/AlGaAs or the InGaAlAs/InP material systems. In the wafer structures used throughout this work, the DBQW-RTD is located in the centre of the waveguide core.

3.5.2 Electro-Absorption Effect in RTD-OWs

According to the Franz-Keldysh effect, the waveguide optical absorption bandedge is broadened by the presence of an electric field, resulting in changes of the optical transmission characteristics at photon energies close to the bandgap [76][77][78]. Under certain bias conditions, a small high

frequency ac signal (<1 V) can switch the RTD-OW operation point well into the two positive differential conductance (PDC) portions of the current-voltage characteristic, with a substantial part of the terminal voltage dropped across the depleted collector region. This corresponds to high speed electric field switching, resulting in high frequency optical waveguide transmission modulation; the RTD-OW is then operating as a *resonant tunnelling diode electro-absorption modulator* (RTD-EAM).

The RTD-OW electric field distribution dependence on the bias voltage can be understood considering the Γ-conduction band profile of the collector spacer layer, Figure 3.30. Below resonance (first PDC region), the applied voltage is dropped mainly across the DBQW, and the electric field in the collector core is rather small. Any optical loss increase with the applied voltage is mainly due to the thermal effects induced by the current flow, which rise with the current.

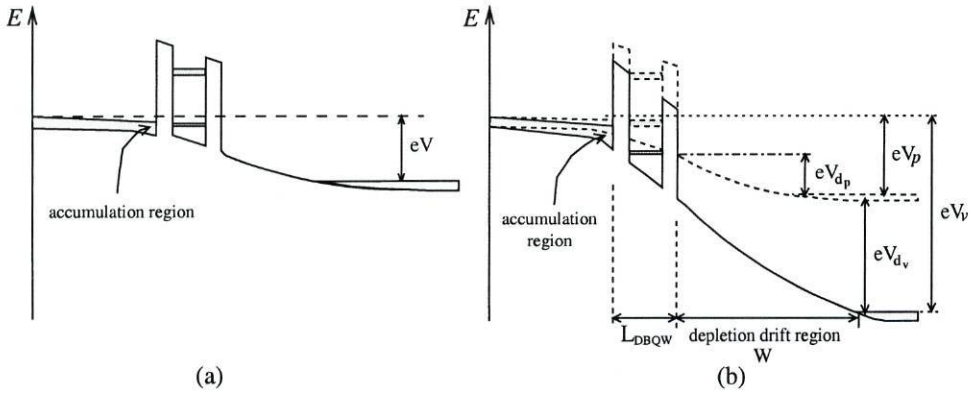


Figure 3.30: Energy band diagram of a RTD with applied bias.

Above resonance (second PDC region), the additional applied bias voltage is dropped mainly across the depleted part of the collector spacer layer, and the electric field magnitude is now much stronger than in the first PDC case, inducing large optical absorption. The thermal optical absorption is less important now because the current is lower. For dc bias in the NDC region, the waveguide electric field distribution can be highly unstable, due to current and voltage instabilities induced by the RTD electrical gain. This is a consequence of the dependence on the bias voltage of the mismatch/match between the electrons energy distribution on the emitter side and the energy levels distribution in the quantum well. When biased in the NDC region, the RTD-EAM can be used to build a *self-oscillation* optical modulator, taking advantage of the RTD electrical gain: incorporating the RTD-EAM in an appropriately designed electrical resonant cavity, the electric field across the depleted collector region can self-oscillate at the resonant cavity characteristic frequency, thus self-modulating the transmission properties of the waveguide.

Figure 3.31 represents schematically the Franz-Keldysh effect in a RTD-EAM biased around the valley point. The modulation depth depends mainly on the overlap between the electric field in the collector depleted region and the optical mode, and on the electric field magnitude.

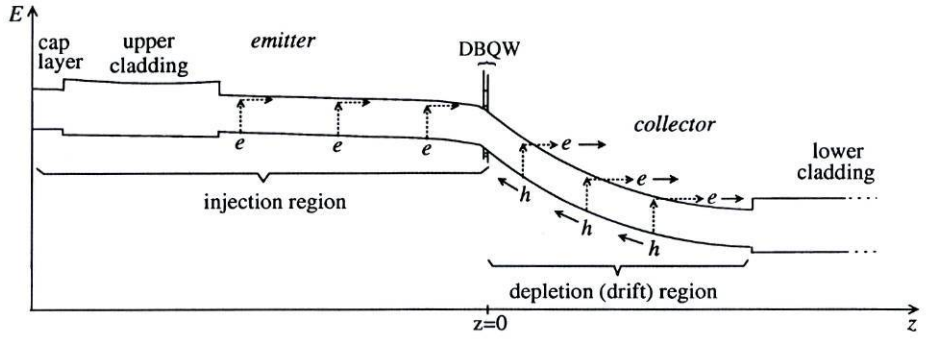


Figure 3.31: Schematic representation of the Franz-Keldysh effect in a RTD-EAM.

Electric Field Enhancement Induced by the RTD Peak-to-Valley Switching

The general approach to determine the magnitude of the electric field change in the depletion region due to peak-to-valley switching follows the well-established theory of conventional transit-time devices [137]. In fact, RTDs with thick spacer layers have many similarities to IMPATT devices: the tunnelling current through the DBQW acts as an *injection source* to the depletion region, analogous to the avalanche region in the IMPATT. There is, however, an important difference relative to avalanche injection, since this essentially acts as an inductance [137]; the tunnelling current responds so fast to the instantaneous voltage across the DBQW that it can be modelled by a pure negative conductance. Resistive injection occurs in yet another device, the BARITT [49][137], which employs a forward-biased p-n junction for injection, giving a positive injection conductance, σ, in parallel with the injection region capacitance.[23]

To determine the magnitude of the electric field change in the depleted core induced by peak-to-valley switching, the injector formed by the emitter and the DBQW, characterized by a negative injection conductance, σ, can be decoupled from the depletion region, on the collector side, since the injector characteristics should not depend strongly on the collector spacer layer design [63][134]. As a first approximation, the electric field in the depletion (drift) region is assumed to be high enough to cause the injected electrons to traverse it at a constant velocity, v; doping in the collector spacer layer (waveguide core doping) is assumed to be such that it is fully depleted at the peak voltage bias.

The magnitude of the electric field in the drift region as function of position, $\mathcal{E}(z)$, at constant current density J can be obtained from the Poisson's equation, Eq. 2.30, $\frac{\partial \mathcal{E}}{\partial z} = \frac{e}{\epsilon}\left(N_d^+ - \frac{J}{ev}\right)$, as:

$$\mathcal{E}(z) = \mathcal{E}_{0,J} - \frac{e}{\epsilon}\left(N_d^+ - \frac{J}{ev}\right) z, \tag{3.59}$$

where $\mathcal{E}_{0,J}$ represents the electric field at the injection plane $z = 0$ (the boundary between the injector and the drift region) at current density J. The change in voltage across the drift region

[23] This type of high frequency operation is completely different from that of transferred electron or Gunn devices [49][137]: in the latter the velocity of electrons is reduced by an increasing field once the field is larger than a certain value (Section 2.1.3, Fig. 2.3), whereas in the RTD the number of electrons participating in the current is reduced by breaking the resonance condition with an increase in the applied voltage.

due to peak-to-valley switching is given by:

$$\Delta V_d = \int_0^W \left[\mathcal{E}(J_v; z) - \mathcal{E}(J_p; z)\right] dz = \Delta\mathcal{E}_0 W + \frac{W^2}{2\epsilon v}\left(J_v - J_p\right), \tag{3.60}$$

where $\Delta\mathcal{E}_0 = \mathcal{E}_{0,J_v} - \mathcal{E}_{0,J_p}$ is the change of the electric field at $z = 0$, between valley and peak; W is assumed to remain the same before and after the current switching. The electric field change, $\Delta\mathcal{E}_0$, at the boundary between the injector and the drift region due to the change ΔJ_{p-v} in the current density flowing through the device is then given by:

$$\Delta\mathcal{E}_0 = \frac{\Delta V_d}{W} + \frac{W}{2\epsilon v}\Delta J_{p-v}. \tag{3.61}$$

This change in the voltage ΔV_d across the drift region can be estimated considering the current-voltage characteristic of the device [41], assuming the peak series resistance, $R_s = \frac{V_p}{I_p}$, does not change due to the current drop:

$$\Delta V_d \simeq R_s(I_p - I_v) = V_p\left(1 - \frac{I_v}{I_p}\right). \tag{3.62}$$

A current decrease ΔJ_{p-v} from the peak to the valley results in an increase of the electric field across the depletion region. Thus, an increment $\Delta V_{v-p} = V_v - V_p$ in the bias around the peak voltage gives rise to a large change in the electric field. As long as the region is fully depleted, and as long as the field inhomogeneity due to space charge or carrier screening and the heterojunction interfaces is negligible, the electric field across the depleted region may be considered, in a first order approximation, to be uniform and given by Eq. 3.61.

Effective Bandedge Shift Induced by the Electric Field Enhancement

The effective absorption bandedge shift to longer wavelengths, $\Delta\lambda_g$, is:

$$\Delta\lambda_g \simeq \frac{\lambda_g^2}{hc}\left(\frac{eh}{8\pi^2 m_r}\right)^{2/3}\mathcal{E}^{2/3}, \tag{3.63}$$

where m_r is the electron-hole system effective reduced mass ($m_r^{-1} = m_e^{-1} + m_h^{-1}$).

Absorption Change Induced by the Electric Field Enhancement

The Franz-Keldysh electro-absorption coefficient, Eq. 2.105, in the weak-field approximation is given by [176]:

$$\alpha\left(\hbar\omega;\, \mathcal{E}\right) = \sum_j A_j \mathcal{E}^{1/3}\left[\left|\left(\frac{d\mathrm{Ai}(z)}{dz}\right)_{\beta_j}\right|^2 - \beta_j\left|\mathrm{Ai}(\beta_j)\right|^2\right], \tag{3.64}$$

where $\mathcal{E}$ is the electric field in V/cm, $\beta_j = B_j\left(E_g - \hbar\omega\right)\mathcal{E}^{-2/3}$, $A_j = \frac{7.65\times10^5}{n\hbar\omega}\left(2m_{r_j}/m_0\right)^{4/3}$, $B_j = 1.1\times10^5\left(2m_{r_j}/m_0\right)^{1/3}$, E_g and $\hbar\omega$ are the bandgap and photon energies in eV, respectively,

n is the refractive index, m_{r_j} is the reduced mass, and Ai is the Airy function. The sum is over the light- and heavy-hole valence bands.

At a given photon energy, $\hbar\omega$, the absorption change induced by the electric field enhancement is, using Eq. 3.64, given by

$$\Delta\alpha\left(\hbar\omega;\,\Delta\mathcal{E}_{p-v}\right)=\alpha\left(\hbar\omega;\mathcal{E}_v\right)-\alpha\left(\hbar\omega;\mathcal{E}_p\right)\simeq\alpha\left(\hbar\omega;\,\Delta\mathcal{E}_{p-v}\right)-\alpha_0\left(\hbar\omega\right), \tag{3.65}$$

where $\alpha_0\left(\hbar\omega\right)$ corresponds to the residual absorption loss of the material. Equation 3.65 assumes $\alpha\left(\hbar\omega;\mathcal{E}_v\right)\gg\alpha\left(\hbar\omega;\mathcal{E}_p\right)\simeq\alpha_0\left(\hbar\omega\right)$, with $\mathcal{E}_v\simeq\Delta\mathcal{E}_{p-v}\gg\mathcal{E}_p$.

3.5.3 RTD-EAM Small-Signal Equivalent Circuit

This subsection presents a small-signal analysis of the RTD-EAM, which includes the transit-time effects associated with the thick depleted spacer layer in the collector core region. Design considerations to maximize the operating frequency for a given modulation depth are also discussed.

Apart from the higher bandgap waveguide cladding layers and the thick emitter spacer layer, the RTD-EAM, Fig. 3.32(a), resembles the quantum well injection transit time (QWITT) diode proposed by Kesan *et al* [133]; the QWITT diode consists of a baseline DBQW structure with a considerably thicker spacer layer on the collector side, Fig. 3.32(b).

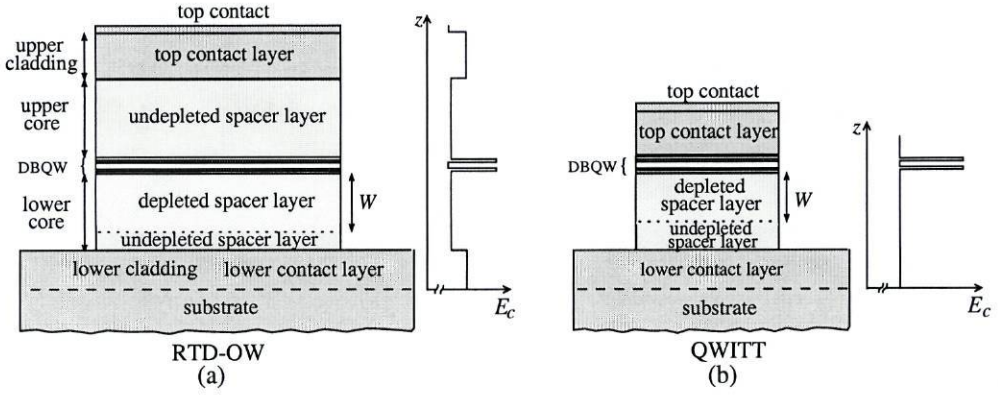


Figure 3.32: RTD-EAM and QWITT schematic diagrams.

Recent work [133]-[137] has demonstrated that the DBQW-RTD cannot be described by the simple equivalent circuit in Fig. 3.13(b), because a thick depletion region is formed "down-stream" from the DBQW, as shown in Figs. 3.30 and 3.31, and a substantial part of the terminal voltage may be dropped across it; the thickness of the depletion layer can be varied by changing doping and width of the spacer layer on the anode side.[24] Because the DBQW is much thinner than the depleted spacer layer length W, the depletion region transit time is much greater than the transit time through the DBQW: transport through the DBQW can be considered instantaneous in this approximation.

[24] An accumulation layer is also located close to the cathode barrier.

The emitter and DBQW can be effectively treated as an injection cathode whose small-signal equivalent circuit is the (geometric) capacitance of the well, in parallel with an injection conductance, σ, that represents the tunnelling current [134]. The lowly doped thick spacer layers produce a depletion region of width W on the collector side of the biased DBQW, much thicker than the tunnelling region L_{DB}, resulting in a transit-time delay [5][134], whose small-signal equivalent circuit corresponds to the drift region specific capacitance, c_{drift}, in parallel with the drift region specific impedance, z_{tt} [134]. Figure 3.33(a) presents the small-signal equivalent circuit for the RTD-EAM, where z_{DB} represents the DBQW specific impedance, c_{well} the effective capacitance, σ the injection conductance, and r_s represents the specific parasitic series resistance due to any undepleted spacer regions, highly doped contact regions, and ohmic contacts; in Fig. 3.33(b), C_{tt} represents the drift region capacitance, R_{tt} the drift region resistance, and R_s the drift region parasitic series resistance.

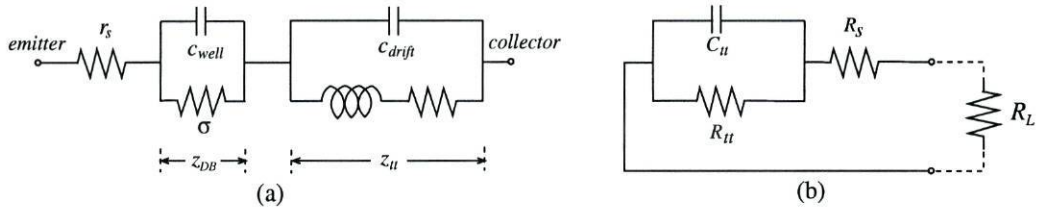


Figure 3.33: (a) Small-signal equivalent circuit for the RTD-EAM [134]. (b) Small-signal equivalent circuit of a RTD-EAM dominated by the transit-time effects.

For small-signal transit-time analysis, the injection region (emitter and DBQW here) can be characterized by a normalized injection conductance σ [5][134], defined as:

$$\sigma = \frac{\partial J_{DB}}{\partial \mathcal{E}_{DB}} \simeq L_{DB} \left(\frac{\partial J_{DB}}{\partial V_{DB}}\bigg|_{V_{appl}} \right), \tag{3.66}$$

where J_{DB} is the instantaneous current density, V_{DB} is the instantaneous voltage, and V_{appl} is the dc bias voltage, each quantity across the DBQW region only (excluding the voltage dropped across the depletion region, contact layers and metal-semiconductor contacts). At present there are no generally accepted accurate theoretical models that predict the transport behaviour (and hence σ) of DBQW structures. However, it is possible to estimate the value of σ using the dc $J-V$ characteristics of RTDs. To apply Eq. 3.66, the terminal $J-V$ characteristic must be corrected for any voltage drops across depletion, accumulation and contact regions. For a first estimate of σ the following expression can be used:

$$\sigma = \frac{L_{DB}}{\Delta V_{W=0}} \Delta J_{p-v}, \tag{3.67}$$

where $\Delta V_{W=0}$ is the voltage dropped across the DBQW in the absence of the drift region ($W = 0$); $\Delta V_{W=0}$ is around 0.1 V for typical DBQWs. Examination of both experimental and theoretical results in the literature indicates that for a wide variety of RTD structures, when biased to pro-

duce negative differential conductance, the room temperature value of σ lies between about -0.05 $\Omega^{-1}\text{cm}^{-1}$ and -0.5 $\Omega^{-1}\text{cm}^{-1}$. To obtain large values of injection conductance, the magnitude of the current density is more important than the PVCR.

The specific impedance of the injection region at frequency ω is given by

$$z_{DB} = \frac{L_{DB}}{\sigma + j\omega\epsilon}, \tag{3.68}$$

where ϵ is the effective dielectric function of the injection region. The specific resistance of the injection region, given by the real part of z_{DB}, is

$$r_{DB} = \frac{\frac{L_{DB}}{\sigma}}{1 + \left(\frac{\omega\epsilon}{\sigma}\right)^2}. \tag{3.69}$$

To obtain the transit-time specific impedance z_{tt}, charge transport through the depleted (drift) region must be considered. It will be assumed here that the depletion region electric field is high enough to cause the injected charge to traverse the depletion region at a constant saturation velocity, v_{sat}, and the injection conductance is assumed to be frequency independent. This permits the application of analytical methods previously established for other transit-time devices [49][137]. Based on these results [134][137], the specific impedance at frequency ω due to the depleted region is given by

$$z_{tt} = \frac{W}{j\omega\epsilon}\left[1 - \frac{\sigma}{\sigma + j\omega\epsilon} \times \frac{1 - \exp\left(-j\theta_d\right)}{j\theta_d}\right], \tag{3.70}$$

where $\theta_d = \omega W / v_{sat}$ is the drift angle. The transit-time specific negative resistance that can be obtained from the depleted spacer region is the real part of z_{tt},

$$r_{tt} = \frac{2\left(\frac{v_{sat}\epsilon}{2\sigma^2}\right)}{\left(\frac{\omega\epsilon}{\sigma}\right)^2\left[1 + \left(\frac{\omega\epsilon}{\sigma}\right)^2\right]}\left[1 - \cos\left(\frac{\omega W}{v_{sat}}\right) + \frac{\omega\epsilon}{\sigma}\sin\left(\frac{\omega W}{v_{sat}}\right)\right]. \tag{3.71}$$

The effective specific resistance of the RTD-EAM (excluding parasitic resistance in the device) is given by the sum of the specific resistance, r_{DB}, and the specific resistance due to the drift region, r_{tt}.

The DBQW injection σ is negative when the device is biased in the negative differential conductance region (NDC). At any given frequency there is an optimum value of the DBQW injection conductance $\sigma_{opt}^{r_{DB}}$, obtained from the solution of the equation $\frac{\partial r_{DB}}{\partial \sigma} = 0$, which maximizes the specific negative resistance of the injection region r_{DB}:

$$\sigma_{opt}^{DB} = -\omega\epsilon. \tag{3.72}$$

The parameter $\omega_0 = |\sigma_{opt}^{DB}|/\epsilon$ represents a characteristic angular frequency. For GaAs devices having $\Delta J_{p-v} \simeq 5$ kAcm^{-2} (from Fig. 3.5), and $|\sigma| = 0.05$ $\Omega^{-1}\text{cm}^{-1}$, Eq. 3.67, the characteristic

angular frequency is $\omega_0 \simeq$43 GHz; InGaAlAs based devices with $\Delta J_{p-v} \simeq$13.5 kAcm^{-2} (Fig. 3.9) and $|\sigma|$ =0.135 Ω^{-1}cm^{-1}, have $\omega_0 \simeq$117 GHz.

As expected for a transit-time device, there is a specific thickness of the depletion region W which yields the maximum negative resistance for any given frequency, injection conductance, and saturation velocity. The optimum width is found by solving the equation $\frac{\partial r_{tt}}{\partial W} = 0$, yielding

$$\tan\left(\frac{\omega W}{v_{sat}}\right) = -\frac{\omega\epsilon}{\sigma}. \tag{3.73}$$

The appropriate branch of the inverse tangent function to be used is $[-\pi/2, \pi/2]$, as σ is negative. At any given frequency there is also an optimum value of injection conductance σ, obtained from the solution of the equation $\frac{\partial r_{tt}}{\partial \sigma} = 0$:

$$|\sigma_{opt}| = \frac{\omega\epsilon}{\sqrt{3}}. \tag{3.74}$$

The DBQW negative resistance is maximized at a value of $|\sigma|$ that is $\sqrt{3}$ larger than that which maximizes the transit-time negative resistance. Assuming Eq. 3.72 is used to determine σ, the maximum negative resistances available from the injection and drift regions of the device are (using 3.73), respectively [134]:

$$r_{max}^{DB} = -\frac{L_{DB}}{2\omega\epsilon} \tag{3.75}$$

$$r_{tt} = -\frac{\sqrt{2}-1}{2}\frac{v_{sat}}{\epsilon\omega^2}. \tag{3.76}$$

Requiring that $|r_{tt}| \gg |r_{max}^{DB}|$, yields the condition [134]:

$$f \ll \frac{\sqrt{2}-1}{2\pi}\frac{v_{sat}}{L_{DB}} \simeq 0.066\frac{v_{sat}}{L_{DB}}. \tag{3.77}$$

If the value of the injection conductance is chosen to maximize the DBQW negative resistance, transit time effects will dominate as long as the frequency is lower than $0.066\frac{v_{sat}}{L_{DB}}$.[25] Examination of Eqs. 3.69 and 3.71 shows that in virtually all cases of interest the specific negative resistance of the RTD-EAM is dominated by the drift region specific negative resistance [134]. Figure 3.33(b) presents the small-signal equivalent circuit for the RTD-EAM controlled by transit-time effects.

For $\omega \ll |\sigma|/\epsilon$, R_{tt} is a constant given by (introducing the cross-section area, A) [137] [Fig. 3.33(b)]:

$$R_{tt} = \frac{r_{tt}}{A} \simeq \frac{W}{\sigma A} + \frac{W^2}{2\epsilon v_{sat} A}. \tag{3.78}$$

The second term represents the *space-charge* resistance familiar in the theory of conventional transit-time devices [11]; the first term represents the quantum well resistance. Overall, R_{tt} is negative if σ is negative and $W < 2\epsilon v_{sat}/|\sigma|$. A maximum negative resistance of $-\epsilon v_{sat}/2\sigma^2 A$ is

[25] Assuming a saturation velocity of 1×10^7 cm/s and a typical DBQW length of 10 nm, the transit-time effects dominate when the operating frequency is less than $\sim$ 660 GHz.

obtained for $W_{opt} = \epsilon v_{sat}/|\sigma|$. Thus, for low frequencies, $\omega \ll |\sigma|/\epsilon$, when the injection conductance σ is negative, the resulting transit-time negative resistance is broadband and independent of frequency; the optimum depletion layer thickness, W_{opt}, is also independent of frequency. Assuming $v_{sat} \simeq 1 \times 10^7$ cm/s, it comes that $W_{opt} \simeq 2.3$ μm and $W_{opt} \simeq 0.85$ μm for GaAs and InGaAlAs based devices, respectively.[26]

An equivalent expression to Eq. 3.78 has been derived for the capacitance [136][137] [Fig. 3.33(b)]:

$$C_{tt} \simeq \frac{A\epsilon}{W}\left[1 - \frac{|\sigma|W}{2\epsilon v_{sat}}\right]^{-1}. \tag{3.79}$$

The drifting space-charge modifies the capacitance, so that it is larger than the parallel plate capacitor value, $C = \frac{A\epsilon}{W}$.

For $\omega \gg |\sigma|/\epsilon$ and $\sigma < 0$, the optimum transit phase angle is frequency independent, approaching $\pi/2$ for high frequencies, corresponding to a depletion region width of $W = \pi v_{sat}/2\omega$. At fixed values of W and σ, as ω is increased just beyond $|\sigma|/\epsilon$, the device negative specific resistance rolls off as $|\sigma| v_{sat}/\epsilon^2\omega^3$ and changes from negative to positive at the value ω_{cutoff} (the intrinsic transit-time cutoff frequency) [134], given by:

$$\omega_{cutoff} \cong \frac{v_{sat}}{2W}\left[\pi + \sqrt{\pi^2 - \frac{8|\sigma|W}{\epsilon v_{sat}}}\right]. \tag{3.80}$$

At this frequency the device can no longer oscillate. Equation 3.80 does not consider any impedance matching constraints. The negative resistance available at the device terminals of area A is that of the "intrinsic" device resistance r_{tt}/A (Eq. 3.71), plus the positive parasitic resistance, r_s/A, associated with any undepleted spacer regions, highly doped contact regions, and ohmic contacts (all these parasitic terms scale with area). If R_L is the resistance of the external circuit, the matching constraints require the magnitude of the negative resistance at the terminals of the device to be R_L, i.e., $|r_{tt}|/A = r_s/A + R_L$. The maximum oscillation frequency is then [134]:

$$f_{max} = \frac{1}{2\pi}\left[\frac{|\sigma| v_{sat}}{\epsilon^2 (r_s + AR_L)}\right]^{\frac{1}{3}}.$$

For a completely optimized device, i.e., optimized for both σ and W, the absolute maximum specific negative resitance is given by $r_{tt_{opt}} = -v_{sat}/4\epsilon\omega^2$ [134]. In this case the maximum oscillation frequency is [134]:

$$f_{max} = \frac{1}{4\pi}\sqrt{\frac{v_{sat}}{\epsilon (r_s + AR_L)}}. \tag{3.81}$$

An upper bound on this frequency is found by requiring that the negative resistance from the device exceeds the contact resistance to the device, i.e., $R_L = 0$ and $r_s = r_{cont}$. Under these conditions, for a typical contact resistance of 10^{-6} Ωcm^2 and a saturation velocity of 1×10^7 cm/s, Eq. 3.81 predicts a maximum oscillation frequency of 234 GHz. In order to achieve such

[26] Using the σ values determined previously: $|\sigma|_{\text{GaAs}}$ =0.05 Ω^{-1}cm^{-1} and $|\sigma|_{\text{InGaAs}}$ =0.135 Ω^{-1}cm^{-1}.

a level of performance, Eq. 3.74 requires an injection conductance magnitude of 0.75 $\Omega^{-1}\text{cm}^{-1}$, much larger than has been reported in the literature. In addition, the depletion layer width for such an optimized device for $\sigma < 0$ is 55 nm (from Eq. 3.73). This corresponds to a very thin depletion region which would give rise to a very small overlap between the electric field and the guided optical field.

Considering the circuit of Fig. 3.33(b) and Eq. 3.29, the highest operating frequency is obtained for $R_L = 0$, when $R_{tt} = 2R_s$ with R_{tt} being (Eq. 3.78) [137]:

$$R_{tt} = \frac{W}{A|\sigma|}\left(1 - \frac{|\sigma|W}{2\epsilon v_{sat}}\right). \tag{3.82}$$

Solving for σ, and inserting in Eq. 3.79, $f_{max} = \frac{1}{4\pi R_s C_{tt}}$ (Eq. 3.30) becomes [137]:

$$f_{max} = \frac{W}{4\pi R_s A\epsilon}\left(1 + \frac{W^2}{4A\epsilon v_{sat} R_s}\right)^{-1}. \tag{3.83}$$

This expression enables the optimum depletion layer width and device area to be found for a particular frequency. The optimization is complicated by the fact that v_{sat} is likely to depend on W - a short drift region will result in greater velocity overshoot (section 2.1.3). The optimum value of W increases as R_s decreases, and f_{max} then becomes higher. It is thus extremely important for high frequency performance of RTD-EAMs, and RTDs in general, to be achieved that a very low series resistance is attained [5][152].

The assumption that carrier transport across the depleted spacer region takes place at a constant saturation velocity is clearly a poor approximation for depletion region widths smaller than 1 μm. Transient transport effects, such as velocity overshoot, can be significant in such circumstances, Fig. 2.4. Due to the quantum well structure, the injected carrier energies into the drift region fall between 0.1 and 0.15 eV, which will give an initial ballistic velocity of about 6×10^7 cm/s or more at the front of the drift region (velocity overshoot) [11]. In this case, the performance of RTD-EAMs operating in the negative injection conductance mode should be significantly enhanced, compared to the constant velocity case described here.

3.5.4 RTD-EAM Expected Performance and Applications

The parameters used to specify intensity modulators (the on/off ratio, the required voltage for a given on/off ratio,[27] the 3-dB bandwidth, insertion loss, and chirp) [13], can be employed to characterize RTD-EAM intensity modulators. Figure 3.34 shows the schematic layout of a generic RTD-EAM. A rough estimate of the performance to be expected from the RTD-EAM can be made considering a planar waveguide approximation. Assuming a $GaAs/Al_{0.3}Ga_{0.7}As$ waveguide (n_{co}=3.590, n_{cl}=3.385, for wavelength λ=0.9 μm), with core 1.0 μm thick, the fundamental mode is efficiently guided, having a confinement factor higher than 0.98 (Eq. 2.52).

[27] In generic electro-absorption modulators the required voltage for a given on/off ratio has also to be considered; however, because the RTD-EAM presents electrical gain this parameter does not apply here.

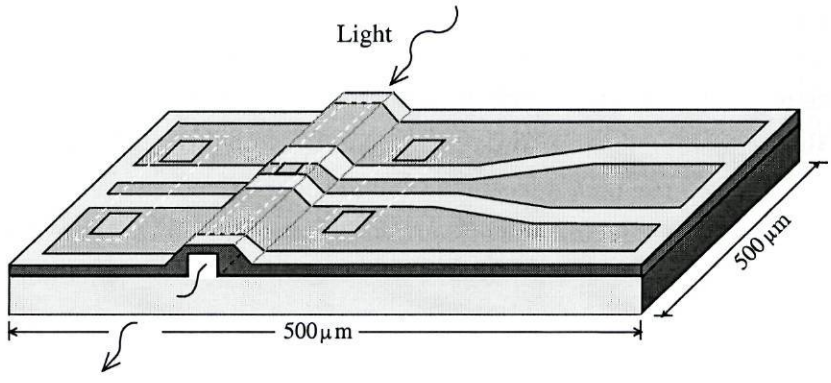


Figure 3.34: Schematic diagram of the RTD-EAM diode layout.

On/Off Ratio

As discussed previously, the absorption coefficient change, $\Delta\alpha = \alpha(V_{off}) - \alpha(V_{on})$, depends strongly on the applied voltage and the wavelength. In the case of the RTD-EAM, $\Delta\alpha$ is mainly determined by the device NDC characteristics (ΔV_{v-p} and ΔJ_{p-v}), as discussed, i.e., $\Delta\alpha = \alpha(V_v; J_v) - \alpha(V_p; J_p)$. Taking into consideration the experimental current-voltage characteristic, Fig. 3.5, of a symmetric GaAs/AlGaAs optical waveguide incorporating a DBQW in the centre of the core 1 μm thick, the electric field change, $\Delta\mathcal{E}_0$ (Eq. 3.61), across the depleted region ($W \approx$0.5 μm wide depletion layer) as a consequence of peak-to-valley switching is approximately 20 kV/cm. Applying Eq. 3.64, and considering $\alpha(\hbar\omega; \mathcal{E}_v) \gg \alpha_0(\hbar\omega)$, to the GaAs case, it comes that $\Delta\alpha \approx$200 cm^{-1} (Eq. 3.65); this gives, for a modulator with an active region 200 μm long, an approximate on/off ratio of 9 dB (Eq. 2.113, assuming $\gamma \approx 0.5$), which is considerable, taking into account the small bias needed, typically less than 0.5 V.

3-dB Bandwidth

The RTD-EAM bandwidth has been discussed in detail in the previous section. A RTD-EAM with a drift region W =0.5 μm wide will have a cut-off frequency around 37 GHz (Eq. 3.83, assuming A =800 μm^2, R_s=5 Ω, $v_{sat} = 1 \times 10^7$ cm/s, and $\epsilon = 13\epsilon_0$). The RTD-EAM cut-off frequency, for a given core doping concentration, is determined mainly by the width of the depleted collector region, device series resistance, together with the saturation velocity. A compromise must be reached between the different physical parameters, in order to be able to achieve given bandwidth and modulation depth values. The RTD-EAM cutoff frequency can be increased if the emitter core thickness is made thinner, which will reduce the series resistance R_s, and decrease the peak voltage. The peak current conductance of the RTD is improved due to a decrease in electron scattering in the shorter path from the emitter highly doped region to the first barrier. The peak-to-valley current ratio will be higher, giving rise to a large absorption change, which will compensate the decrease in the fundamental mode optical confinement factor.

Insertion Loss

The insertion loss is mainly due to the mode mismatch (in the growth direction) between the incident and the fundamental guided mode, which is considerable because of the small spot size of the fundamental mode of a 1 μm core waveguide (due to the large refractive index contrast). (The reflection loss can be eliminated by anti-reflection coatings applied on the waveguide facets.) The transfer of beam power to a given waveguide mode is accomplished by matching the beam-field to the waveguide mode field.

Assuming waveguide excitation with a standard single mode optical fibre (in the gaussian mode profile approximation), the overlap integral in the growth direction is $\eta_{C_1} \simeq 0.2$ (Eq. 2.118); the coupling loss (in dB) is $\alpha_c \simeq 8$ dB (Eq. 2.117).

Chirp

The Franz-Keldysh effect produces primarily electro-absorption [19][77]. The typical value for refractive index change in GaAs for a field of 2×10^4 V/cm is less than $\Delta n = 1 \times 10^{-3}$ @ $\lambda \sim 900\,\text{nm}$ [77]. Chirping in the RTD-EAM is expected to be small, because the operating wavelength is close to the absorption bandedge energy, where $\Delta\kappa = \Delta\alpha/(4\pi/\lambda)$ is expected to be considerably larger than Δn ($\Delta\kappa \simeq 1.3 \times 10^{-3}$).

Applications [13][14][15]

The RTD-EAM can be used as a short optical pulse generator with a high repetition rate, using its non-linear response to applied voltage. As will be shown, the repetition frequency can be imposed by the external circuitry, and a high duty ratio of pulse width to repetition period achieved, which can be employed to implement an optical soliton source. At the same time, its integration with other functional devices can be used to encode generated optical pulses [74][170].

60 GHz fibre-radio systems in which optical subcarrier signals from a base station are fed through fibre into each terminal station have been proposed [3], not only because the 60 GHz band has not yet been allocated to public communications, but also because the penetration distance of rf signals in this band is short enough to allow cells a few hundred metres apart, and thus increase the system capacity. In principle, the RTD-EAM can detect optical signals, recovering the original transmitted rf signal, and also modulate the optical carrier with the rf signal to be transmitted [74][75]. As this kind of system needs a large number of terminal stations, the RTD-EAM could lead to significant cost reduction.

In chapter 4, the implementation steps required to achieve a suitable RTD-EAM configuration will be described in detail. The specific RTD-EAM structures designed and fabricated for operation around 900 nm and 1550 nm will be presented in chapters 5 and 6, respectively, together with the experimental results obtained.

3.6 Final Comments

The electrical characteristics of DBQWs were discussed in terms of the relevant current transport mechanisms, stressing the NDC. The influence of device parameters and bias conditions on high frequency operation was addressed. A small-signal equivalent circuit was discussed, and self-oscillation and stability analyses were presented.

Specific aspects of the characteristics of the GaAs/AlAs, and InGaAs/AlAs material systems for implementation of RTDs were presented, together with a discussion of the device main structural design parameters leading to high performance.

Some applications of RTDs were summarized, with particular relevance given to optoelectronics (sources, detectors, bistable devices). These discussions allowed an understanding of the physical principles of the RTD-EAM, including its high frequency behaviour. Experimental results obtained from fabricated devices, such as I-V curves and relaxation oscillation waveforms, were used to illustrate some concepts introduced in this chapter.

A detailed model was adapted from the one used with transit-time devices, which provides a good insight into the relevant design parameters (such as doping profile and collector space layer width). Device designs can be extracted and fabrication principles derived from this model, as described in the next chapter. The model points to some relevant characteristics: wide-band optical modulation (up to 60 GHz), low power consumption, and capability of optical modulation by self-oscillation due to the intrinsic device gain.

Chapter 4

RTD-EAM: Design, Fabrication and Testing

This chapter intends to be a guide to the various steps involved in design, fabrication and characterization of the RTD-EAM, produced from a semiconductor wafer grown by MBE and containing a double barrier quantum well structure in the centre of the highest refractive index layer (waveguide core). For operation in the 900 nm spectral region, GaAs is used to form the waveguide core and the quantum well, AlAs and the AlGaAs ternary material system are employed to form the barrier and waveguide cladding layers, respectively. In the 1300/1600 nm optical transmission windows, the InGaAlAs quaternary material system can be used to form the waveguide core and the quantum well, with AlAs and $In_{0.48}Al_{0.52}As$/InP being employed in the barriers and in the waveguide cladding layers, respectively. The alloy composition of the semiconductor has to be appropriately selected to provide the required waveguide bandedge and index profiles.

4.1 RTD-EAM Design

The RTD-EAM waveguide structure should be designed to enable operation over a wide bandwidth, require low drive voltage, give a large absorption change and introduce a low insertion loss, i.e., the RTD-EAM waveguide I-V characteristic must show considerable negative differential conductance (NDC), and at the same time provide efficient light guidance with low optical background loss.

A compromise must be reached between the different parameters (such as the required waveguide NDC characteristic, the core thickness and optical background loss), in order to achieve given values of bandwidth and modulation depth, without inducing high coupling and propagation loss. The electric field enhancement due to the peak-to-valley transition increases with the PVCR, and

so the absorption change, which depends mainly on the DBQW structure, specially through the barriers thickness. As discussed in the previous chapter, the width of the depleted core collector region, together with the electron saturation velocity, determines the cut-off frequency of the device. Also, in order to achieve high frequency operation, the device series resistance must be as low as possible, which requires considerable doping across the structure [5].

Important design parameters are the core and claddings thicknesses, the waveguide refractive index profile, the doping concentration across the structure, and the barriers and well widths.

Coplanar waveguide transmission lines were used to connect the RTD-EAM to the external world, as high frequency operation is sought for.

4.1.1 Waveguide Structure

In ridge RTD-EAM waveguides, optical confinement (waveguiding), Fig. 4.1, is achieved in the vertical direction, perpendicular to the substrate plane, by a refractive index difference due to material compositional changes, and in the lateral direction by "effective index differences" due to a ridge etched into the material layers. The ridge waveguide structure provides both light confinement along the DBQW plane and a suitable path for the injected (applied) current (electric field) in the growth direction (perpendicular to the tunnelling plane), as represented schematically in Fig. 4.1, for the case of the GaAs based RTD-EAM. Light confinement in the tunnelling plane results in a much larger interaction between the depletion region electric field and the optical field.

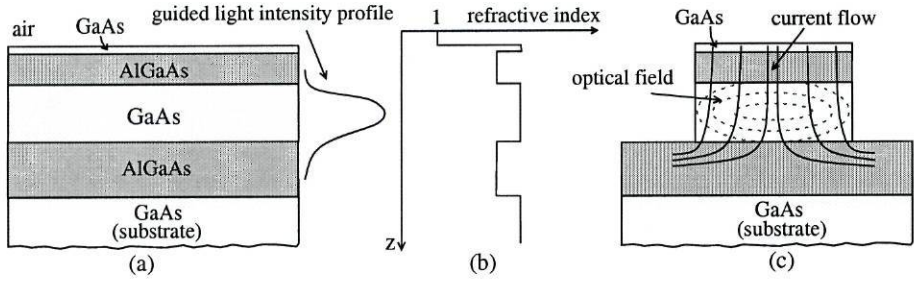


Figure 4.1: GaAs-$Al_xGa_{1-x}As$ double heterostructure waveguide, and corresponding sketches of guided light intensity and refractive index profiles. Light and injected current confinement are also represented.

Waveguide Geometry

The waveguide core of the RTD-EAMs was made 1 μm thick to allow easy endfire light coupling, without excessive coupling loss; the cladding layers were made around 0.3 μm thick with a core/cladding refractive index difference of $0.2-0.3$. The planar waveguide obtained is essentially single mode in the growth direction.[1] Practical lateral dimensions of the optical waveguide (ridge

[1]In order to increase the modulator bandwidth and operation with much lower power, the waveguide core thickness can be made smaller, i.e., the optical field can be confined in a thinner active region ($\sim$0.5 μm).

widths larger than 1 μm) and high index contrast between air and the semiconductor materials involved, $[(n_{core} - n_{air})/n_{core}] \simeq 0.7$, give rise to multimode waveguides. However, when a single-mode fibre is used to launch light into the ridge, it is possible to excite only the fundamental mode, as long as there are not many propagating modes and the distance of propagation is short (typically less than 1 mm) [13][58]; this was verified experimentally. As an approximation, the propagation and absorption properties of the optical radiation were treated in terms of only the fundamental mode, throughout this work.

The roughness of the ridge side walls must be minimized in order to reduce the scattering loss. Free carriers due to the core and claddings doping levels can cause some attenuation of the guided modes. However, as long as the device length is less than 1 mm, and the core doping level is below 10^{17} cm^{-3}, the free-carrier absorption is negligible [58] when compared to the material background loss. The light coupling efficiency from a single-mode fibre into a semiconductor waveguide supporting a few modes is generally very small. It can be improved by designing the waveguide to operate near the fundamental mode cut-off, in order to increase the waveguide mode size, or by fabricating 3D tapers [13][58].

As there is no exact analytical solution of the wave equation applied to a ridge profile waveguide, approximations such as the effective index method and the beam propagation method can be used to study light propagation [59]. The ridge waveguides to be used in this work were analysed using *FWave*,[2] a software based on a finite difference implementation of the vector electromagnetic wave equation [59].

Waveguide Background Doping

Once the waveguide composition and geometry are defined, a parameter that can significantly affect the RTD-EAM J-V characteristic is the background doping across the waveguide.

Doping, specially in the cladding layers, has to be kept as high as possible to minimise the device series resistance, allowing high frequency operation without inducing large optical losses through free-carrier absorption. Taking into consideration that the fraction of the optical power guided in the claddings is expected to be small, the claddings doping concentration, $N_{d_{cl}}$, can be considerably higher than the core background doping, and still not give rise to large free-carrier absorption loss.

The depleted portion of a RTD collector spacer layer is mainly determined by its doping concentration. The waveguide core (RTD spacer layers) doping concentration, $N_{d_{co}}$, should allow complete depletion of the collector spacer layer (the drift region) and induce low free carrier absorption.[3] However, too low core doping can have a deleterious and even a catastrophic effect [11][63]. There exists an optimum collector spacer layer (drift region) length, W_{opt}, for a given background doping density, $N_{d_{co}}$. Figure 4.2 shows an illustration of the electric field profiles,

[2] M. R. S. Taylor, "FWave IV: a vector E-M wave solver," user manual, University of Glasgow, Glasgow, United Kingdom (URL: http://www.elec.gla.ac.uk/~michaelt/software/fwave/).

[3] The spacer layer on the emitter side of the DBQW region will be undepleted under normal bias, and its main effect is to increase the total parasitic series resistance of the device.

corresponding to the peak and valley points of the RTD J-V characteristic, across the drift region; the shaded area is related to the difference between the peak and valley voltages, ΔV.

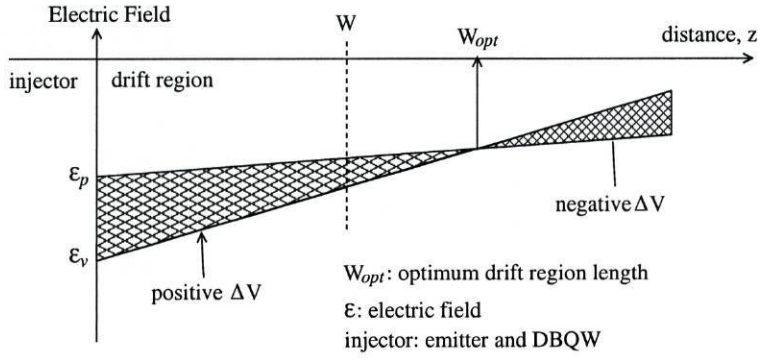


Figure 4.2: Illustration of the electric field across the drift region, at the peak and valley points [63]. Because the current density is lower at the valley, the slope of the field is greater than at the peak.

In the following analysis, the electric field in the drift region will be assumed to be high enough to cause the injected electrons to traverse it at a constant saturation velocity, v_{sat} [5][63]. For the slope of the electric field in the drift region to be positive, attending to Poisson's equation, $\frac{\partial \mathcal{E}}{\partial z} = \frac{e}{\epsilon}\left(N_{d_{co}} - \frac{J}{ev_{sat}}\right)$, the injected electron density necessary to support the current density J, $N_J = J/ev_{sat}$, must be less than the background doping density, $N_{d_{co}}$. However, if the drift region background doping concentration is smaller than the electron concentration required to support the current density, the magnitude of the electric field in the collector region increases, and can reach a sufficiently high value to cause breakdown, mainly due to impact ionization events (which are highly probable specially for narrow bandgap materials) [11][160]. Electrons entering the drift region gain energy from the rising electric field and participate in impact ionization, resulting in more electrons in the drift region, which contribute to a further increase in the field; this positive feedback process eventually leads to breakdown [49]. The holes that are created in the impact ionization events drift towards the DBQW and become trapped at the valence band heterojunction between the barrier and the collector spacer layer [160]. The presence of positively charged holes will then locally enhance the electric field at this heterojunction and also result in a positive feedback mechanism, again leading to breakdown [63].

Rigorous self-consistent models incorporating accurate impact ionization processes are required to understand breakdown [2][11]. From a device design point of view, breakdown can be avoided by ensuring that the drift region doping is higher than the expected peak electron concentration, $N_{d_{co}} \geq \frac{J_p}{ev_{sat}}$ [63].

4.1.2 RTD-EAM Structure

The RTD-EAM waveguides based on the GaAs/AlGaAs system were implemented in "semi-insulating" (SI) GaAs substrates and consist of two thick low doped layers of GaAs (core) with

absorption bandedge around 900 nm, and a GaAs/AlAs DBQW in between.[4] The waveguide cladding layers are formed by two highly doped layers of $Al_{0.33}Ga_{0.67}As$. Because the waveguide core and the substrate have similar real refractive index (dashed line in Fig. 4.3), the lower cladding layer is made thicker than the upper cladding in order to avoid radiation leakage towards the substrate.

The RTD-EAM waveguides operating at 1550 nm were implemented in InP substrates, which also act as the lower cladding region, with two low doped layers of $In_{0.47}Ga_{0.42}Al_{0.05}As$ with absorption bandedge around 1520 nm, separated by a $In_{0.47}Ga_{0.53}As/AlAs$ RTD, forming the core region. The upper cladding is formed by a $In_{0.48}Al_{0.52}As$ layer.

Both RTD-EAM structures were grown by Molecular Beam Epitaxy in a Varian Gen II system. The generic wafer structure is shown schematically in Fig. 4.3, together with the refractive index and the Γ-valley depth profiles (the dashed line applies to GaAs wafers only).

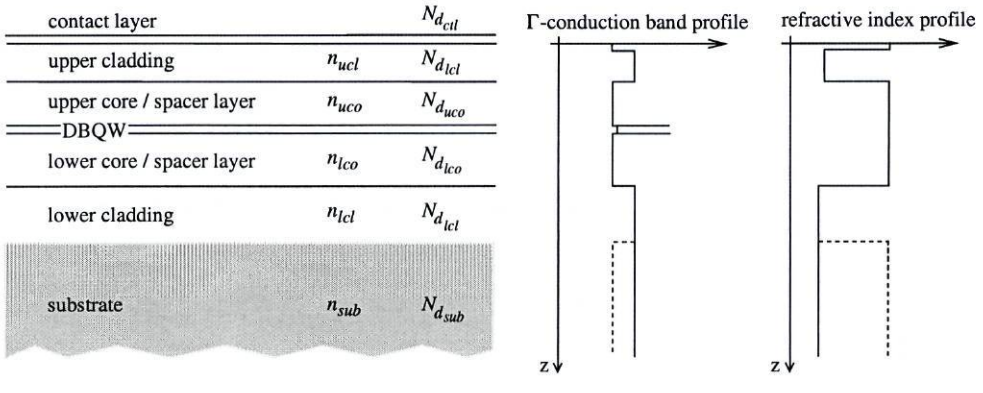


Figure 4.3: RTD-EAM waveguide wafer diagram, Γ-valley and refractive index depth profiles.

4.1.3 High Frequency Bonding Pads

Low frequency circuit theory is based on the assumption that all circuit dimensions are small compared with the electromagnetic wavelengths of interest, with any propagation phase delays being ignored. The opposite approximation is used in classical geometric optics, where all dimensions are assumed to be much larger than the wavelength; this assumption ignores diffraction and interference effects associated with the wave nature of the electromagnetic field. Microwave and millimetre-wave circuits operate in the regime where neither approximation holds valid; simple interconnects have to be treated as transmission lines, and interference cannot be ignored [10][178]. The propagation time of the signal energy, and electromagnetic field radiation, absorption and reflection taking place at discontinuites, have to be considered [10][178]-[183]. In the domain of ultra-high frequency circuitry, "stripline" techniques that involve parallel-conductor transmission lines as part of the actual circuit are routinely used [10][178]-[183].

[4]The SI substrate reduces device parasitic capacitance to ground and interconnecting track parasitic capacitance on top of the substrate, which are important features when considering operation at high frequency.

Transmission Line Categories [179][180]

There are four general categories of transmission lines: two-wire (parallel), coaxial, waveguide, and printed transmission lines. Two-wire lines are used only at low frequencies (the low end of the microwave band); their primary limitation is due to radiation loss, but they also exhibit dielectric and skin effect losses. Above 1 MHz, coaxial lines are more efficient, as they overcome the radiation loss problems due to their inherent shielding provided by the outer conductor.

Coaxial lines connectorization characteristics should also be considered. The typical reflection coefficient of a well made BNC connector rises to about 0.1 at 1 GHz, and BNCs should not be used above this frequency. SMA connectors[5] for use up to 18 GHz have an outer conductor diameter of 3.5 mm, with a reflection coefficient rising to about 0.1 at 20 GHz. Good quality cables and connectors are difficult to make, and the higher the frequency, the smaller their outer conductor diameter becomes. Currently, the highest frequency coaxial connector available works at up to 60 GHz, and it has an outer conductor diameter of 1.8 mm. The performance of this connector is strongly dependent on the torque which is applied in setting up the connection, and is limited by dielectric and skin effect losses. However, its main drawback is a limited power handling capability.

Waveguides have better all-around characteristics than most other transmission lines. They are used almost exclusively in radar, with its high power requirements.

Near the millimetre wave band, printed transmission lines, mainly microstrip and coplanar waveguide (CPW) types, have distinct advantages. Particularly useful characteristics include: d.c. as well as a.c. signals may be transmitted; active devices, diodes, and transistors may be readily incorporated (shunt connection is also quite easily made); in-circuit characterization of devices is implemented straightforwardly. The signal wavelength in the line is reduced considerably from its free-space value, because of the permitivity of the substrate; hence, distributed component dimensions are relatively small. The structure is quite rugged and can withstand moderately high voltages and power levels.

Figure 4.4 shows the equivalent circuit model of a uniform transmission line. The unit length constants of a transmission line are: the resistance, R, and the inductance, L, along the line, respectively; the conductance, G, and the capacitance, C, shunting the line (all specified "per loop metre", which means that each quantity is determined on a "go and return" basis).

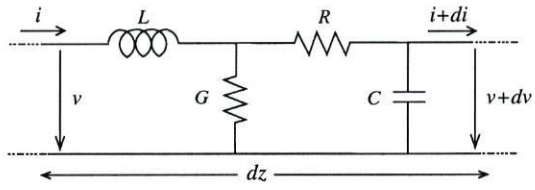


Figure 4.4: Equivalent circuit model of a uniform transmission line.

[5]High performance subminiature connector, widely used in demanding military systems and test equipment, as well as where miniaturization is desired [179].

Propagation of a sinusoidal wave along the line is characterized by the (complex) propagation coefficient:

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)\,(G + j\omega C)}. \tag{4.1}$$

The wavelength of the propagating field in the line is reduced according to the square root of the relative permitivity of the material in which the line is embedded, $\lambda_g = \lambda_0/\sqrt{\epsilon_r}$, and consequently the wave is slowed by $\sqrt{\epsilon_r}$. The characteristic impedance Z_0 is given by[6]

$$Z_0 = \sqrt{\frac{(R + j\omega L)}{(G + j\omega C)}}. \tag{4.2}$$

In microwave transmission lines the effects due to L and C tend to dominate, because of the relatively high inductive reactance and capacitive susceptance, respectively. In such cases the lines are referred as *loss-free*, although in practice some information about R or G may be necessary to determine actual power losses. At sufficiently high frequencies, $\beta \simeq \omega\sqrt{LC}$, and $Z_0 \simeq \sqrt{\frac{L}{C}}$. The last expression, which is a purely resistive result, can be written in the form $Z_0 = 1/v_p C$, where $v_p = \omega/\beta$ is the phase velocity.

For coaxial lines, two values of the characteristic impedance are commonly encountered: at V.H.F (30 - 300 MHz), $Z_0 = 75\ \Omega$; at UHF and SHF (300 MHz - 30 GHz), $Z_0 = 50\ \Omega$.

Coplanar Waveguide Transmission Line [10][180]

A coplanar waveguide (CPW) transmission line consists of a strip of a thin metallic film deposited on the surface of a dielectric slab, with two ground electrodes running adjacent and parallel to the strip on the same surface [181], as shown in Fig. 4.5.

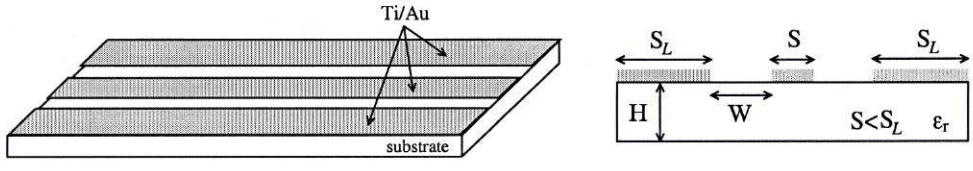


Figure 4.5: Coplanar waveguide transmission line schematic geometry.

There is no low frequency cut-off, and the coplanar configuration of all conducting elements permits easy connection of external shunt elements, such as active devices, as well as fabrication of series and shunt capacitance. It is also ideal for connecting various elements in monolithic microwave integrated circuits built in semiconducting substrates. The most important parameters are the central conductor width S ($S \ll S_L$), the gap between the central and the ground conductors W, the substrate thickness H and its relative permitivity, ϵ_r. The width of the ground planes S_L should be larger than the gap width W to ensure single mode operation. The thickness of the metallic film strips is generally less importance and may quite often be neglected in the

[6]By definition, the characteristic impedance, Z_0, is the value of wave impedance, v/i, measured at the input of the transmission line when its length is infinite [180].

analysis. However, to ensure good operation, the minimum thickness recommended is at least four times the skin depth in the metal at the lowest frequency of interest [10][182].

CPWs have been analysed using quasi-static as well as full-wave methods [181][182]. The quasi-static methods provide information about characteristic impedance and effective line permitivity, ϵ_{eff}, neglecting the frequency dependence of these parameters. A full-wave analysis of coplanar lines provides information regarding frequency-dependence of phase velocity and characteristic impedance.

The following closed form design equations can be used in most of the practical implementations of coplanar line circuits [183]:

$$Z_0 = \frac{30\pi}{\sqrt{\epsilon_{eff}}} \frac{K(k')}{K(k)} \tag{4.3}$$

$$\epsilon_{eff} = 1 + \frac{\epsilon_r - 1}{2} \frac{K(k')K(k_1)}{K(k)K(k_1')}. \tag{4.4}$$

where $K(k)$ represents a complete elliptic function of the first kind [82], and $K'(k)$ its complementary function $[K'(k) = K(k')]$, with $k_1 = \sinh(\pi S/4H)/\sinh[\pi(S+4W)/4H]$, $k = S/(S+2W)$, and $k' = \sqrt{1-k^2}$; $K(k)/K(k')$ can be approximated as

$$\frac{K(k)}{K(k')} \simeq \left[\frac{1}{\pi}\ln\left(2\frac{1+\sqrt{k'}}{1-\sqrt{k'}}\right)\right]^{-1} \qquad 0 \le k \le 0.7 \tag{4.5}$$

$$\frac{K(k)}{K(k')} \simeq \frac{1}{\pi}\ln\left(2\frac{1+\sqrt{k}}{1-\sqrt{k}}\right) \qquad 0.7 \le k \le 1. \tag{4.6}$$

It is customary to have input and output impedances of a high speed device matched to those of coaxial cables or rectangular waveguides over a range of frequencies, rather than at a single frequency. For matching to coaxial cables or broadband matching, the impedance of choice is normally 50 Ω. This standard allows cascading of devices without any concern for multiple reflections and instabilities caused by interconnection lines, as well as efficient transfer of power from one device to the next. Several elements, such as attenuators and filters, require controlled inpedance levels for their proper operation [179].

The above expressions were used to design RTD-EAM high frequency contact pads and package transmission lines, consisting of wide bandwidth (beyond 60 GHz) 50 Ω CPW transmission lines with adiabatic impedance preserving tapers, as represented in Fig. 4.6. The CPW ground planes were short circuited periodically to ensure that they have the same potential.

4.1.4 Design and Production of Masks

Once the design of a device was complete, a computer aided design software programme (Wavemaker) allowed direct generation of the pattern design files, which could then be analysed and "written". The designed patterns can be classified as windows, stripes and lines. The windows may be used to make connections to a conductor through an insulator. The conducting intercon-

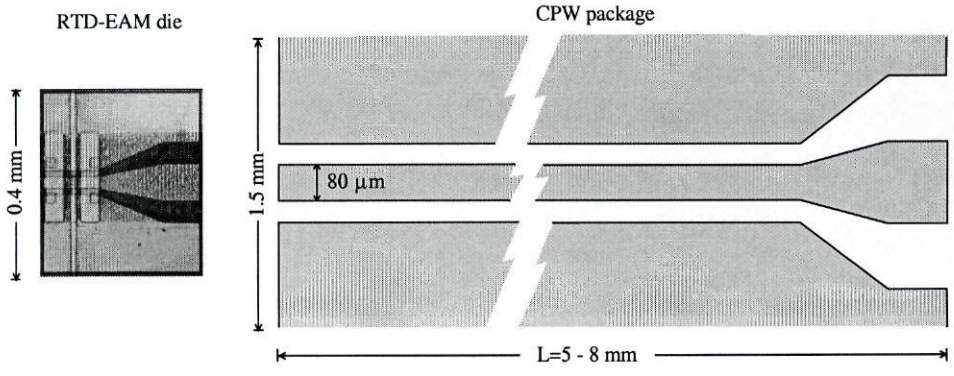


Figure 4.6: RTD-EAM and package CPW lines with impedance preserving tapers.

nects are usually formed into stripes and line patterns. Lines are also employed to define optical waveguide widths.

As the wafer, in its virgin form, has little irreversible sensitivity to optical or electron beams, the designed patterns are transferred to an electron active polymer layer (*resist*) previously spread uniformly, in a process known as *spin coating*, on the wafer epilayer or on a master mask plate, employing e-beam lithographic techniques [20][21]. Because the smallest feature encountered in the device is about 1 μm wide, which is well within the resolution of the photolithography apparatus available ($\sim$0.25 μm), e-beam lithography was only used for generating the masters of the photolithographic mask patterns.

A master mask consists of a plate of a material transparent in the ultaviolet (UV), generally fused silica or borosilicate glasses, having good mechanical and thermal properties, and covered on one side by a thin layer of metallic chrome [20][21]. After development, the resist layer reproduces the written pattern, which is then transferred to the mask metallic layer by a chrome chemical etch; the regions of the master mask protected by the resist remain opaque, covered with metallic chrome. The master plate is then copied optically, using a contact printer, which employs a collimated UV light source - mercury lamps are the common light sources - that exposes, through the master mask, a different mask plate covered by an UV active resist, previously brought into contact with the master plate [20][21]. These new mask plates (photolithographic masks) have a thin layer of iron oxide, instead of a metallic chrome layer, to facilitate alignment during the photolithographic processes. Once the master pattern is transferred, after development, the resist layer reproduces the pattern, which is then copied to the plate by an iron oxide chemical etch.

Typical photolithography techniques are proximity printing, contact printing, projection printing, and optical stepping [20][21]. Contact printing was the technique used throughout this work. In this technique, the photolithographic mask is brought into contact with the sample resist layer, and exposed to UV light using a mask aligner. The exposed areas of the resist are next washed away when the samples are soaked in an appropriate alcaline solution [20][21]. This technique is capable of defining geometries below 1 μm. Intimate contact between mask and sample is absolutely essential to obtain good resolution and pattern definition in contact printing; optical interference fringes should be visible (if using see-through masks) when the sample

is clamped to the mask. Contact between sample and mask must be carefully realized, however, because it tends to damage small semiconductor samples.

4.2 RTD-EAM Fabrication

The RTD-EAM wafers were diced into 1 cm^2 samples, followed by cleaning. The samples were subsequently patterned by optical lithography, defining the sample epilayer regions corresponding to the emitter and collector contacts, followed by the n-type ohmic contact stripes metallization. The sample epilayer was again patterned to form the waveguide etching mask, followed by epilayer etching using either a Reactive Ion Etching (RIE) or a wet etching process, forming the mesas and channel waveguides. A dielectric layer (SiO_2 or SiN) is then deposited and used to act as an insulating film along which the metallization tracks are run; contact access windows are opened on the waveguide and mesas, allowing connection to be made through high frequency bonding pads (CPW transmission lines).

Figure 4.7 shows the schematic of the RTD-EAM.

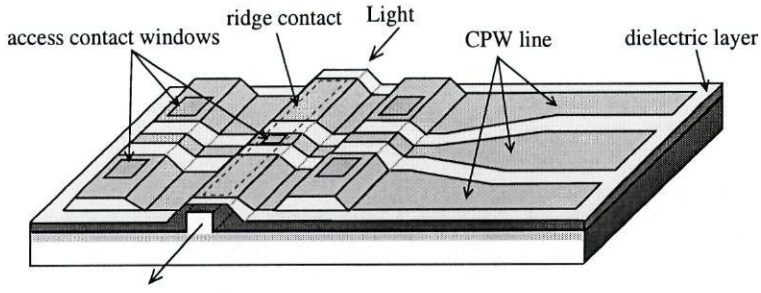


Figure 4.7: Schematic of the RTD-EAM.

Techniques such as thinning, cleaving, encapsulation, bonding and lead attachment, are used in device packaging.

4.2.1 Photolithography

In total, four photolithographic steps employing four different patterned mask plates were used in the RTD-EAM fabrication process: patterning of ohmic contacts; definition of waveguides and mesas; creation of access contact windows; patterning of the high frequency bonding pads. A Karl Suss mask aligner was used throughout this work, providing good alignment and contact between sample and mask, achieving high resolution (below 0.5 μm) in patterning of the resist. In this machine, the UV light output of the mercury lamp is continuously monitored to mantain constant intensity. All the lithography fabrication phases occurred in a sterile and dust-free environment (clean room class 100).

Sample Preparation [20][21]

Sample preparation consisted basically of wafer cleaving into small samples and their cleaning in ultrasonic baths. A quarter of a 3 inch wafer was cleaved into samples of suitable size, which depended on the dimensions of the devices and the number of devices desired per sample. It is not recommended to use large samples because this would result in waste of material when a fabrication step goes wrong. However, samples should not be less than 5 mm across, otherwise their handling becomes difficult; a square sample 1 cm across will produce a suitable yield of devices. Before cleaving, it is necessary to ascertain which is the epilayer side (mirrored); in case both sides are polished, the substrate side must be scratched (for identification) before wafer cleaving. A portion of the wafer samples was used for calibration at different fabrication steps. In order to avoid damage to the samples, they must be handled with plastic tweezers.

The samples were then cleaned in opticlear, acetone, and methanol ultrasonic baths, before proceeding further. Five minutes in the ultrasonic bath stage for each was adequate. The samples must be completely immersed at each stage and were not allowed to dry in between stages. After the final solvent (methanol), the samples were flushed out with deionised water, then were blow dryed with a jet of filtered air. Subsequently, they were allowed to dry on a glass dish for 5 minutes in an oven at 90 ^{0}C, to ensure that the samples were completely dry.

Resist Coating[7] [20][21]

For a resist to be reliable, it must have good bonding properties to the substrate, its thickness must be uniform and be reliably controlled over different sample runs. Both the resist and the sample must be exceptionally clean; any particle contamination will result in poor imaging or defects. Relative humidity must be also controlled, and it should generally be below 50%. The sample must be completely dry before resist is applied; a bake step or hot plate step should also be included to ensure complete dryness.

The sample was then placed on the vacuum chuck of the spinner, and a small puddle of a suitable resist (S1818) was applied to the centre of the sample, which was held to the spindle by vacuum. The chuck was previously tested with a sample mounted to verify operation and spinning speed; this also helped to remove any dust that the sample may have acquired. The resist was applied using a syringe fitted with an appropriate filter to remove particles as small as 0.2 μm. The spindle was rotated at a rate of around 4000 revolutions per minute (rpm) for 30 seconds, which resulted in a film thickness of around 1.8 μm. During the first couple of seconds of spinning, most of the resist is thrown off and drained away. The remaining resist forms a thin layer whose thickness is controlled by the spin speed (the thickness is proportional to $1/\omega$, where ω is the spinning frequency). The spinning operation also pre-dries the resist, so that it stays in place after spinning ceases. The photoresist tends to pile up seriously on the edges of small samples and fails to form a uniform film when it is spun on. The edge bead formed can be several times the thickness of the film, and must be carefully removed. The resist requires a

[7]Throughout the fabrication, Shipley Microposit resists were used.

post-spin bake to harden it by removing more of the remaining solvents. The samples were soft baked in an oven at 90 ^{0}C for 30 minutes, to improve resist adhesion. To promote resist bonding over an oxide, or any dielectric, an adhesion promoting primer is usually applied before resist coating.

Mask Alignment and Pattern Exposure [20][21]

The first photolithographic step defines the ohmic contacts regions. Before mask and sample alignment, the ferric mask plate pattern had to be cleaned in an acetone ultrasonic bath, flushed out with deionised water and blow dryed. The ohmic contact mask pattern consisted of a set of transparent lines of different widths (2-6 μm) spaced by 500 μm, and arranged in groups of variable length (100, 150, 200, 300 μm), Fig. 4.8(a). Later on, a different ohmic contact pattern was used, consisting of a similar set of transparent lines but including rectangular windows (50 μm $\times$ 100, 150, 200, 300 μm), on both sides of each line and 50 μm away from it, Fig. 4.8(b). Figure 4.8 schematically shows the ohmic contact pattern corresponding to a single device.

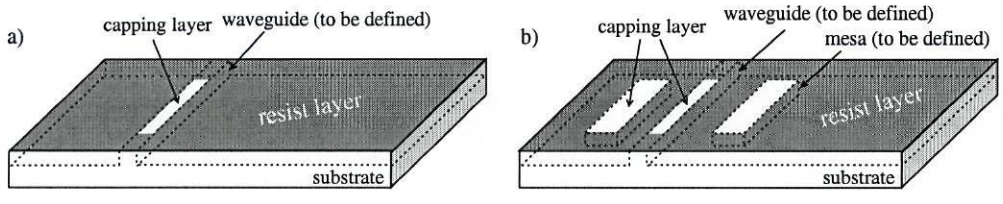


Figure 4.8: RTD-EAM ohmic contact pattern, without and with mesa configurations.

The mask is placed pattern side down in the mask aligner holder and is firmly held in place by a vacuum clamp. The mask pattern is brought down towards the sample resist-coated side up and held in place by another vacuum clamp. The mask stays close to the sample without contact; the mask and sample alignment is checked using a microscope and adjusted using micrometer drives. Once properly aligned, the sample is put into contact with the mask and is then exposed to the UV light for the required time (around 7 seconds).

Pattern Development [20][21]

A resist which becomes more soluble when exposed to illumination is called positive, and its pattern is identical to the opaque layout of the mask plate (the inverse applies to negative resists). Positive resists are normally preferred because they offer higher resolution and allow more appropriate edge profiles for *lift-off* processes.

A typical positive resist is composed of three major components: a photoactive compound (inhibitor), a base resin, and a suitable organic solvent system. All or most of the solvent is lost after spinning and baking. The base resin alone is moderately soluble in an aqueous alkaline developer, such as dilute sodium hydroxide. The removal rate of unexposed resist is near 1 nm/s. Radiant energy in the wavelength range of 200-450 nm essentially destroys the photoactive component and results in an increased removal rate, up to 200 nm/s.

The resist pattern was developed in a 1:1 mixture of Microposit developer with deionised water. The sample was placed in a beaker and the developer solution added; the developing time was approximately 90 seconds. The sample was then removed and rinsed in deionised water for 2 minutes before blow drying with an air gun. The sample was inspected under a microscope with a yellow filter in place to check if it had been properly developed; the development was repeated, whenever necessary. Post baking for 5 minutes at 90 ^{0}C followed, and the result is schematically represented in Fig. 4.8.

4.2.2 Metallization

Metallization provides metal to semiconductor contacts (ohmic contacts and Schottky barriers), interconnects between various points on the chip (transmission lines), and compatible bonding pads for packaging.

Metal Deposition [20][21]

Metal films are typically deposited on semiconductor and dielectric surfaces either by evaporation or sputtering in a vacuum chamber. Evaporation requires heating of the metallic source either by direct thermal means, using a resistive heater, or by an electron beam. Resistive heating melts a solid piece of metal placed on a coil, resulting in high temperature atoms impinging on the wafer; but this process provides non-conformal profiles, meaning that the side walls of steep surface structures are poorly covered (the metal impinges on the sample from essentially one direction, although tilting samples during metal deposition allows improved step coverage). Electron beam sources use an electron beam to melt the metal, or a metallic compound, in a crucible. Electron beam machines may have multiple metal targets that can be rotated into position without breaking the vacuum; thermal evaporators may have multiple sources placed side-by-side, and the current can be switched from one to another, to deposit a sequence of metals. Sputtering is a technique that uses the kinetic energy from inert atoms in a plasma, usually Ar, to evaporate atoms from a metal target, which will then deposit on the wafer surface. This method does not require high temperatures, thus reducing the amount of substrate or photoresist heating, and yields conformal profiles. In all thermal evaporation techniques, a shutter is placed over the metal source to prevent the metal atoms from reaching the wafers during the heating process.

Evaporation techniques are appropriate for *lift-off* applications, but inadequate for step coverage. Evaporators must operate at high vacuum, usually in the 10^{-6} to 10^{-7} torr range, while sputtering (being a plasma process) occurs at much lower vacuum. Calibrated crystal oscillators were used to monitor the deposition rate; they utilize the piezoelectric properties of quartz and the fact that the crystal resonance frequency is influenced by the accumulation of mass of the coating material on its surface.

Semiconductor/dielectric surface cleanliness is essential to ensure good metal adhesion and fabrication of reliable and reproducible interconnects and contacts. A final cleanup must be employed after the pattern is defined in resist and immediately before the sample is placed in

the evaporator. These cleanups in semiconductors may include a light etch or other procedures to remove any oxide layer present. Because of the high temperatures necessary to produce any appreciable atomic flux, the deposition rate must be chosen with care to ensure that the resist will not be heated to the extent that it would flow or become unremovable. If the metal application heats the resist to the flowing point (around 150 ^{0}C), the resulting pattern will be cracked or wrinkled.

Step coverage is a critical consideration for optoelectronic device metallization due to the sharp side walls of the waveguides and mesa contacts; even deposition conditions produce a dense structure on a flat surface, but yield a porous structure on the step side walls, and an open boundary or cusp emanating from the base of the step as a result of self-shadowing. The coating atoms approach a substrate interface along directions that are dependent on the geometric configuration of the deposition apparatus. In the case of evaporation, the atoms path is essentially line-of-sight from the source to the substrate. In the case of sputtering at elevated pressures, the atoms approach the substrate in line-of-sight from the point of the last collision, which is nominally one mean free path from the substrate (which is, at 50 mtorr, on the order of 1 mm), proving better sidewall coverage. The technique used to achieve conformal interconnect (CPW) line profiles was angled evaporation, where the sample holder was tilted with respect to the line-of-sight from the source to the holder.

Lift-Off [20][21]

Metal patterning in semiconductor processing usually employs metal lift-off, which permits a certain layout to be created with thin metallic films by photolithography, Fig. 4.9. The film to be patterned must be considerably thinner than the photoresist layer, which is applied to the sample, exposed, and developed to define the desired pattern; the metal film to be patterned is deposited next. The sample is then soaked in a solvent, usually acetone, which dissolves the resist. The film on top of the photoresist is therefore removed ("lifting off" the metal), but remains in the areas where the metal is in contact with the semiconductor. Unlike etching processes, the success of lift-off is highly sensitive to the edge profile of the patterned resist.

To obtain good lift-off, the photoresist, after being developed, must show an undercut at the edges, so the deposited metallic film brakes easily. A protruding lip, Fig. 4.9(b), at the top edge of the resist profile is desirable. A post-exposure bake tends to form a crust on the resist and provide a lip. This process works by retarding the action of the developer on the upper layer of the resist; lower layers are developed more vigorously, and result in an undercut profile. A well succeeded photolithography is generally sufficient to ensure good lift-off; however, the resist undercut can be further improved by chemically attacking the resist surface for an appropriate period of time, further reducing its sensitivity to light or to the developer, so that a less sensitive layer is formed at the surface of the photoresist.

For the S1818 photoresist ($\sim$1.8 μm thick, when spun at 4000 rpm for 30 seconds), the following procedure, after resist coating, gave good results: 1. baking at 90 ^{0}C for 15 minutes;

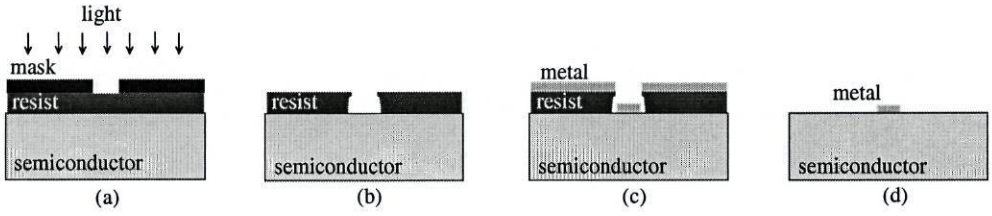


Figure 4.9: Lift-off process used in RTD-EAM fabrication.

2. exposure; 3. soaking in chlorobenzene for 15 minutes; 4. baking at 90 ^{0}C for 15 minutes. Soaking the sample in chlorobenzene removes residual solvents and the low molecular weight resin from the upper layers of the resist film; this retards the developing action in these layers, and therefore yields an undercut profile, Fig. 4.9(b).

Multi-layer Metallization [20][21]

Typical metal interconnects and contacts requirements in optoelectronics include high electrical conductivity, good adhesion, effectiveness as diffusion barrier or as etch stop (to reactive ion etching, for example), deposition easiness, and *lift-off* capability. In general, it is not possible to achieve all requirements with a single metal. In principle, high electrical conductivity can be met by gold, aluminium, silver, and copper. Copper, however, is a rapid diffuser, tends to getter onto crystalline defects, and exhibits high chemical reactivity.[8] Silver tends to be highly reactive, and is little used. Gold and aluminium have more desirable properties. Aluminium seems to be more attractive because it combines high electrical conductivity with good adhesion. However, aluminium has one serious disadvantage: it oxidizes or reacts easily and its electrical conductivity, although large, is not as high as gold. Gold is the dominant metal used, particularly with high current density regimes; gold wires (usually 25 or 12.5 μm in diameter) or gold ribbons are used to connect chips to the outside circuit. Gold also has relative chemical inertness, low melting point, is malleable, does not tarnish, but exhibits poor adhesion to GaAs, InP, and dielectrics.

Such conflicting requirements often lead to multi-layer metallizations consisting of two or more layers of different metals. Gold is usually deposited over another metal placed on the substrate/dielectric for adhesion; titanium or chromium exhibit good adhesion to almost any surface. However, under elevated temperatures or high current densities, gold diffusion can result in intermetallic compounds or alloys with gallium or arsenic; platinum has been found to be a good barrier metal, and is used between the titanium and gold layers; palladium has also been used for this purpose. The Ti-Pt-Au metallization, where the metals are sequentially evaporated without breaking the vacuum, has become very popular, being used for many purposes including p-type ohmic contact, power bus, gates, overlay metal, and transmission lines.

Gold-germanium based metallizations were found to be superior for contacting n-type GaAs and InGaAs, as well as InP, over variable doping ranges (1×10^{17} to 1×10^{19} cm^{-3}), demonstrating high reliability. They are usually applied with an overlay of another metal, such as nickel. Gold-

[8]There have been reports of successful application of copper in micro-electronics by IBM and others.

germanium is evaporated in proportions that represent an eutectic alloy (88% Au, 12 % Ge by weight), which melts at 360 ^{0}C. High quality ohmic contacts are obtained because germanium diffuses into the semiconductor and Ga diffuses towards the metal layer, creating a heavily doped ($> 10^{19}$ cm^{-3}) layer at the semiconductor surface, giving rise to a very thin contact barrier. Ohmic contacts in n-type GaAs, InGaAs, as well as in InP, are obtained with a sandwich of Au-Ge-Au-Ni-Au (again, the metals are sequentially evaporated without breaking the vacuum). Nickel acts as a wetting agent and prevents the Au-Ge metal from "balling up" during alloying, which could give rise to poor contact resistance, poor morphology, and very irregular edges. The alloying operation is usually performed using thermal profile programable rapid thermal annealler furnaces in the presence of an inert gas, such as helium or argon, or forming gas (15% hydrogen, 85% nitrogen). The alloying temperature is generally in the range 360-450 ^{0}C. Usually, the n-type metallization includes a thick gold overcoating, which enhances the sheet conductivity, results in improved surface morphology, also helps reproducible results from test probing of the device.

Catastrophic failure can occur in metallic conductors at sufficiently high current densities. Long before this occurs, however, metal migration (caused by electron scattering) can begin, resulting in metal depletion in one area and accumulation in an adjacent area. Once the metal cross section becomes diminished, resistance increases and catastrophic failure can follow.[9] (It should be noted that the thin-film properties of metals differ, in general, from those of the bulk material; electrical condutivity is lower.)

First Level Metallization [20][21]

Achieving high quality ohmic contacts in electronic and optoelectronic devices is very important, because even with the best presently known procedures, the associated parasitic resistance may be half of the total parasitic resistance, limiting device performance (the noise figure is particularly sensitive to such resistance) [10][11]. The requirements for ohmic contact formation are good conductivity and adhesion, chemical inertness and lack of undesirable interaction with the substrate over a long operating lifetime.

If the surface layer of the semiconductor, such as GaAs, is highly doped, almost any metal deposited on it will result in an ohmic contact without the need to be alloyed. This has certain advantages: the wafer does not have to be exposed to elevated temperatures; possible difficulties arising from alloying (irregular metal edges, poor surface morphology) are eliminated; and a complex process step can be skipped. Unfortunately, the necessary doping level is around 1×10^{19} cm^{-3}, which is not easily achieved by most growth or doping techniques. A metal in contact with a wide bandgap III-V semiconductor, with $n < 1 \times 10^{19}$ cm^{-3}, generally results in a rectifying contact, since a Schottky barrier tends to form. To overcome this problem, a multiple metal layer is usually required to self-dope the semiconductor surface through alloying. During

[9]MIL-S-19500F lists the maximum allowable continuous current density for metal conductors which are in thermal contact with a surface along their entire length (for gold, the maximum current density is 6×10^5 A/cm^2) [179].

the alloying and cooling period, a component of the metallization difuses into the semiconductor, thus creating a highly doped semiconductor layer in contact with the metallization.

Ohmic contacts to the RTD-EAM capping layers were achieved by e-beam depositing a Au-Ge-Au-Ni eutectic, followed by an extra layer of Au, which were patterned by lift-off. Metal stripes were formed on the sample epilayer regions to be defined as the waveguides and mesa emitter/collector contacts, Fig 4.10.

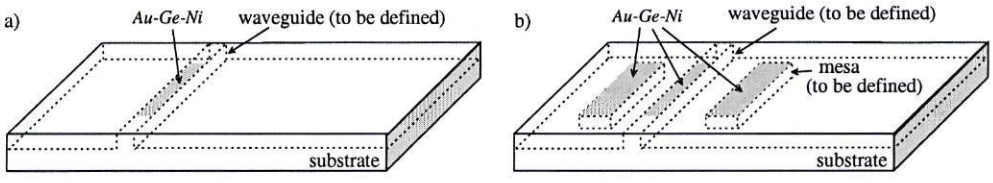


Figure 4.10: RTD-EAM ohmic contacts, without and with mesas.

The RTD-EAM multi-layer ohmic contact sequence consisted of the following thicknesses: 14 nm gold, 14 nm germanium, 14 nm of gold, 11 nm of nickel, and finally 240 nm gold capping layer. To ensure low contact resistance and ohmic behaviour, the GaAs ohmic contacts were alloyed at temperatures in the range of 360 - 390 ^{0}C in a neutral gas (N_2) atmosphere for 60 seconds. The epilayer, in the case of the InGaAs based RTD-EAM, was delta doped ($\sim 10^{19}$ cm^{-3}), making contact alloying unnecessary.

4.2.3 Etching

Etching techniques range from pure chemical removal, also known as *wet etching*, to combined physical plus chemical removal (such as reactive ion etching) and pure physical removal (such as ion milling) [20][21]; these lattest mechanisms are known as *dry etching* processes. The etchants allow removal of the material in a selective manner, once the resist has been patterned. An ideal etching process must be able to remove only the specified layer(s) of semiconductor from the selected regions. The etchant should neither attack the resist or the metals, nor penetrate under the mask causing undercuts; instead, it should attack only the layer(s) to be etched and be self limiting (e.g., should etch InGaAlAs and not InP). The etching solution must exhibit oxidizing characteristics and be free of metallic cations to avoid ionic contamination of the etched surface, which may produce undesired surface current leakage [11][20][21].

Reactive Ion Etching [20][21]

Dry etching techniques use plasma-driven chemical reactions or energetic ion beams to remove material. Most dry etching processes are based on sputtering by energetic ion or particle bombardment, evaporation of volatile compounds created through interactions with a reactive gas or, most frequently, through a combination of the two mechanisms. Etching proceeds by three steps: absorption of the necessary species on the material surface, relevant chemical reactions, and desorption of the products; any of these may be the rate-limiting factor. Dry etching tech-

niques usually yield almost vertical sidewalls regardless of crystallographic orientation, because vertical etching rates can be much higher than the horizontal ones.

Reactive ion etching (RIE) techniques employ radio frequency power to maintain a plasma between at least two electrodes. A schematic of a RIE reactor is illustrated in Fig. 4.11.

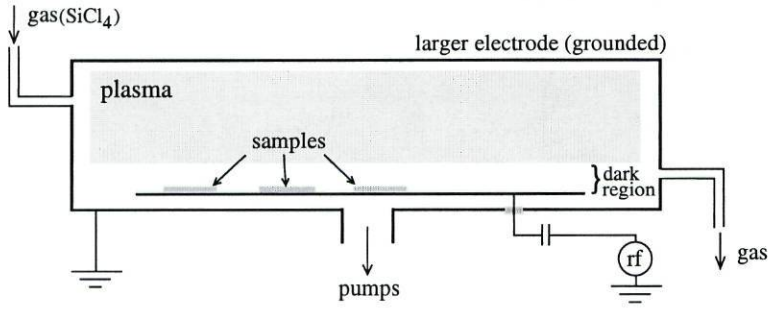


Figure 4.11: Typical parallel plate configuration of a RIE machine [21].

The electrode that holds the samples is usually the rf powered electrode (it can also be used to control sample temperature), so the chamber may be rf grounded. (Generally, the entire chamber is used as the second electrode.) The rf voltage is applied through a large blocking capacitor, so that no dc bias is intentionally applied. The plasma is excluded from the immediate vicinity of the electrode sample surface (which is called the plasma sheath or the dark region, Fig. 4.11, because it does not glow, being on the order of 0.1 to 10 mm thick, depending on operating conditions) by electromagnetic fields.

The plasma is generally neutral. However, the plasma sheath is a region of positive space charge and the electrode surfaces are negative with respect to the plasma; this is due to the higher mobility of the electrons, which move rapidly to the surface of the electrodes. The electric sheath-field formed at the powered electrode vicinity accelerates positive ions to the sample and enhances the etching rate by stimulating the formation and desorption of volatile products. The etched profile depends on the specific gas employed, its pressure (or gas flow) and the rf power applied. When the ratio of the electrode areas is large, most of the voltage drop will appear across the plasma sheath at the smaller electrode; a large asymmetry in electrode area will enhance ion energy and hence etching directionality. Directionality is also enhanced in RIE systems by operating at lower pressures (0.01-0.1 torr). The volatile products created are driven away from the surface by the gas flow. Higher pressure will result in undercutting, because the ions have a higher chance to collide with each other, gaining lateral momentum and hitting the side walls of the etched profile.

The interaction between experimental variables such as presssure, rf power, bias voltage (the sheath-field magnitude), flow rate, sample temperature, the resulting etching rate, anisotropy, and surface smoothness are very complex. Plasma chemistry can be complex, and a wide variety of chemistries has been studied [20][21]. Generally, the incoming gases supply a highly reactive species, such as O, Cl, or F. Etching of III-V semiconductors using Cl-containing gases has been very successful, as they etch spontaneously when exposed to Cl_2 or to Cl atoms created in an

upstream plasma.

Dry etching techniques, such as RIE, are able to produce anisotropic and non-selective etching (i.e., almost constant etching rate along different crystal directions and for different material compositions), yielding highly vertical side walls. Unfortunately, anisotropic etching has its downside, the damage that is almost inevitably imparted to the surface, bringing about degradation in the optical and electrical properties of the material. Dry etch-induced damage in III-V semiconductors takes the form of lattice damage and "chemical" damage, such as dopant passivation. Damage to depths of 100 nm to 200 nm can occur in GaAs when operating Cl-based RIE processes at low pressure (4 mtorr) and bias voltages between 280 V and 550 V. In many applications a final wet-etch removal of such damaged layers does not compromise dimensional integrity.

AlGaAs/GaAs Dry Etching [20]

An Oxford Plasmatechnology RIE 80 machine with silicon tetrachloride ($SiCl_4$) as the etching gas was used to define the AlGaAs/GaAs based waveguide and mesa structures. This machine was adapted to allow *in situ* monitoring of the etching depth, by addition of an optical system which detects changes in the reflected portion of a laser beam being shone on the sample during etch, due to different layer refractive indices. Comparing in real time the *reflectance* with the calculated reflectance of the structure being etched, precise control of the etching depth can be obtained. The rf power was kept at 100 W, the etching pressure at 12 mtorr, with a gas flow around 9 sccm, producing nearly rectangular and reproducible waveguide profiles at an etching rate of $\simeq$200 nm/min, with neglible material damage.

InGaAlAs/InP Dry Etching [21]

The standard dry etching process developed at the University of Glasgow is based on CH_4/H_2, using an Electrotech SRS Plasmafab 340 RIE machine, and requires a dc bias around 850 V. This is likely to introduce considerable material damage, such as hydrogen passivation of dopants and degradation of the material mobility. To avoid such damage, InGaAlAs/InP based RTD-EAMs were defined through wet etching.

Wet Etching [20][21]

In wet etching the sample is simply soaked in a liquid chemical solution that dissolves away the semiconductor. Wet etching has traditionally been used mainly because it is simple and inexpensive, and causes low levels of damage. Wet etching can be isotropic (giving rise to mask undercut) or crystallographic (due to differences in the atomic composition of the crystal planes); adequate dimensional and profile control can be difficult to achieve. Wet etching procedures are important for removal of damaged material, to form mesa and channel structures, to aid in polishing, as part of cleaning procedures, or as an integral part of diagnostic techniques. As part of device fabrication procedures, other materials, such as metals and dielectrics, can be also wet

etched.

A chemical etchant usually contains three basic constituents: an *oxidizing* agent, a *complexing* agent and a *dilutant*. There are different types of oxidizing and complexing agents, depending on the etching rate required and other factors, such as mask integrity, crystallographic orientation (the etching rate depends on the density of semiconductor atoms); the dilutant is usually water. The etching rate increases with temperature as the underlying physical phenomena (diffusion, surface adsorption) have an $\exp(-\Delta E_A/k_BT)$ dependence on temperature, where ΔE_A is their respective activation energy. The etching rate can also be changed by doping and by irradiation. Some etchants can produce bubbles, usually denoting hydrogen, which will prevent etching if left to adhere to the surface, causing nonuniform etching; constant agitation is usually employed to alleviate this problem.

Wet etching is not compatible with submicron technology and depends heavily on intrinsic material properties (such as preferential directions). This makes it unsuitable for devices where very sharp sidewalls are to be produced, as can be the case of optical waveguides. However, in this work, wet etching was used for channel waveguide and mesa formation in InGaAlAs/InP based RTD-EAMs. Wet etching was also employed as damage removal and surface cleaning in AlGaAs/GaAs and in InGaAlAs/InP based devices.

AlGaAs/GaAs [20]

GaAs etchants are usually either acids (hydrofluoric acid - HF; nitric acid - NHO_3; acetic acid - $H_4C_2O_2$; sulphuric acid - H_2SO_4; phosphoric acid - H_3PO_4) or alkaline solutions of ammonium hydroxide (NH_4OH). The most popular etching system for GaAs consists of sulphuric acid - hydrogen peroxide - water, which is often used for etch-polishing substrates and mesa etching. Oxides formed on GaAs tend to be amphoteric, and so can be dissolved in either acidic or basic solutions. Alkaline systems generally use NH_4OH in combination with H_2O_2 and H_2O.

The n^+ GaAs capping layer was removed, after waveguide formation (to minimize current spreading across the waveguide axis), using the following wet etching recipe: 1 ammonia - 1 H_2O_2 - 200 H_2O, giving an etching rate of 60 nm/min.

InGaAlAs/InP [21][184]

The most used etching solutions for InGaAlAs are sulphuric acid - hydrogen peroxide - water, hydrobromic acid - nitric acid - water, orthophosphoric - hydrogen peroxide - water. Orthophosphoric acid (H_3PO_4) based solutions do not etch InP; the H_2SO_4 and H_3PO_4 systems etch InGaAs/InAlAs [001], forming an angle of 55^0 with the $[1\bar{1}0]$ direction (in the [110] direction, the angle obtained is $\simeq 115^0$). For device applications, the appearance of large angles may result in difficulties in the technological steps that follow chemical etching, such as surface passivation or dielectric coating, and metallization of the structures; mask undercutting can also be significative. For this purpose, the best results, when selective etching is not required, are obtained with the Br_2-CH_3COOH etching system.

The InGaAlAs waveguides and mesas were defined using a InP selective wet etching solution: 3 H_3PO_4 - 3 H_2O_2 - 200 H_2O, giving an etching rate of 60 nm/min.

Surface Cleaning [20][21]

Surface cleaning before metallization is necessary to remove contaminants (such as carbon) or native oxides that begin to grow immediately as the semiconductor sample is exposed to air. Although no compound semiconductor oxidizes as readily as silicon, the few nanometers of oxide that are typically present on a III-V semiconductor surface prior to any cleaning can be a major obstacle to a good ohmic contact.

The exposed surface of the compound semiconductors used in this work (AlGaAs/GaAs and InGaAlAs/InP) was treated with diluted HCl (1:4 HCl:H_2O), after the metallization lithographic step involving the chlorobenzene soak is completed, to remove any native oxides and other depositions formed during previous fabrication steps. The samples were next rinsed in deionised water, blow dried and left in the oven at 90 ^{0}C for 5 minutes.

Gold Etching [20][21]

The first devices employed stripes of Au-Ge-Au-Ni-Au ohmic contact metallization as waveguide etching mask. Standard etches for gold are based on potassium iodide and iodine (iodine crystals can be dissolved in a saturated solution of potassium iodide and give an excellent etch). The recipe 200 ml deionised water + 340 g potassium iodide + 7 g iodine, etches the n-type ohmic contact metallization with an etching rate of 1 μm/min; it also etches GaAs.

SiO_2 Etching [20][21]

When it was necessary to remove the SiO_2 film used for device passivation and electrical insulation, the standard isotropic etch for silicon dioxide (a hydrofluoric acid (HF) solution) was used: a diluted HF (1:4 HF:H_2O) solution typically achieves an etching rate around 1 μm/min.

RTD-EAM Waveguide Formation

The sample was resist coated and patterned using the waveguides mask to form the channel waveguides etching mask, which consisted of sets of resist stripes of different widths (2-6 μm), spaced by 500 μm, and 10 mm long. Figure 4.12 shows a portion of the sample corresponding to a typical device. The mesa metal contacts acted as etching masks.

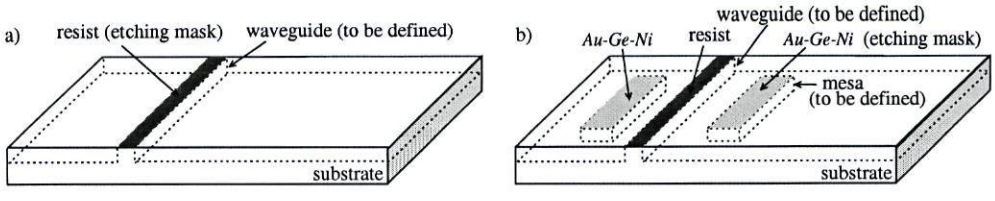


Figure 4.12: RTD-EAM waveguide pattern configuration.

The sample unprotected epilayer was then etched down to form channel waveguides and

mesas, by means of wet chemical etching (InGaAlAs) and/or reactive ion etching (AlGaAs). Figure 4.13 presents a scheme of the etched RTD-EAM waveguide.

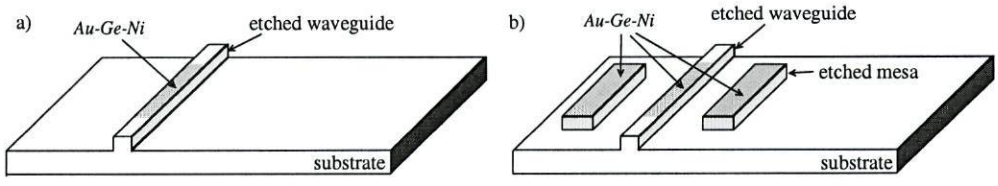


Figure 4.13: Scheme of etched RTD-EAM waveguide.

4.2.4 SiO_2 Passivation and Second Level Metallization

Thin dielectric films grown by plasma-assisted processing are used for encapsulation, scratch and environmental protection, surface passivation (to reduce leakage current), as metallization interlevel dielectrics, as capacitor dielectrics, as transfer layers in multilevel resist procedures, and as etching or ion implantation masks. These dielectric films should adhere well to the semiconductor, be chemical and physically stable. An insulator film should have high electrical resistivity, typically $>10^{12}$ Ωcm, and its impurity content should be low; structurally, the dielectric should be amorphous, devoid of pinholes, its surface morphology should be smooth and its dielectric function independent of frequency, and possess a high breakdown voltage.

The dielectric materials mostly used in semiconductor processing are silicon dioxide and silicon nitride. The low dielectric constant of silicon dioxide makes it a popular choice for use as an interlevel spacer to separate metal crossovers with minimum parasitic capacitance.

Plasma-Enhanced Chemical Vapour Deposition [20][21]

A common silicon dioxide (SiO_2) deposition technique is the plasma-enhanced chemical vapour deposition (PECVD): the plasma generates chemically reactive species, and the temperature of the atoms, molecules, or radicals produced, as represented by their translational and rotational energy, is generally near room temperature. The electrons, however, can exhibit temperatures of tens of thousands of kelvin, sufficient to break molecular bonds and create chemically active species in the plasma. Any of these species can be excited to higher electronic energy states by further interaction with the electrons; hence, chemical reactions that usually occur only at high temperature can be made to occur at low temperature or even room temperature in the presence of an activating plasma state. The regime of interest to semiconductor processing is the low-pressure plasma or glow discharge. These plasmas are characterized by gas pressures on the order of 0.1 to several torr, free-electron densities of 10^9 to 10^{12} cm^{-3}, with energies of 1 to 10 eV. Most of the species remain neutral in glow discharges (the degree of gas ionization is typically less than 10^{-5}).

The major advantage of PECVD is the ability to grow films at relatively low substrate

temperatures (compared to chemical vapour deposition, CVD),[10] usually well under 300 ^{0}C. This is obtained at a price: compositional control of the thin film material is difficult. PECVD yields amorphous films with very little short-range structural ordering. Chemical bonding within the film may vary, with implications on the etching rates. These thin films can easily have "pinholes" (often on the order of 1 μm) that expose the underlying material, caused by particles present on the sample (or formed during deposition, if contaminant gases are present). It is necessary to clean reactors regularly, as thin film materials will tend to deposit on the electrodes and on other parts of the chamber. These materials are most easily removed by operating the reactor in an etching mode, using appropriate gases and higher power.

A silicon dioxide film was formed on the RTD-EAM samples, as schematically shown in Fig. 4.14, by plasma assisted deposition (PECVD) or sputtering below 300 ^{0}C. Silicon dioxide film thicknesses between 500 and 1000 nm were employed; because of the low deposition temperature, these films can be applied over the completely fabricated devices, and can serve as their final environmental passivation/encapsulation.

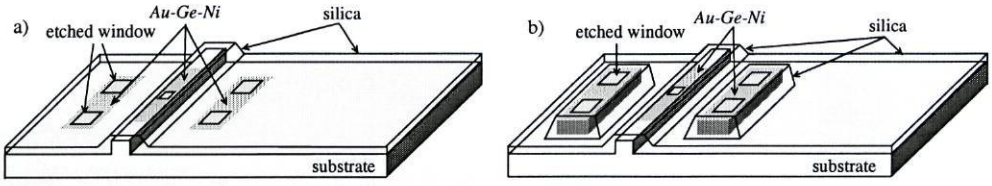


Figure 4.14: RTD-EAM SiO_2 passivation/insulation, showing access contact windows.

Etching of Via Holes [20][21]

Via holes (access contact windows) through the silicon oxide film allow electrical continuity between different levels of metal interconnects at appropriate locations, where necessary. A C_2F_6 plasma dry etching was used to etch SiO_2 contact windows in order to allow access to the mesas and waveguide ohmic contacts. At a rf power of 100 W and a gas flow rate of 20 sccm, the etching rate was about 50 nm/min. Figure 4.14 shows schematically the etched contact windows over the ohmic contacts.

Second Level Metallization [20][180]

Most high speed test fixtures require that individual devices are to be mounted and electrical connections made with bonding wires. Advances in high speed monolithic circuits generated the need to perform tests on the processed sample and avoid the cost of placing inacceptable devices in fixtures. Also, for characterization and modelling of active devices for subsequent circuit insertion, it is advantageous to have the measurement port as close to the device as possible [10].

[10] Chemical vapour deposition of dielectric films is not used because of the high temperature required ($\sim$ 500 ^{0}C), which is incompatible with GaAs/InP materials (leading to decomposition) and ohmic metallizations used.

The idea was to fabricate a controlled impedance transmission line extending to the contact point of the probes, allowing at the same time easy connection to the outside world.

Figure 4.15 presents the high frequency contact pad (50 Ω CPW transmission line) lithography pattern. The Ti-Pd-Au metallization was sequently evaporated without breaking the vacuum: 33 nm of Ti + 30 nm of Pd + 240 of Au.

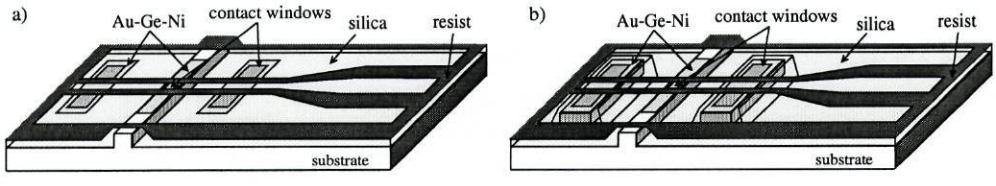


Figure 4.15: RTD-EAM high frequency bonding pads lithography.

The CPW transmission line, with 50 Ω impedance preserving tapers as shown in Fig. 4.16, was formed. (Figure 4.16 presents the top view picture of a fabricated RTD-EAM, showing the coplanar waveguide transmission line.) It also serves as high frequency bonding pads and power bus. In order to ensure low series resistance, the metallization thickness is made considerably thick by deposition of an extra Au layer, being typically 0.7 - 1.0 μm. An angled metallization was employed to improve step coverage on the waveguide wall of the central CPW track.

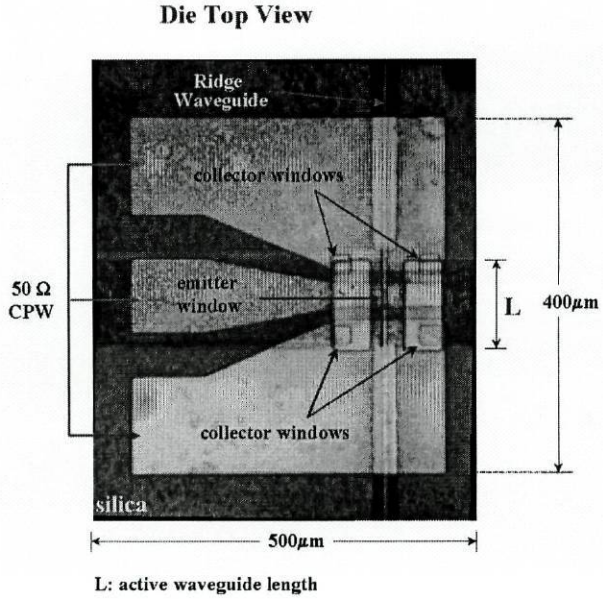


Figure 4.16: RTD-EAM die top view, showing the CPW contact-pad/transmission-line.

On-sample probing was then relatively easy with the sample keeping its full thickness, because the formed CPW does not need a backside ground metallization; indeed, the RTD-EAM does not require back side metallization.

4.3 Post Processing and Packaging

No less important than the main aspects of device fabrication is the transformation of a sample containing many devices into a large number of packaged devices, ready for testing. The small size of devices, their fragility and the need for optical alignment do not propitiate fully automated procedures.

4.3.1 Back-Sample Lapping

Substrates of III-V semiconductors are made typically 2-3 times thicker than a silicon wafer of the same diameter, because of the much higher fragility of the compound semiconductor materials. This thickness, whilst necessary to limit the number of breakages during processing, may not be desirable in the finished device. In this work, the purpose of sample thinning was mainly to facilitate device separation (dicing) and improve characteristics such as heat dissipation (reduce thermal resistance); however, thermal gradients across the surface of the chip will become more severe with sample thinning.

The extra thickness was removed by lapping the back of the sample (great care must be taken not to damage the devices). The sample was glued to a glass slide, epilayer down, with photoresist, which also protected the epilayer, and baked at 90 ^{0}C for 15 minutes. The glass slide was then mounted substrate up on a stainless steel jig with dental wax, avoiding spilling any wax on the exposed face (substrate) of the sample. The substrate was hand lapped on a grinding plate, in a liquid suspension of an abrasive powder, until its thickness was $\simeq$150 μm. Then, the glass slide was removed by heating the jig and melting the wax; the sample was removed from the glass slide using acetone, thoroughly rinsed in deionised water and dried, being ready for device cleaving and separation.

4.3.2 Device Separation

The device separation procedure typically uses mechanical scribing or sawing, although laser scribing techniques have sometimes been used on an experimental basis. Device scribing takes advantage of the crystallographic nature of the semiconductor sample, its purpose being to weaken the structure along one cleavage plane. In this work, a home designed mechanical scribing technique was used.

There are difficulties with this technique, as a fine balance has to be determined when scribing the sample, so as to introduce enough surface damage to weaken the structure but not damage the device. If the sample is not weakened sufficiently, the subsequent cleaving process may result in small device yield. This is specially true for this RTD-EAM device because the guiding region is not embedded into the semiconductor, but instead corresponds to a protuberance on the sample surface, as shown in Fig. 4.17. In fact, this was one of the most critical frabrication steps due to the need to obtain good quality device facets, because of the high risk of waveguide facet destruction in the scribing process.

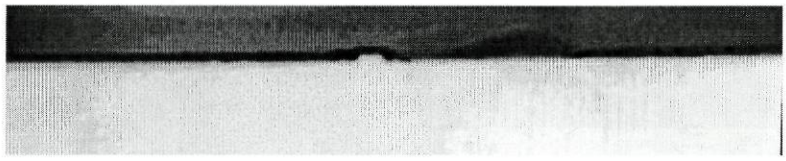


Figure 4.17: Side view picture of a RTD-EAM ridge waveguide (ridge: 1.4 μm).

The samples were placed on adhesive foil, substrate side down. A diamond scribe was then applied on the sample edge area, left clear, along a scribe plane parallel to the waveguides, cutting slightly into the sample. The scribe was then stepped and repeated in a parallel plane between waveguides, spaced 500 μm apart.

The sample was broken into individual stripes by covering the scribed sample with a scratch protection foil, and applying pressure on the foil with a razor blade edge. Stretching the foil separated the stripes even further. The scratch protection foil was removed and the stripes were then picked from it and placed on a new adhesive foil, and then individually scribed along planes perpendicular to the waveguides, spaced by $\approx$500 μm. The stripes were broken into individual devices, Fig. 4.16, using the same procedure. The scribe process creates devices with perfect crystal facets, which act as good end mirrors in a semiconductor waveguide. The dies ($\approx$500 μm long, $\approx$500 μm wide), Fig. 4.16, were then picked up from the foil, being handled with sharp tweezers and avoiding contact with the mirrored facets perpendicular to the waveguide.

The optical quality of the facets perpendicular to the waveguides was then inspected using a die bonding commercial equipment (EDB 80 Eutectic die bonder, Dynapert-Precima Ltd.), containing a pick-up arm which can be rotated through 180^0 for die inspection (the four facets of the die could then be examined). Good quality devices went through electric testing to verify again their NDC characteristics.

4.3.3 Package and Device Connection

Devices showing good facets and having appropriate dc current-voltage characteristic were mounted on a suitable test package. Factors to be taken into account in this process are orientation of the RTD-EAM waveguide to ensure good alignment to a fibre (where needed) or a lens, and the possibility of heat sinking. Of principal interest to high speed operation is the design of a package with reduced parasitics: a good common ground plane is essential, bonding wire bends should be minimal to eliminate stray inductances, and multiple wire bonds should be used to minimize resistance.

Test Package

The package should provide good electrical and thermal contact between the die back-side and the metal holder, and must be designed to allow light coupling into the device through microscope objective end-fire (or optical fibre arrangement) in an optical set-up using micro-blocks. Figure

4.18 presents a picture of the RTD-EAM package implemented. Brass is used in the package on account of its excellent thermal conduction properties and ease of machining.

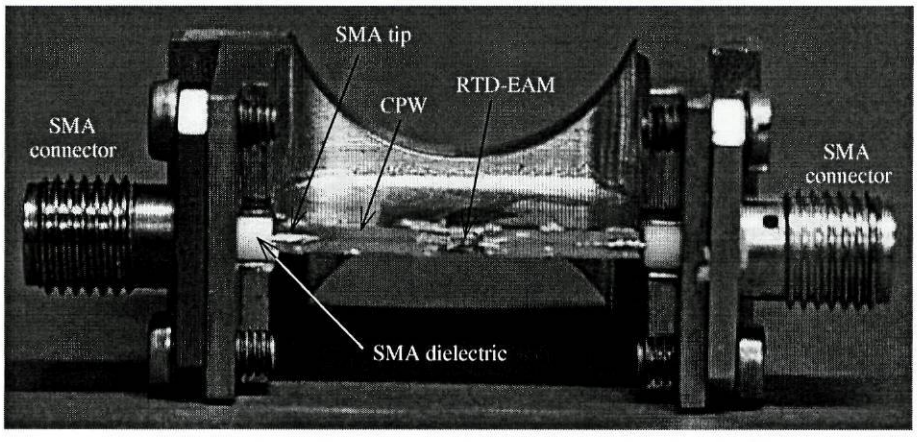


Figure 4.18: Photograph of the RTD-EAM package implemented.

CPW Chip Bonding and SMA Connectorization

At high frequency, the package is an integral part of the overall circuit design. Coplanar waveguide (CPW) transmission line chips with 50 Ω impedance, implemented on SI GaAs, were mounted on the optical test package and were used to connect the RTD-EAM die to the outside of the package, as shown schematically in Fig. 4.19. Mounting was performed with the CPW contact pads face up for subsequent wire bonding, using indium solder at a typical temperature of 175 ^{0}C. A SMA connector was then attached to the test package and connected in line to the package CPW line, Fig. 4.19. The SMA connector limits the package bandwidth to around 18 GHz; however, the RTD-EAM can still be rf probed while in this fixture.

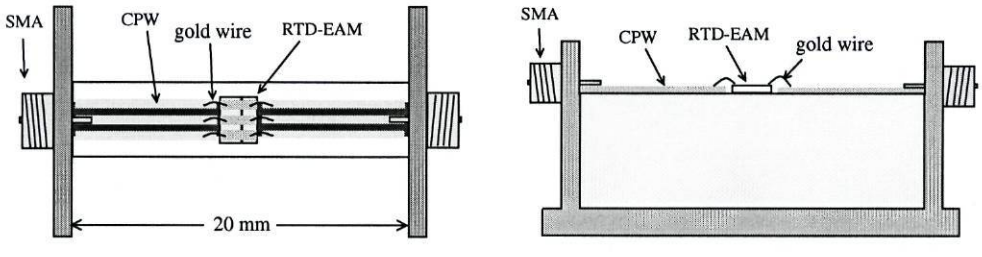


Figure 4.19: Schematic representation of a packaged RTD-EAM device.

RTD-EAM Chip Bonding

The RTD-EAM die was bonded by attaching it, substrate down, to the package pedestal with a high conductivity adhesive, Fig. 4.19, using a die inspection and placement machine (EDB 80 die bonder). When the die was properly aligned relative to the package, the pick-up arm was lowered to bring the die into contact with the package. Silver epoxy glue, cured at 100 ^{0}C for

5 minutes to provide good adhesion, was used instead of indium because any amount of indium that overflows under the facets or contacts will make the device useless; besides, silver epoxy glue, if not yet cured, can be easily removed using acetone.

Wire Bonding [20][21]

The final step in completing device preparation was the connection between the device high frequency pads and the package CPW transmission line, as represented in Fig. 4.20. The connections were realized with gold wire, about 25 μm diameter, that linked the package terminals to the microscopic metal contact pads on the device chip. The process was carried out under a microscope, using a commercial ultrasonic wire bonding machine (Wedge Bonder, Model 4123, Kulicke and Soffa Industries, Inc.).

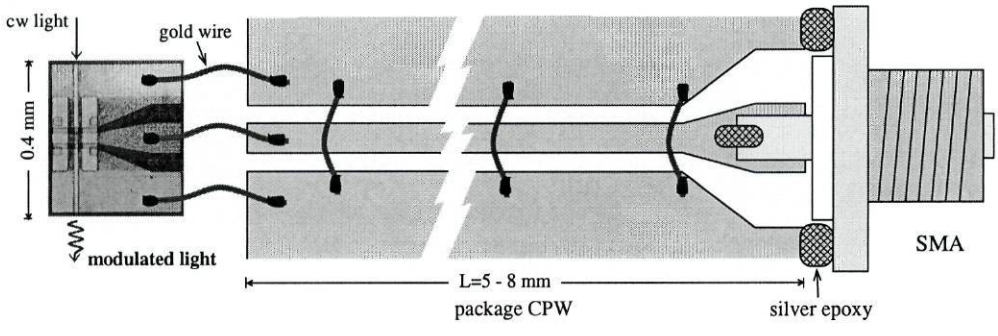


Figure 4.20: Schematic of the RTD-EAM and SMA connection to the CPW package.

There are several factors involved in creating a good bond, such as ultrasonic frequency, time, temperature, surface cleanliness and the metals used. Failure of the connection is usually due to underbonding (insufficient contact), overbonding (too much contact, leading to squashing) or even lifting off the pad contact.

4.4 Electrical Characterization

Virtually every processing step can alter the electrical properties of the device and, consequently, process development or modification must be closely coordinated with electrical characterization.

The capacitance-voltage (C-V) technique [55] was employed to profile the doping concentration $N_d(z)$ of the wafers grown, when justified.

4.4.1 dc and Low Frequency Testing

The Hewlett-Packard Semiconductor Parameter Analyser probe station (model 4145A with a residual resistance in voltage mode of 0.4 Ω and capacitance load $\leq$1000 pF) was used to characterize the dc electrical behaviour of devices. This is a fully automatic, high performance, programmable test instrument designed to measure, analyse, and graphically display the dc characteristic of a wide range of semiconductor devices. On-sample measurements at dc and low

frequencies use needle probes to make contact to the appropriate device terminals; such measurements are used either to monitor the various processing steps or to screen devices before fixturing. The inductance associated with the needles makes them unusable at higher frequencies.

Typical current-voltage characteristics of GaAs and InGaAlAs based RTD-EAMs are shown in Figs. 3.5 and 3.9, respectively. From them the voltage dropped across the waveguide core depletion region is estimated, hence the electrical field and the resulting depth of optical modulation. The RTD-EAM switching time, t_R, can also be estimated from its I-V characteristic: $t_R = 4R_dC_d \simeq 4\,(\Delta V/\Delta J)\,C_v$ [152], where ΔV and ΔJ are the peak-to-valley voltage and current differences, respectively; C_v is the capacitance at the valley point per unit area ($C_v = \epsilon/W$, ϵ is the depleted region permitivity, and W is the depletion region width). The NDC characteristic gives also a measure of the quality of the DBQW structure.

4.4.2 High Frequency Testing

Basic measurements of voltage and current are extremely difficult at microwave frequencies (as high as 30 GHz and beyond) in terms of accuracy and stability [3][10]: high gain microwave devices may not be stable if their terminals are shorted (i.e., they may oscillate); to obtain accurate short or open circuits at the input or output ports at microwave frequencies is rather difficult, because appreciable impedance can result from very small parasitic capacitances and inductances; for example, a shorting wire or strip across the output terminal of a device, even if a few tens of microns long, will have a nontrivial impedance at microwave frequencies. Conversely, an open circuit formed by simply leaving the terminals open can easily have nontrivial capacitance and impedance. In the mid-1980s, microwave transmission line probes were introduced that bring the test ports for high frequency measurements directly to the contact points on the sample [10]; devices to be tested by such probes need to have input and output ports that accommodate the contact point of the probe (foot-testing) [3].

Wave scattering parameters (S-parameters) measurement [10] circumvents the need for true shorts and opens at the terminations, and also avoids the stability problem as it is based on signal power ratios. In this case, the device under test (DUT) is connected to transmission lines that extend to an impedance termination on the output side and a signal source on the input side [3]. The source and impedance termination can be at some distance from the DUT (their impedance is usually 50 Ω). When the signal source is turned on, there will be incident and reflected voltage waves which can result in a standing wave pattern [3]. Calibration steps are needed to establish the phase reference planes and to map the measurement space to the Smith chart[11] of the device [10][179], for correct calculation of the phases and to extract any components of the reflected signal due to imperfections in the hardware, such as reflections from connectors, bends, finite cross talk between isolated channels, imperfect matching of the signal source, and

[11] The Smith chart gives a graphical transformation between the reflection coefficient and impedance [179]. The complex reflection coefficient $\left(S_{11} = \rho e^{i\phi}\right)$ is represented in polar coordinates [10]. Superimposed on that are circles and lines of constant real and imaginary parts of normalized impedance ($Z = R + iX$). It provides a graphical tool for designing distributed matching networks [179].

so on. Calibration for S_{11} might involve measurement of an open-short-standard load circuit to establish the reference phase plane and give enough information to map three points on the Smith chart (S_{11} =1, -1, 0, respectively) [10]. This allows the removal of most errors associated with imperfect directivity and mismatch.

High frequency characterization of the RTD-EAMs used a Network Analyser, which can measure the complex value (magnitude and phase) of the S_{11} parameters at selected frequencies, by harmonic down conversion and heterodyne detection [10]. A source providing a stable signal at precise frequencies (synthesized source) supplied the incident signal to the RTD-EAM over a broad frequency range. The RTD-EAM exhibits dynamic negative conductance to the measurement system, so the reflected power is greater than the incident power when it is operating at a frequency lower than the cut-off frequency, i.e., $|S_{11}\,(f < f_c)| > 1$; from the measured frequency dependence of the S_{11} parameter, the device equivalent circuit parameters and bandwidth can be estimated.

4.5 Optical Characterization

Optical characterization of GaAs based RTD-EAMs employed light from a Ti:sapphire laser (cw Ti:sapphire laser, model 3900, Spectra-Physics),[12] tuneable in the wavelength region around the absorption edge of the GaAs waveguide (850-950 nm) with a linewidth around 0.1 nm (40 GHz) [185]. Light from a Tunics diode laser, tuneable in the wavelength region 1480-1580 nm, with a linewidth below 10^{-5} nm (100 kHz) [186], was used to characterize InGaAlAs based RTD-EAMs.

RTD-EAM optical characterization included observation of the propagation modes, followed by determination of propagation loss. In studying their application as intensity modulators, the electric field induced absorption change was measured, followed by high frequency optoelectronic characterization. Optical characterization also included determination of InGaAlAs waveguide core bandgap energy, employing photoluminescence techniques.

4.5.1 Waveguiding Characterization

The most efficient method to couple light into channel waveguides is through focusing, with a microscope objective or a lensed single-mode optical fibre, into a waveguide cleaved end (endfire coupling).[13] (Two other methods of efficiently coupling light into a channel waveguide are through gratings and tapers [13][58].) The lenses/fibres and the packaged waveguide were mounted on micro-block translation stages, in an endfire arrangement [58]. The input and the output could be both independently adjusted to optimise focusing and alignment, and the sample could also

[12]Titanium doped sapphire is a solid-state laser medium capable of tuneable laser action over a broad range, near infrared wavelengths. The broad absorption band in the blue and green makes the argon ion laser an ideal pump source for this material. The laser linewidth was < 40 GHz, the spatial mode was TEM_{00}, and the polarization was horizontal (100:1)

[13]The lensed fibres were prepared by fusing the cleaved end of a single-mode optical fibre, using a commercial fusion splicer.

be translated and rotated independently of the lenses/fibres. The endfire experimental set-up employed is schematically represented in Fig. 4.21.

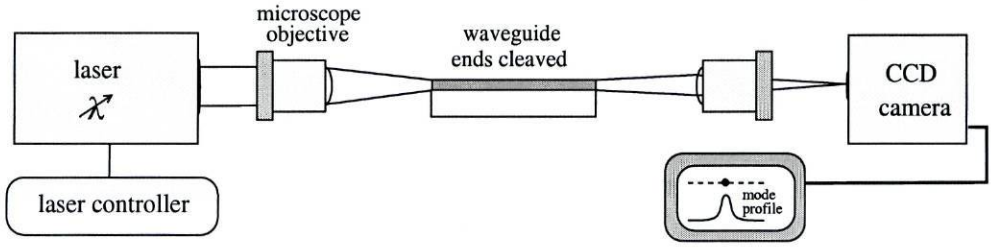


Figure 4.21: Experimental apparatus for observation of optical waveguiding [58].

With GaAs based RTD-EAMs, lenses (×20) were used instead of fibres due to the higher power coupling achieved; lensed monomode fibres were employed to couple light into the InGaAlAs RTD-EAMs. The output light was collimated/focused by a microscope objective lens and sent to an infra-red camera. Input coupling is critical, as is locking of that alignment once it has been determined. The infra-red camera line scan facility was used to monitor the intensity, to observe the optical output power distribution and hence to maximize coupling.

4.5.2 Loss Measurement

Loss sources in waveguides are generally attributable to three mechanisms [58][69]: absorption, scattering, and radiation leakage, Fig. 4.22 (Fig. 2.23, section 2.2.4).

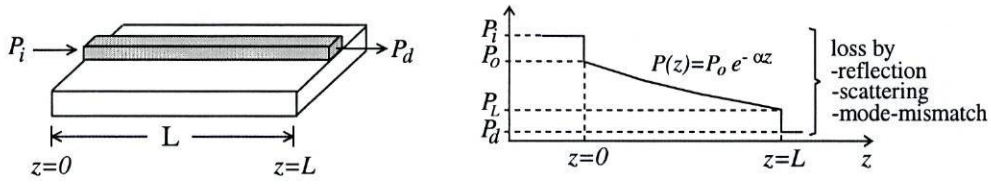


Figure 4.22: Optical power attenuation in the waveguides.

To measure waveguide loss, a known optical power is introduced into one end of the waveguide and the power emerging from the other end is measured [58][177]; the output power from the waveguide can be written as

$$P_L(z = L) = P_0(z = 0) \exp(-\alpha L), \tag{4.7}$$

where P_0 and P_L are the input and output powers, respectively, within the waveguide of length L; α is the attenuation coefficient.

There are many problems and inaccuracies inherent in using such approach [58][177]: coupling losses at the input and output are generally not known, and can largely obscure the true waveguide loss. Also, if the waveguide is multimode, loss attributable to individual modes cannot be separately determined. (Direct lens endfire coupling excites continuum modes as well

as guided modes, which complicates the interpretation of results [177].) The proper choice of a measurement technique depends on the type of waveguide being used, on the predominant loss mechanism, and on the magnitude of the loss being measured [58][177]. In this work, sequential cleaving was used to measure AlGaAs RTD-EAM waveguide loss [58]. The Fabry-Pérot method was adopted to estimate the loss of the InGaAlAs RTD-EAM waveguides [177].

The sequential cleaving method is not suitable for propagation loss measurements of low-loss (<1 dB/cm) semiconductor waveguides, as it normally exhibits large uncertainties [58]. In this case, the transmitted output power will depend simultaneously on the waveguide loss and on the Fabry-Perot étalon effect [177], since the reflectivity of the cleaved facets is as high as 0.3. Another disadvantage of this method is that it is destructive. Also, no direct discrimination is made between the losses of different modes.

In the Fabry-Pérot method, loss is determined by changing the optical path length of a waveguide cavity [177], for example by modifying the waveguide temperature, and measuring the light output as the temperature changes. The Fabry-Pérot method cannot, in principle, be applied to multimode waveguides because the superposition of the response functions of the individual modes yields a reduction of the measured overall contrast. However, it is possible to excite a single mode by fibre endfire coupling [58], which can relax this constraint; an optimized experimental setup with a polarization maintaining fibre coupler is also expected to reduce the measurement errors [177]. If the light source coherence length is too short, the Fabry-Pérot resonances cannot be perfectely resolved and the measurement will predict a higher loss; in order to make an accurate measurement of loss, the laser linewidth must be substantially narrower than the spectral range of the waveguide étalon, $\Delta\nu = \frac{c}{2n_{eff}L}$, where n_{eff} is the effective refractive index of the guided mode and c is the speed of light [177].

4.5.3 Waveguide Transmission Spectrum Change

Knowledge of the waveguide transmission spectrum at diferent bias points allows definition of the operating wavelength which gives the maximum modulation depth. Figure 4.23 shows a schematic representation of the experimental set-up used for measurement of the waveguide optical absorption spectrum change, induced by the Franz-Keldysh effect.

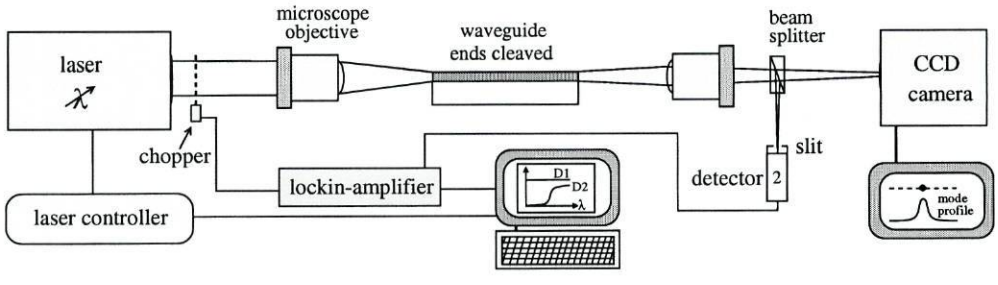


Figure 4.23: Experimental set-up used for bandedge shift measurement.

The RTD-EAM waveguide transmission spectrum was determined with zero applied voltage,

with the device dc biased at slightly below the peak voltage, and in the valley region; the device was not dc biased in the NDC region in order to avoid self-oscillation. The change in the optical absorption spectrum induced by the peak-to-valley transition was also measured: the RTD-EAMs were dc biased at slightly below the peak voltage, and a low frequency rf signal was injected to induce switching into the valley region.

4.5.4 High Frequency Characterization

Typical techniques and hardware for high speed optical detection and short optical pulse characterization can be divided into two classes, based on their response time. For pulse widths longer than 15 ps or frequencies below 60 GHz, solid state photo-detectors offer a combination of small size and ruggedness, and can be used with sampling oscilloscopes [3][10]. On shorter time scales, the characterization techniques are currently confined to the laboratory environment, and rely on streak cameras (that can resolve pulses longer than 1 ps), or autocorrelation techniques,[14] that can be used to analyse pulses as short as a few femtoseconds.

A Hamamatsu C5680 streak camera, with a minimum time resolution of around 2 ps, was employed to characterize the RTD-EAMs optical response at around 900 nm of rf modulated optical signals. A InGaAs Schottky photodetector was employed to characterize RTD-EAM operation at around 1550 nm.

Fast Photo-Detector

At the present time, solid state detectors offer a time response slower than streak cameras and autocorrelators; commercial devices with bandwidths up to 60 GHz and rise times of less than 10 ps are available from several commercial vendors [60][188]. While design details vary among manufacturers, fast response times are achieved by optimising the detector structure with respect to a number of important variables. Principal among these are the decay time for electrons and holes generated by input light pulses and the capacitance of the detector structure and associated circuitry. In general, it is obviously desirable to minimise both the carrier lifetime and the capacitance. Using a material in which the carriers have short lifetime and high mobility, in combination with an electrode design that minimises transit-time effects and capacitance, a bandwidth of 375 GHz was achieved in a laboratory prototype [187] and devices of similar design are currently being manufactured [188].

The photodetector employed to characterize RTD-EAM devices was a Schottky InGaAs (12 μm diameter) photodetector with the following characteristics [60]: wavelength range 950-1650 nm, typical 3-dB bandwidth 45 GHz, rise time 9 ps, maximum conversion gain 10 V/W, typical maximum responsivity 0.4 A/W, output impedance 50 Ω, cw saturation power 1 mW, maximum pulse power 100 mW, optical input single-mode ST or FC-fibre-optic connector, electrical output Wiltron K connector.

[14] In addition to simple intensity autocorrelation, advanced techniques have been developed that completely describe the time evolution of the electric field in a short pulse [190].

Streak Camera [189]

Based on photo-electron emission in an image-converter tube, streak cameras produce an output image in which time is represented as a linear displacement. This is accomplished by rapidly scanning the electron beam inside the image-converter tube. The input optical system of a typical streak camera images light entering the camera into a narrow slit on the photocathode of the streak tube. While tightly focusing the input beam along one axis, the input optics has no effect on the intensity distribution in the orthogonal direction. As a result, the camera can accurately reproduce the spatial distribution of light in the exit plane of the spectrograph or along one axis of a spatially inhomogeneous source. Figure 4.24 shows the operating principle of the streak tube used in the C5680 Hamamatsu Universal Streak camera.

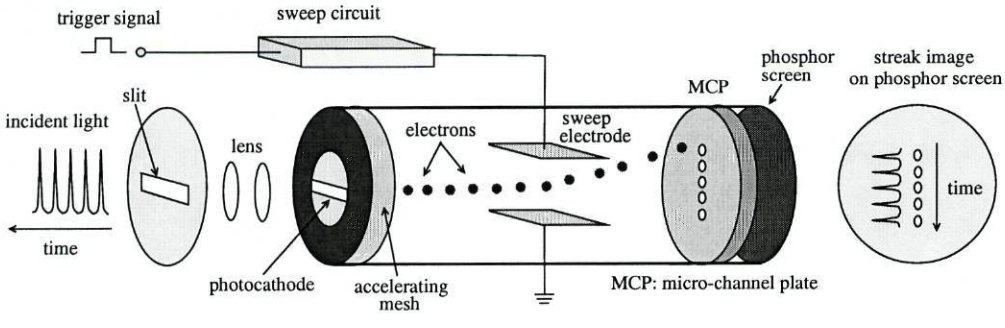


Figure 4.24: Operating principle of the streak tube [189].

The light impinging on the photo-cathode produces a narrow beam of electrons that are accelerated by a grid electrode towards the output end of the tube (phosphor screen). Deflection/sweep electrodes, located after this grid, are used to rapidly scan the electron beam across the input face of a micro-channel plate (MCP) near the output end of the tube. This element amplifies the electron current while preserving its spatial distribution. A visible image is generated by a phosphor screen at the output end of the micro-channel plate.

Applying a short, high voltage pulse to the deflecting/sweeping electrodes, the electrons released by the photo-cathode are scanned across the output screen in a few hundred picoseconds. Because linear deflection of the tube image is proportional to time, the streak image on the output phosphor screen is a record of the temporal evolution of the arrival light intensity, i.e., in the direction perpendicular to the scan direction, the output image faithfully records the variation of the incident intensity on the photo-cathode.[15] In those cases where the sample "emission" is periodic, repetitive scanning is used to increase the camera sensitivity.

In modern instruments, the two-dimensional output images are digitally recorded using a detector array for subsequent analysis. In a typical commercial device, a time resolution of 2 ps can be achieved when operating in a single-scan mode. Ultimately, the time resolution is limited

[15] In a typical application, a streak camera can also be used to record the time evolution of the spectrum from a pulsed sample. By using a streak camera in place of a photographic plate or optical multichannel analyser in the output plane of a spectrograph, it is possible to observe variations in the source spectrum during the camera scan period.

by space-charge effects within the tube, that make the electron transit-time dependent on the input intensity.

The C5680 streak camera employs a S-1/borosilicate glass streak tube (spectral range 300-1600 nm), with radiant sensivity of $\approx 10^3$ μA/W at 900 nm and $\approx 10^{-1}$ μA/W at 1550 nm [189]. Because of the low power available from the laser and the low radiant sensivity of the borosilicate glass streak tube at 1550 nm, it was not viable to employ the streak camera to characterize RTD-EAMs operating at 1550 nm.

4.6 Final Comments

Several aspects to be considered in the realization of RTD-EAM devices were discussed, including design of the wafer structure for good integration of the DBQW-RTD within the semiconductor waveguide, and layout of high frequency coplanar waveguide transmission lines, employed as device bonding pads and for testing the package. Special relevance was given to specification of the spacer layers and their doping levels, in order to optimize the behaviour of the RTD in conjunction with the semiconductor optical waveguide, and to keep the operation away from catastrophic failure due to current breakdown.

The device fabrication procedure used to obtain a ridge waveguide diode incorporating the DBQW was described. A modified approach was employed to the electrical contact problem, avoiding the technique initially used (with ohmic contacts applied directly on AlGaAs layers). This was replaced by a configuration using mesas, i.e., contacts applied on a GaAs capping layer, which provide higher quality ohmic contacts - lower contact resistance, better stability and superior thermal behaviour). The description was kept as general as possible, covering the requirements of the two material systems employed (AlGaAs/GaAs and InGaAlAs/InP).

The steps taken to package the devices for high frequency optoelectronic testing were also described, including chip and wire bonding, and connectorization to allow high frequency operation. A description of the tests routinely applied to packaged and unpackaged devices comprised dc electrical measurements for control of fabrication, and low and high frequency measurements of packaged devices.

A description of optical and optoelectronic characterization techniques employed (modal profile, loss and transmisssion spectrum change, and light modulation) was also included in this chapter.

This chapter laid the foundations for discussing the design, fabrication and testing of RTD-EAMs. The specific details, procedures and results concerning GaAs devices (operating at around 900 nm) and InGaAlAs components (operating around 1550 nm) are the object of chapters 5 and 6, respectively.

Chapter 5

AlGaAs Based RTD-EAM

In this chapter, AlGaAs/GaAs RTD-EAM waveguide design, fabrication and characterization procedures are presented. The ridge optical waveguide configuration was employed due to its ability to laterally confine both electrical and optical fields, thus maximizing their interaction. Operation of the device as a high frequency intensity optical modulator is demonstrated.

5.1 AlGaAs RTD-EAM Design and Fabrication Procedures

In the wafer structures used throughout this work, the DBQW-RTD was located in the waveguide core centre. As discussed in chapters 3 and 4, there is a compromise between the required waveguide I-V characteristic (high peak current density, large PVCR, and low series resistance) and the optical waveguide properties (efficient light guiding, favourable light coupling, and low loss). These impose structural limitations on the DBQW (barriers and well thicknesses) and on the waveguide structure (core and claddings thicknesses and background doping profile).

5.1.1 Ridge Waveguide Design

The GaAs-$Al_xGa_{1-x}As$ double heterostructure, shown in Fig. 4.1(a), provides a useful buried guide, which can be qualitatively described as nearly symmetric, with abrupt refractive index steps [Fig. 4.1(b)]. (This versatile structure has been put to use on lasers, modulators, and detectors.) The refractive index of the $Al_xGa_{1-x}As$ alloy as a function of the wavelength, λ, is given by [15][191]:

$$n_{Al_xGa_{1-x}As}(\lambda) = \left[1 + \gamma\left(E_f^4 - E_g^4\right) + 2\gamma\left(E_f^2 - E_g^2\right)E^2 + 2\gamma E^4 \ln\left(\frac{E_f^2 - E^2}{E_g^2 - E^2}\right)\right]^{\frac{1}{2}}, \qquad (5.1)$$

where $E = 1.2398/\lambda$ eV (λ in μm), $\gamma = E_d/4E_0^3\left(E_0^2 - E_g^2\right)$, $E_f = \left(2E_0^2 - E_g^2\right)^{1/2}$, $E_d(x) = 36.1 - 2.45x$ eV, $E_0(x) = 3.65 + 0.871x + 0.179x^2$ eV, and $E_g(x) = 1.424 + 1.266x + 0.26x^2$ eV ($x < 0.4$).

The waveguide core of the AlGaAs/GaAs RTD-EAM was implemented with a layer of GaAs 1 μm thick to allow easy endfire light coupling. The AlGaAs/GaAs RTD-EAM cladding layers were implemented with $Al_{0.33}Ga_{0.67}As$ alloys; an aluminium concentration of 33% gives rise to an index difference, $n_{core} - n_{clad}$, of $\simeq$0.224 at 900 nm (the GaAs refractive index at 900 nm is $\simeq$3.593), which is sufficiently high to obtain efficient light confinement with relatively thin cladding layers, keeping the device series resistance low. Besides, $Al_{0.33}Ga_{0.67}As$ is a direct bandgap compound with conduction band valley discontinuities $\Delta E_{C_\Gamma} = 0.211$ eV, $\Delta E_{C_L} = 0.285$ eV, and $\Delta E_{C_X} = 0.312$ eV, that reduce the problem of conduction complexity associated with L and X valleys (avoiding the lower conductivity associated with indirect bandgap AlGaAs compositions) in higher aluminium content compounds.

The upper cladding layer thickness was made approximately equal to twice the reciprocal of the first mode exponential decaying factor, $\simeq$300 nm. Because the waveguide core and the substrate have similar real refractive indices, the lower cladding layer was made considerably thicker, acting as an isolation layer separating the GaAs guide core from the GaAs substrate. An isolation layer of around 600 nm was adopted in order to avoid radiation leakage into the substrate.

Transversal light and electric field confinement were obtained by defining a ridge waveguide structure, Fig. 4.1(c), etching away unwanted semiconductor material (upper cladding and core) by means of reactive ion etching (RIE), forming a ridge of $\simeq$1.4 μm, as represented in Fig. 5.1. Because air surrounds the ridge, the refractive index discontinuity in the transverse direction is $\Delta n \simeq 2.593$ (in the vertical direction it is $\simeq$0.224), and several lateral modes are allowed to propagate.

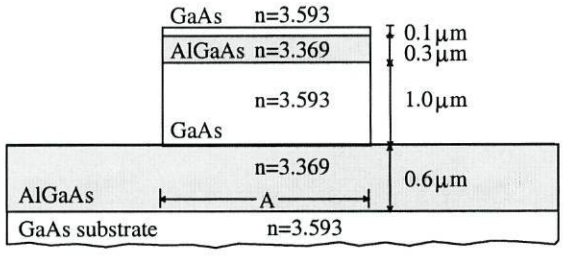


Figure 5.1: AlGaAs/GaAs ridge waveguide cross section schematic.

Table 5.1 shows the effective refractive index, n_{eff} (for TE and TM polarizations), and the profiles of the first three guided modes, for the case of a 4 μm wide ridge waveguide, obtained using *FWave IV* software.[1]

5.1.2 Wafer Specification

For a given material system, the parameters that significantly affect the RTD-EAM current density - voltage (J-V) characteristic are the barriers, quantum well and spacer layers widths,

[1] M. R. S. Taylor, "FWave IV: a vector E-M wave solver," user manual, University of Glasgow, Glasgow, United Kingdom (URL: http://www.elec.gla.ac.uk/~michaelt/software/fwave/).

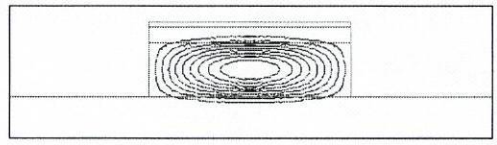
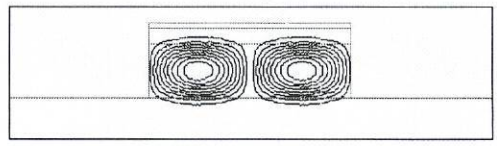
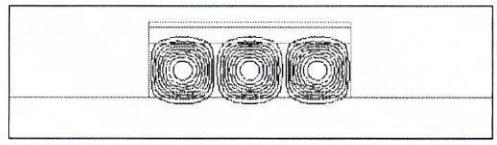
mode	$n_{\text{eff,TE}}$	$n_{\text{eff,TM}}$	*FWave IV* mode profile
1st	3.573	3.572	
2nd	3.567	3.567	
3rd	3.558	3.558	

Table 5.1: First three guided modes effective refractive index and profile, for the case of a 4 μm wide 1.4 μm ridge waveguide ($\lambda = 900$ nm).

contacts and spacer layers doping concentrations.

Because the DBQW is embedded in the centre of a 1 μm thick waveguide core, the spacer layer regions are much thicker than the DBQW ($\approx$500 nm versus $\approx$10 nm). Their background doping, $N_{d_{co}}$, has to be kept low to minimize the optical loss due to free-carrier absorption, and also because of the need to achieve complete collector spacer layer depletion under operation at the valley voltage. However, the emitter and collector spacer layers background doping concentration cannot be too low in order to ensure small series resistance compatible with wide bandwidth operation, and obtain low peak voltage operation [5]. The spacer layer on the emitter side of the DBQW region is expected to be practically undepleted, and in future designs it should be made thinner.

Because of the low fraction of the fundamental mode optical power guided in the claddings, a doping concentration in the claddings of $N_{d_{cl}} = 2 \times 10^{18}$ cm^{-3} was adopted that still does not give rise to large free-carrier absorption loss [58]. In the case of the core, a background doping concentration, $N_{d_{co}}$, smaller than the electron density associated with the peak current density, $\frac{J_p}{ev_{sat}}$ (v_{sat} represents the electron saturation velocity), has deleterious (high series resistance) and even catastrophic (current breakdown mainly through impact ionization) effects [63].

The value of the resonant tunnelling component of the peak current density, determined using Eq. 3.20, was used to estimate the collector spacer layer minimum background doping concentration. For a GaAs/AlAs RTD consisting of a DBQW heterostructure with 1.4 nm barriers and a 7 nm quantum well, and emitter doping concentration of the order of 4.7×10^{17} cm^{-3}, peak current density around 20 kA/cm^{-2} is expected, which requires an electron effective concentration around 2×10^{16} cm^{-3}, taking $v_{sat} \simeq 0.6\times10^{7}$ cm/s. A core background doping

concentration of 2×10^{16} cm^{-3} was adopted.[2]

Following the design rules discussed above, wafers #B634, #B757, #B710, and #B791, were grown by Molecular Beam Epitaxy in a Varian Gen II system, on a "semi-insulating" (SI) GaAs substrate, by the MBE Group of the Department of Electronics and Electrical Engineering of the University of Glasgow.

Wafers #B634 and #B757

Figure 5.2 shows schematically the #B634 and #B757 AlGaAs/GaAs RTD-EAM waveguide wafer structures, together with their conduction Γ-valley and refractive index profiles. The wafers #B634 and #B757 consisted of a DBQW, formed by two 1.4 nm thick AlAs barriers separated by a 7 nm wide GaAs quantum well, embedded between two 500 nm thick moderately doped (Si: 2×10^{16} cm^{-3}) GaAs spacer layers (refractive index ≃3.593), which are surrounded by heavily doped (Si: 2×10^{18} cm^{-3}) $Al_{0.33}Ga_{0.67}As$ cladding layers (refractive index ≃3.369). A 100 nm thick n+ GaAs cap layer was provided for formation of Au-Ge-Ni ohmic contacts.

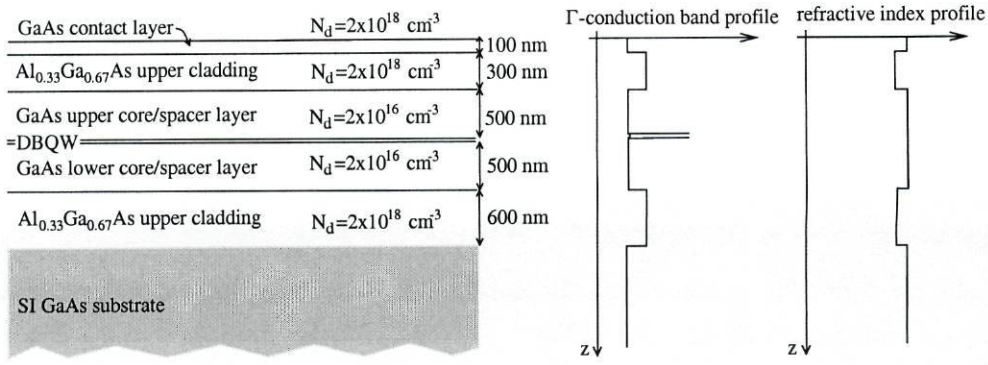


Figure 5.2: AlGaAs RTD-EAM wafer structure, Γ-valley, and refractive index profiles.

Wafer #B710

In order to improve the PVCR of the RTD, and hence the modulation depth and cut-off frequency, the wafer #B710 was grown with a DBQW consisting of a 1.7 nm thick barrier and a 5 nm well embedded in the same waveguide structure, while keeping the doping profile unchanged. However, the I-V characteristics of fabricated devices did not show stable NDC, breaking down at voltages slightly higher than the peak voltage. (Due to the thinner quantum well, the resonance width is larger, giving rise to a peak current density higher than 20 kA/cm^{-2}). This instability was not initially associated with breakdown originated by the lower background doping concentration, but only later on when a better understanding of the device behaviour was achieved through the models described in chapter 3.

[2]The AlGaAs/GaAs RTD-EAMs fabricated have shown peak current densities up to 13.5 kAcm^{-2}, which corresponds to an electron effective density of $\simeq 1.4 \times 10^{16}$ cm^{-3}, assuming $v_{sat} = 0.6 \times 10^7$ cm/s.

Wafer #B791

In order to have a more clear understanding of the effect of the DBQW on the diode waveguide I-V characteristic, the wafer #B791 was grown, and consisted of a $Al_{0.33}Ga_{0.67}As$-GaAs optical waveguide structure without DBQW; the doping profile used was the same of previous wafers.

5.1.3 Fabrication and Testing Procedures

In this subsection, processing of the GaAs RTD-EAM is described in detail, step by step, starting from scribing of the wafer up to the point where the device is packaged, according to the general procedures and technologies introduced in chapter 4.

The fabrication routine can be divided into 7 phases:

Phase 1: sample preparation

Wafer scribing into samples approximately 10 mm × 10 mm. Due to their greater fragility compared to silicon wafers, great care must be taken when handling GaAs substrates. The samples were cleaned in an ultrasonic bath, 5 min for each solvent: opticlear, acetone and methanol, followed by deionised water rinsing and nitrogen blow drying.

Phase 2: ohmic contacts formation and contact annealing

The samples were coated with AZ 1400-31 photoresist @ 6000 rpm for 30 s, then baked in an oven for 30 min @ 90 ^{0}C. Exposure of the samples used a mask aligner and the ohmic contact mask: the ohmic contact stripes were aligned parallel to the cleaved edge of the sample, in order to guarantee waveguide facets perpendicular to the cleavage planes. Exposure with the ultra-violet light source lasted for 10 s. Development for 75 s was followed by a postbake in the oven for 5 min @ 120 ^{0}C and HCl-H_20 (1:4) de-oxidisation etching for 30 s. The gold-germanium n-type ohmic contact recipe was evaporated using a Plassys machine, followed by lift-off in acetone. The ohmic contact patches were then examined under a microscope.

The metallization quality was verified through analysis of the corresponding Schottky I-V characteristic (the contacts had not yet been annealed), obtained using the HP parameter analyser with the help of needle probes. Before annealing, a rectifying behaviour is found identified by the characteristic (a) in Fig. 5.3, as a consequence of the relatively high GaAs-metal Schottky barrier (~0.8 eV for common metals) and of the doping concentration of 2×10^{18} cm^{-3}, not high enough to obtain a very thin barrier.[3] When annealed at a temperature above 360 ^{0}C, the metal-semiconductor contact becomes ohmic, as represented by the curve (b) of Fig. 5.3.

[3]The lower solubility for the n-type dopants used in MBE grown GaAs makes difficult to achieve a dopant concentration higher than typically 2×10^{18} cm^{-3} [56].

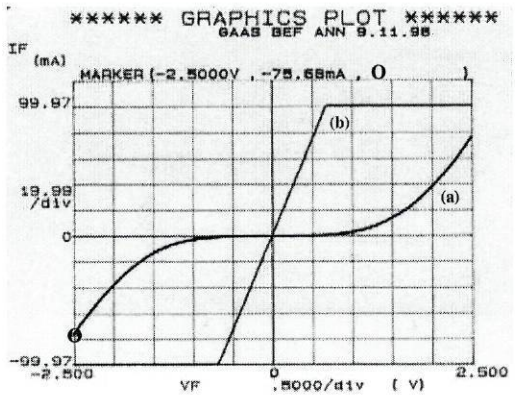


Figure 5.3: I-V curves of GaAs-AuGe metallization (a) ante- and (b) post-annealing.

Phase 3: waveguides definition, and wet etching of the $n+$ GaAs cap layer

The sample was coated with 1400-31 photoresist, spun at 4000 rpm for 30 s, followed by baking for 30 mins @ 90 ^{0}C. The waveguide mask consisted of sets of stripes 2, 3, 4, 5, and 6 μm wide, and 8 mm long. The sample was aligned with the waveguides mask and exposed (mask aligner) for 7 s, followed by development for 120 s and post-baking for 15 min @ 120 ^{0}C. The resist pattern obtained, consisting of stripes of photoresist, acted as dry etching mask.

The unprotected epilayer was etched using the RIE 80 machine with $SiCl_4$ gas and 100 W rf power, to a depth of $\approx$1.4 μm, as shown in Fig. 5.4. (The etching depth was verified using a step profiler.) The average etching rate was 0.12 nm/min, but varied slightly from run to run. A quarter wafer of GaAs was placed under the sample to ensure reproducible and good quality etching.

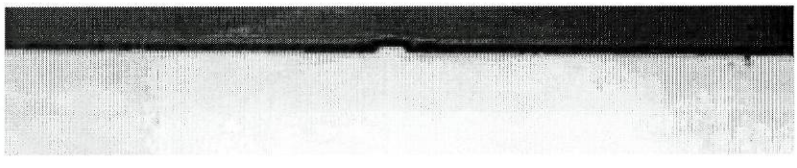


Figure 5.4: Side view of the AlGaAs waveguide (ridge depth: 1.4 μm; width: 4μm).

The sample was then soaked in acetone to remove the resist dry etching mask. The top $n+$ GaAs ridge capping layer region uncovered by the ohmic metallization was etched to reduce current spreading along the ridge axis. An alkaline solution of NH_4OH in combination with H_2O_2 and H_2O (1:1:200) was used, with an etching time of 60 s. The expected negative differential conductance characteristic of the RTD-EAM, Fig. 5.5, was then observed using the HP parameter analyser.

Phase 4. SiO_2 deposition (passivation) and access contact windows formation

The sample was covered by a layer of dielectric SiO_2 around 500 nm thick, deposited by plasma assisted chemical vapour deposition (PECVD), and then coated with 1400-31 resist, spun @ 4000

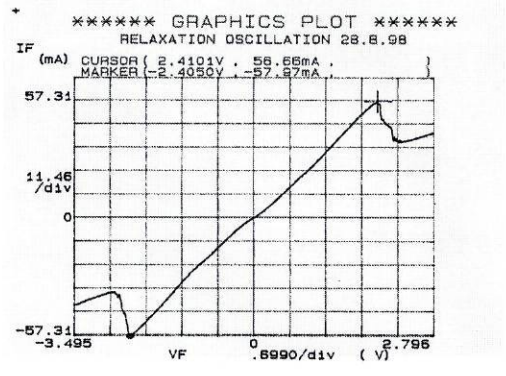


Figure 5.5: I-V chatacteristic of a AlGaAs RTD-EAM showing regions of NDC.

rpm for 30 s, followed by baking over 30 mins @ 90 ^{0}C. The access contact window mask and sample were aligned in order to expose defined areas of the resist above the mesas and ridge ohmic contact patches; the exposure time was 10 s. The resist was developed for 75 s, and post baking lasted for 30 min at 120 ^{0}C. The uncovered SiO_2 was etched with C_2F_6 gas using the BP80 machine, with an etching rate about 50 nm/min (for a rf power of 100 W and a gas flow rate of 20 sccm).

Phase 5: contact pad (CPW) formation

The sample was next coated with 1400-37 photoresist, spun @ 4000 rpm for 30 s (giving a film $\simeq$3.2 μm thick), followed by baking for 30 mins @ 90 ^{0}C. Alignment of the sample to the contact pad mask and exposure for 22 s were followed by development for 75 s.

Evaporation of a three-layer metallization of Ti-Pd-Au was performed. A final layer of gold 300 nm thick was then applied using an evaporator allowing angled deposition, in order to ensure continuous step coverage, providing good connection between the top and the bottom of the waveguide. The sample was first tilted by about 45^0 to one side of the waveguide, and then tilted by 45^0 to the other side. Lift-off in acetone followed. Device I-V characterization was then realized.

Phase 6: sample thinning and device separation

The purpose of sample thinning was to facilitate device separation (dicing) and reduce thermal resistance. The GaAs substrate was hand lapped on a grinding plate until its thickness was $\simeq$150 μm, as described in chapter 4. The devices were separated as 500 μm long by 500 μm wide dies, with perfect crystal facets.

Phase 7: device packaging

The optical quality of the die facets perpendicular to the waveguide was then inspected using the EDB 80 eutectic die bonder. The devices showing good facets and having appropriate dc

current-voltage characteristic were mounted on the test package described in chapter 4. Correct waveguide orientation is needed to facilitate light coupling. The die was bonded by attaching it, substrate down, to the package pedestal, using silver epoxy glue, which was then cured at 100 ^{0}C for 5 minutes to provide good adhesion and low thermal resistance. The chip CPW microscopic patches were gold wire connected to the corresponding package CPW pads terminals, as schematically represented in Fig. 5.6.

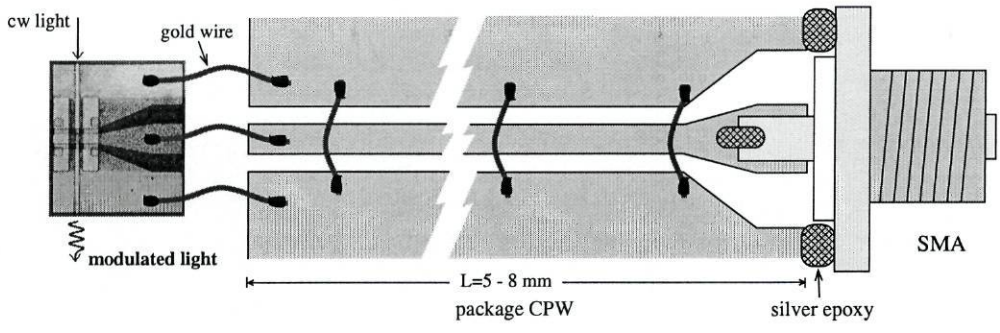


Figure 5.6: Diagram of the RTD-EAM and SMA connection to the CPW package.

5.2 AlGaAs RTD-EAM Electrical Characterization

The Hewlett-Packard model 4145A semiconductor parameter analyser was employed to graphically display the dc characteristics of the RTD-EAM devices throughout the fabrication steps. High bandwidth oscilloscopes, together with a HP 8564E rf spectrum analyser (9 kHz - 40 GHz bandwidth), were used to characterize the relaxation oscillations.

5.2.1 dc I-V characteristics

RTD-EAM optoelectronic properties, such as modulation depth and electrical bandwidth, can be extrapolated from the negative differential conductance region parameters; the dc current-voltage characteristic also gives a measure of the quality of the wafer growth and of the contacts.

Wafers #B757 and #B634

These wafers were grown with 1.4 nm thick barriers and a 7 nm quantum well. The dc I-V characteristics of the AlGaAs-GaAs RTD-EAMs fabricated according to the procedure described above display typical RTD behaviour, shown in Figs. 5.5 and 5.7, with peak current densities up to 13.5 kAcm^{-2}, and peak-to-valley current ratios in the range of 1.4-1.8; $\Delta J = J_p\,(1 - PVCR^{-1})$ is around 5 kAcm^{-2}. The peak voltages are in the range 1.5 - 3.0 V, and the valley voltages vary from 1.8 to 3.5 V. Because the GaAs based RTD-EAM structure is practically symmetric in relation to the DBQW and the device is unipolar, the current-voltage characteristic is anti-symmetric.

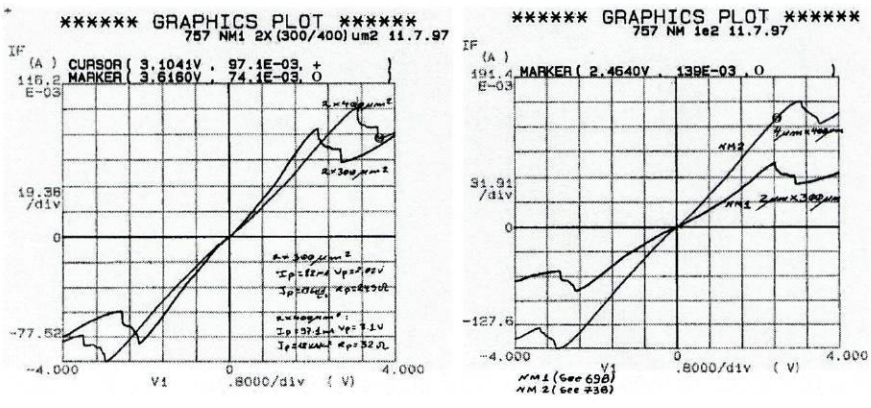


Figure 5.7: Typical I-V curves of AlGaAs RTD-EAMs.

The smaller area devices show higher peak current density, larger PVCR and lower peak voltage. As the active volume of the devices increases the electric field is less effective on bending the energy bands due to non-uniformity of the ohmic contacts, and the number of scattering events rises, which leads to lower PVCR and higher peak voltage as consequence of the higher leakage current. Devices implemented using material away from the centre of the wafer showed poor characteristics, possibly due to lack of wafer uniformity in the periphery. From a RTD application point of view, the issue of wafer uniformity is of paramount importance.

Wafer #B710

The DBQW in wafer #B710, consisting of 1.7 nm (6 ML) barriers and a 5 nm (18 ML) quantum well, was designed to increase the PVCR, and hence improve RTD-EAM high frequency operation and modulation depth. The devices fabricated presented peak-to-valley current ratios up to 3.3. However, as pointed out before, devices showed catastrophic current breakdown at voltages slightly higher than the peak voltage, with peak current densities up to 30 kAcm^{-2}, Fig. 5.8. This measured peak current is equivalent, assuming v_{sat} =0.6×10^7 cm/s, to an electron effective concentration higher than 2×10^{16} cm^{-3} (which was the requested wafer background doping). At the time this behaviour was not well understood, being originally associated with growth errors.

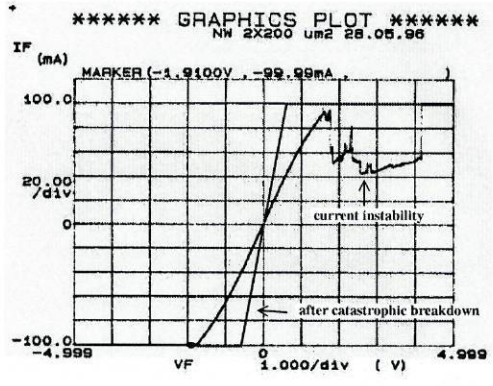


Figure 5.8: Example of catastrophic current breakdown in a device from wafer #B710.

These peak current densities, with resonant voltages as high as 4 V, correspond to large power dissipation, which can result in thermal instability for a large-area diode, leading to burn out in a few seconds after biasing around the peak voltage. Also because the requested wafer background doping (waveguide core background doping concentration) is smaller than the electron concentration required to support the peak current density, the magnitude of the electric field in the collector region reaches a sufficiently high value to cause breakdown.

Wafer #B791

In order to have a clear understanding of the influence of the DBQW on the current-voltage characteristic and on the waveguide optical transmission, this wafer was grown with a structure not including the DBQW. The I-V characteristic of the corresponding diode optical waveguide, with the same waveguide geometry as the RTD-EAM, is presented in Fig. 5.9.

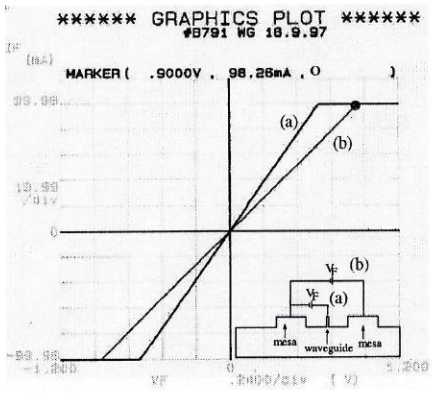


Figure 5.9: Typical I-V curve of AlGaAs optical waveguide diodes without the DBQW.

The diode waveguide I-V characteristic shows a clear absence of NDC, being linear in the voltage range employed. From the I-V curves it does not seem possible to obtain sufficiently high electric field and any significant absorption change when biasing the waveguides (which was confirmed by studying the waveguide optical transmission as a function of the applied bias).

Dc Bias Oscillations

Under certain circumstances, the RTD in conjunction with the extended wiring found in probe stations[4] functions as an active resonant circuit supporting low frequency oscillations. Figure 5.10(a)-(c) shows the test equivalent circuit schematic and the ac equivalent circuit, including the RTD and the bias circuit, together with the I-V load diagram.

If the source V_s and the resistor R are such that the load line crosses the NDC region once, Fig. 5.10(c), the RTD is dc bias stable in the NDC region: when the imaginary part of the total impedance of the circuit is zero and the real part is zero or negative, noise present in the circuit reinforces a given frequency (the resonant tank circuit fundamental frequency, or the

[4]Which corresponds mainly to an inductance L.

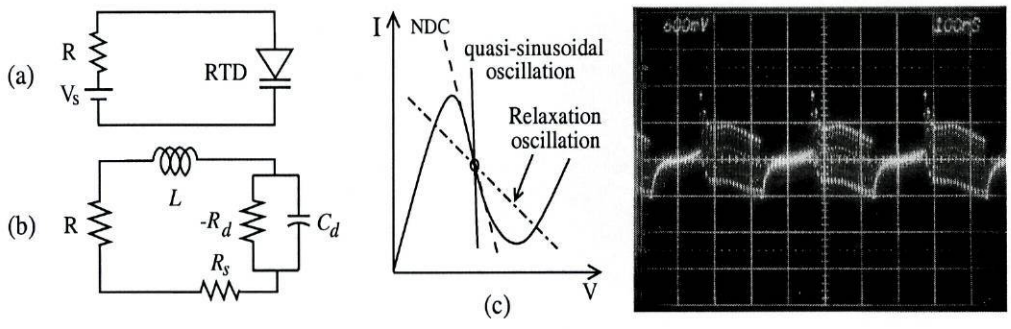


Figure 5.10: (a) Schematic of the dc test circuit. (b) ac equivalent circuit. (c) I-V load diagram. (d) Experimental result of bias circuit oscillation (around 4 MHz).

most favourable harmonic), which builds up until the magnitude of the oscillation exceeds the extremes of the NDC region. The energy is supplied from the dc source and converted into ac oscillation.

The steep slope of the NDC region found in RTDs makes it very difficult to achieve dc bias stability in this region, leading to bias circuit oscillation at low frequencies. Only small-area RTDs, having a suitably low differential conductance, are easily stabilized. Figure 5.10(d) presents experimental results of low frequency (around 4 MHz) RTD-EAM self-oscillation associated with the instability of the dc power supply. This oscillation can be understood considering that typical power supplies have control circuits which are designed to mantain a constant voltage; when feeding a NDC load, the power supply stability fails and it bursts into oscillation. This can be minimized by employing ferrite bead inductors as high frequency oscillation suppressors.

5.2.2 Relaxation Oscillation

As discussed in chapter 3 (section 3.3.4), RTDs are well known for their high frequency generation and high speed switching characteristics (associated with the NDC region). The essential conditions for quasi-sinusoidal oscillations in RTDs are [143]: (*i*) monostable dc bias point, (*ii*) unstable ac load and (*iii*) resonant ac load circuit, with a loaded quality factor (including the RTD resistance) that exceeds unity at some frequency less than the f_{max} of the RTD. The first two conditions are given simply by $R_L < \Delta V/\Delta I$, where ΔV and ΔI are the width and height, respectively, of the NDC region [143]. Under these conditions, an ac fluctuation of the RTD voltage in the vicinity of the resonant frequency will grow exponentially until the amplitude roughly approaches ΔV. At this point, saturation sets in and a steady-state quasi-sinusoidal waveform occurs across the load resistor.

As discussed before, the typical steep slope in the NDC region makes it difficult to achieve dc bias stability. Connection of the RTD-EAM to a high frequency energy-storage element, such as a length of transmission line, as indicated in Fig. 5.11, results in a RTD-EAM transmission line relaxation oscillator [157], which does not require dc bias stability in the NDC region.

The relaxation oscillation corresponds to repetitive switching of the RTD-EAM through the NDC region, between dc stable points lying below the peak point and above the valley point,

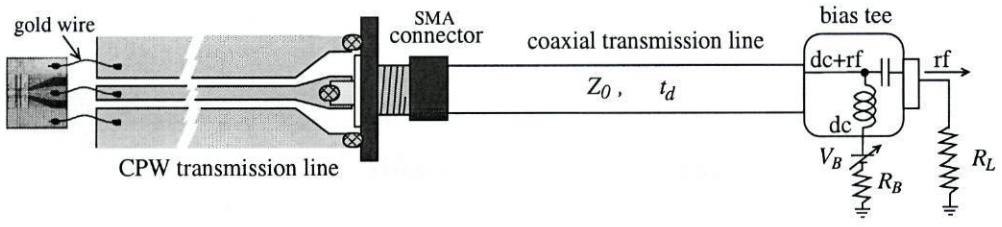


Figure 5.11: RTD-EAM transmission line relaxation oscillator schematic.

where the period of the output voltage waveform - the voltage across the RTD-EAM terminals - is determined by the electrical delay of the transmission line [157]. The left side of the modulator CPW is unterminated (open circuit), as indicated in Fig. 5.11. The RTD-EAM is biased, through the bias resistor R_B, by the supply voltage V_B. The bias tee, with a 26 GHz bandwidth, was employed to couple the RTD-EAM to the bias voltage used to induce and mantain the oscillations, through a transmission line with a characteristic impedance Z_0 and an electrical delay t_d. The RTD-EAM ac load R_L was the 50 Ω input impedance of an oscilloscope. The quiescent operation point was initially just below the peak point, and then the RTD-EAM was induced to switch towards the valley point by slowly increasing the dc voltage V_B.

Switching generates a waveform a which travels down the line, Fig. 5.12. The existence of an impedance discontinuity in the circuit, giving a negative current reflection coefficient, r, generates a backward wave b with inverted polarity, $r = b/a$, Fig. 5.12, which arrives back at the RTD-EAM after a delay $2t_d$.

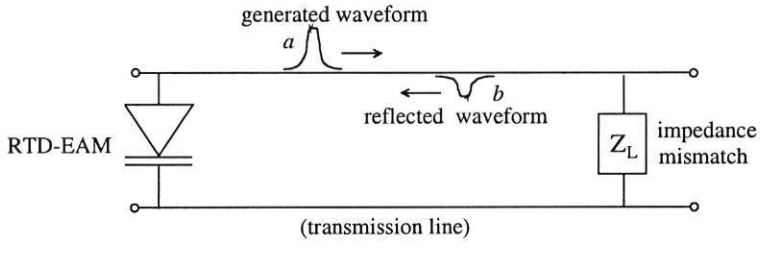


Figure 5.12: RTD-EAM transmission line relaxation oscillator working principle.

If $2t_d$ is sufficiently larger than the RTD-EAM switching time, t_R, the returning waveform will then induce the RTD-EAM to switch back to the initial operating point. The subsequent returning waveform then induces a similar switching action to the first one, and the process repeats itself at a rate of approximately $(4t_d)^{-1}$. The RTD-EAM dwells in either of its positive differential conductance regions between switching events.

Figure 5.13 presents the typical voltage waveforms of RTD-EAM transmission line relaxation oscillators implemented with coaxial cables of different lengths (1 m; 50, 15 and 10 cm). The devices were initially biased slightly below the peak point, and then the voltage was increased until oscillation starts.

As the length of the cable decreases (repetition rate increases), the oscillating waveform loses its square wave character, acquiring a quasi-sinusoidal shape; the voltage swing decreases,

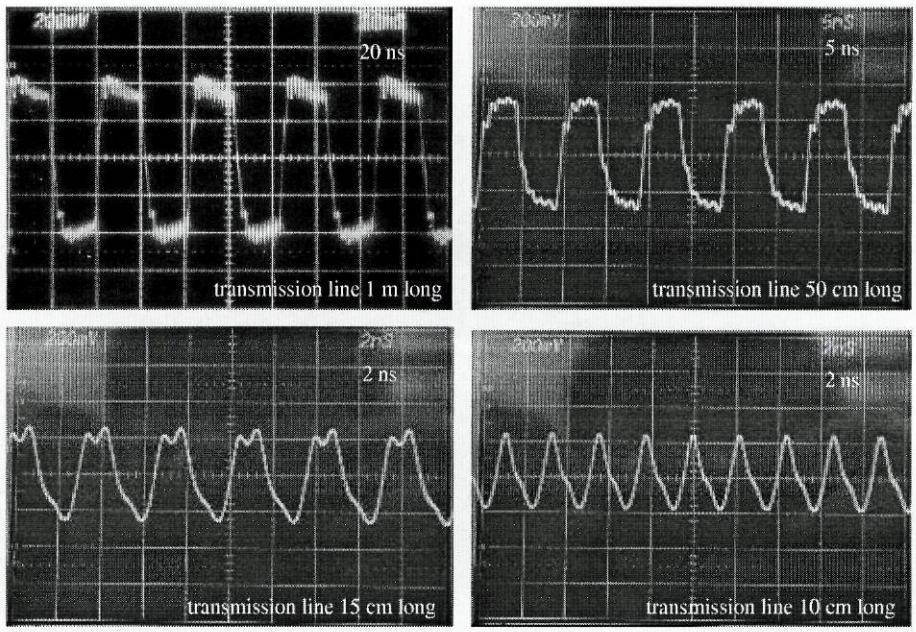


Figure 5.13: Experimental voltage waveforms of a RTD-EAM connected, as indicated in Fig. 5.11, to coaxial cables of different lengths: 1 m; 50, 15 and 10 cm, respectively.

approaching the value of the NDC region voltage width, ΔV.

The repetition rate can be slightly changed by varying the optical power coupled into the RTD-EAM; the corresponding frequency decreases when the optical power rises. For an oscillation frequency around 470 MHz, a tuning range of 10 MHz was observed. Figure 5.14 shows the rf spectrum analyser display of the relaxation oscillation signal with and without optical radiation being guided. (The laser output power was kept low in order to avoid waveguide facet damage.)

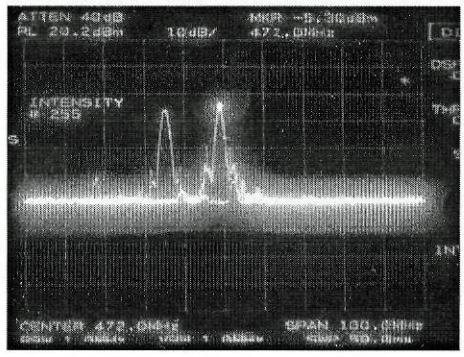


Figure 5.14: Rf spectrum of the relaxation oscillation signal with and without optical radiation being guided. (RTD-EAM connected to a coaxial line 10 cm long, as in the case of Fig. 5.13.)

Figure 5.15 shows the frequency spectrum of the RTD-EAM relaxation oscillations generated when a 15 cm long coaxial cable was employed, at two slightly diferent dc bias points in the NDC region.

Frequency components up to 16 GHz are clearly observed, probably limited by the package

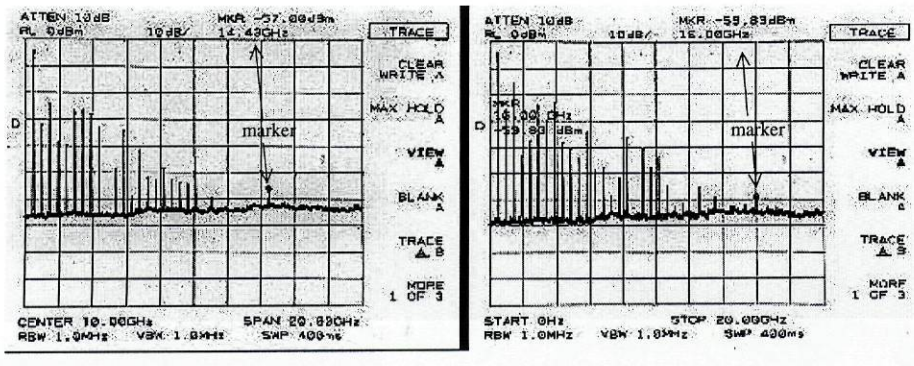


Figure 5.15: Frequency spectrum of the relaxation oscillation waveform, represented in Fig. 5.13, for the case of the RTD-EAM connected to the coaxial cable 15 cm long.

bandwidth. Changing the dc bias point (in the NDC region) modifies the negative dynamic conductance value, altering the gain bandwidth.

5.2.3 High Frequency Response

The switching time of the RTD-EAM, and hence the cut off frequency, $f_{max} = 1/4\pi t_R$, can be estimated from the shape of the NDC region. The RTD-EAM modulator 10%-90% switching time, t_R, between the peak and valley points can be estimated from $t_R = 4R_dC_d \simeq 4\,(\Delta V/\Delta J)\,C_v$ [152][157], where C_v is the capacitance at the valley point per unit area ($C_v = \epsilon/W$, where ϵ is the electrical permitivity, and W is the depletion region width). For 4 μm $\times$ 200 μm active area devices, with $\Delta V \simeq$0.4 V, $\Delta J \simeq$5 kAcm^{-2}, $W = 0.5$ μm and $\epsilon = 13\epsilon_0$, $t_R \simeq$7.4 ps. From this switching time, a cut-off frequency $f_{max} \approx$10.8 GHz is to be expected.

5.3 AlGaAs RTD-EAM Optical Characterization

Optical characterization of GaAs based RTD-EAMs included analysis of modal distribution, determination of propagation loss, followed by measurement of the electric field induced bandedge broadening, using light from a cw Ti:sapphire laser, tuneable in the wavelength region around the absorption edge of the GaAs waveguide (850-950 nm) [185].

5.3.1 Optical Modes

The laser light was coupled in and out using a microscope objective end-fire arrangement, Figs. 4.21 and 5.16. The coupling efficiency into the waveguide fundamental mode, which can be estimated from the overlap integral between the laser beam distribution and the mode distribution (Eq. 2.118), must be as high as possible. Microscope objective lenses of $\times$20 magnification were used because they make the device more resistant to excessive incident intensity, which can be one of the most serious problems shorting the lifetime of the device (the waveguide facets are easily destroyed by the high power density launched into them). Important factors are the initial

presence of any flaws in the facets, and the width of the waveguide; damage initiated at flaws will spread to other regions of the facet. The safe operating level can be increased by use of an anti-reflection dielectric coating on the waveguide facets, which also reduces the effects of erosion, or gradual degradation of the facets [58]. (This is a moisture induced phenomenon which reduces the mirror reflectivity.)

The set-up employed, Fig. 5.16 (Fig. 4.21), did not allow a clear distinction of the various guided optical modes supported by the GaAs-AlGaAs waveguide. However, the CCD camera line scan facility made possible to observe the intensity distribution associated with the first two propagated modes. The widest ridge waveguide implemented, 6 μm wide, shows clearly three guided modes. The *FWave* analysis (section 5.1.1) shows a larger number of modes, but because the claddings and the core are lossy (due to free carrier absorption), and the background material absorption, waveguide scattering loss and the loss due to the metallization across the ridge walls, the number of modes effectively propagated is reduced to the first two or three , which have more confined modal distributions. Waveguides with 4 μm $\times$ 500 μm dimensions show an absorption bandedge around 890 nm.

5.3.2 Propagation Loss

The GaAs-AlGaAs waveguide propagation loss was determined employing the method of sequential cleaving [58][177]. The experimental set up used is represented in Fig. 5.16. Low optical power density was employed in order to avoid/neglect any nonlinearities and the risk of facet damage.

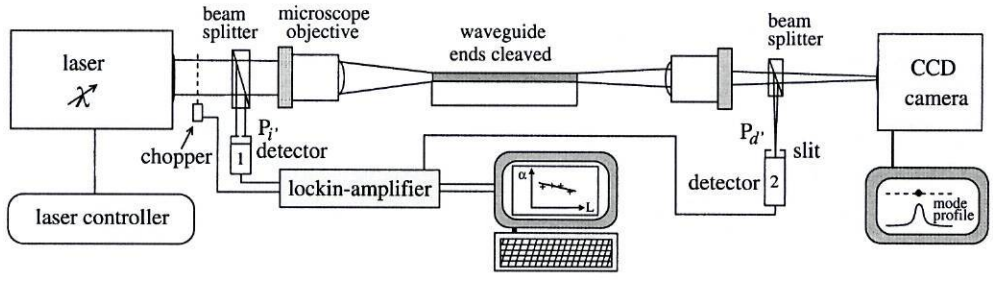


Figure 5.16: Propagation loss measurement set up.

In the method of sequential cleaving used, light of the desired wavelength was focused directly onto a cleaved input waveguide facet, and known fractions of the input and of the transmitted output powers were measured [58], as shown in Fig. 5.16. Care must be taken before each measurement to efficiently align the laser beam and the waveguide for optimum coupling, maximizing the observed output power.

The powers $P_{i'}$ and $P_{d'}$ (the input and output power fractions, respectively) measured by two detectors, as represented in Fig. 5.16, are proportional to the coupled and the decoupled optical

powers (see Fig. 4.22 and Eq. 4.7), P_0 and P_L, respectively, with [58]

$$\ln\left(\frac{P_{i'}}{P_{d'}}\right) \propto \alpha L, \tag{5.2}$$

where L is the length of the waveguide, and α is the attenuation coefficient. Measurement for a given polarization (TE or TM) were repeated for different lengths of the waveguide (L_j), beginning with a relatively long waveguide sample, and then repetitively shortening the sample by cleaving [58]. After the measurements were made for two different waveguide widths (3 and 5 μm) and for several waveguide lengths (L_j: 1.5, 2, 4, and 5.5 mm), the resulting data was used to plot $\ln\left(\frac{P_{i'}}{P_{d'}}\right)$ against L_j, Fig. 5.17.

The attenuation coefficients for TE and TM polarizations, α_{TE} and α_{TM}, obtained for two ridge waveguide widths, are given by the linear fit represented in Fig. 5.17, showing $\alpha_{\mathrm{TE}} \simeq 5.5$ cm^{-1} and $\alpha_{\mathrm{TM}} \simeq 6.7$ cm^{-1} for a ridge 3 μm wide, and $\alpha_{\mathrm{TE}} \simeq 3.9$ cm^{-1} and $\alpha_{\mathrm{TM}} \simeq 4.3$ cm^{-1} for a ridge 5 μm wide. The data points fall closely into the straight line, showing that sufficient consistency has been achieved.

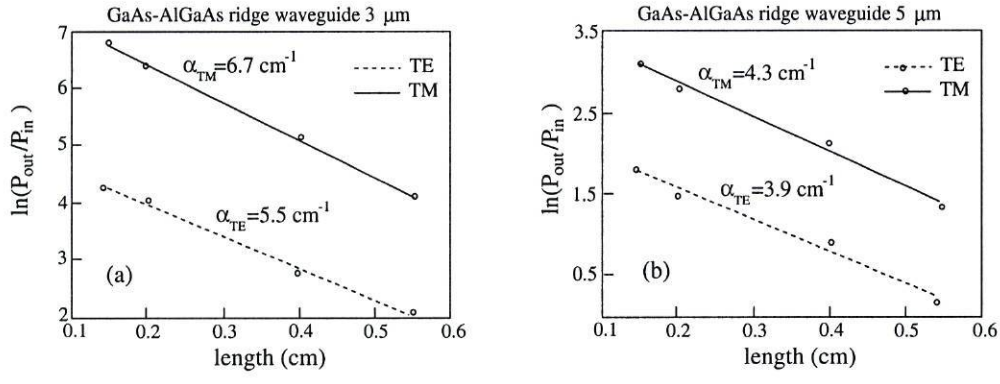


Figure 5.17: Sequential cleaving experimental data at around 908 nm for a 3 μm wide waveguide, and at around 915 nm for 5 μm waveguide, respectively.

TM-polarized light sustains higher loss owing to heavy doping of the cladding layers, due to the larger overlap of the field with the absorptive cladding materials. The free carrier absorption (intraband absorption), responsible for losses in the waveguide conducting (highly doped) cladding/metal contact layers, estimated using Eq. 2.75 [58], is around 2.7 cm^{-1}, for a cladding doping concentration of $N_{\mathrm{C}} \simeq 2 \times 10^{24}$ m^{-3}, assuming $\mu_{\mathrm{ac}} \simeq 0.2$ m^{2}V^{-1}s^{-1}, $n_{\mathrm{r}} = 3.6$, $\lambda = 908 \times 10^{-9}$ m.

5.3.3 RTD-EAM Transmission Spectrum

Application of voltage across the RTD-EAM terminals induces, via the Franz-Keldysh effect, broadening of the waveguide absorption edge. Knowledge of the waveguide transmission spectrum at several bias points allows the determination of the operating wavelength that gives the maximum modulation depth.

Absorption Bandedge "Shift"

To measure the absorption bandedge broadening due to the Franz-Keldysh effect, a silicon photodetector was employed. The RTD-EAM waveguide transmission spectra at zero bias, slightly below the peak, and just above the valley points, are shown in Fig. 5.18. (The device was not dc biased in the NDC region in order to avoid self-oscillation.)

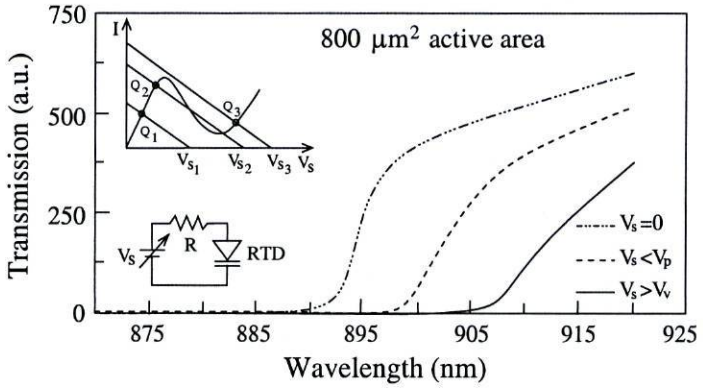


Figure 5.18: GaAs RTD-EAM transmission spectrum at several dc bias voltages.

It is clear from Fig. 5.18 that as the applied voltage increases from the peak to the valley point, there is a sharp drop in the waveguide transmission at wavelengths in the range 900 – 910 nm. The results obtained can be explained with the help of the inset diagram of Fig. 5.18. Considering an initial condition with $V_s = V_{s_1}$, Q_1 becomes the operating point, giving an electric field $\mathcal{E}_1$. As V_s increases to V_{s_2}, the operating point moves to Q_2 (slightly below the peak point), and the electric field increases to $\mathcal{E}_2$, due to the longer collector depletion region, with a corresponding decrease in the waveguide transmission (following the Franz-Keldysh effect). When V_s increases to V_{s_3}, the operating point shifts to Q_3 (on the second PDC region, above the valley point), with a large rise in the voltage across the RTD-EAM and a sharp drop in the current flow. As a consequence, there is a steep increase in the magnitude of the electric field to $\mathcal{E}_3$. In the RTD-EAM, when biased close to the peak point, a relatively small increment in V_s leads to a large increase in the electric field ($\mathcal{E}_3 \gg \mathcal{E}_2$), and so to a large drop in the waveguide transmission for wavelengths slightly higher than the waveguide core bandedge wavelength.

From Eq. 3.61, $\Delta\mathcal{E}_{p-v} = \frac{\Delta V_d}{W} + \frac{W}{2\epsilon v_{\text{sat}}}\Delta J_{p-v}$, and using the GaAs RTD-EAM I-V characteristics presented in section 5.2.1, the estimated electric field change induced by peak-to-valley switching, $\Delta\mathcal{E}_{p-v}$, is of the order of 19 kV/cm, taking $\epsilon = 13\epsilon_0$, W=500 nm and $v_{\text{sat}} = 1.0\times10^7$ cm/s. The absorption bandedge "shift" due to electric field enhancement in the collector depletion region, as a result of the Franz-Keldysh effect, is given approximately by (Eq. 3.63)

$$\Delta\lambda_g \approx \frac{\lambda_{g'}^2}{hc}\left(\frac{e^2h^2}{8\pi^2 m_r}\right)^{\frac{1}{3}}\Delta\mathcal{E}_{p-v}^{\frac{2}{3}},$$

where $\lambda_{g'}$ is the wavelength corresponding to the waveguide transmission edge when biased

slightly below the peak voltage, which is around 895 nm; $m_{r_{lh}} \simeq 0.0369m_0$, and $m_{r_{hh}} \simeq 0.0583m_0$. The observed bandedge "shift" associated with peak-to-valley switching was approximately 12 nm; considering the approximation used, the value measured is close to the value of ≈ 9 nm given by the expression above, confirming that $\Delta\mathcal{E}$ is around 19 kV/cm.

Modulation Depth

The RTD-EAM modulation depth, at low frequency, due to the transition between the two positive differential conductance (PDC) regions was determined by dc biasing the RTD-EAM slightly above the valley point (to minimize thermal effects), and injecting a low frequency (100 MHz) negative impulse square wave *rf* signal, with amplitude slightly higher than $\Delta V = V_v - V_p$. The Hamamatsu C5680 streak camera was employed as photo-detector, as illustrated in Fig. 5.19. A fraction of the injected *rf* signal was required to trigger the streak camera.

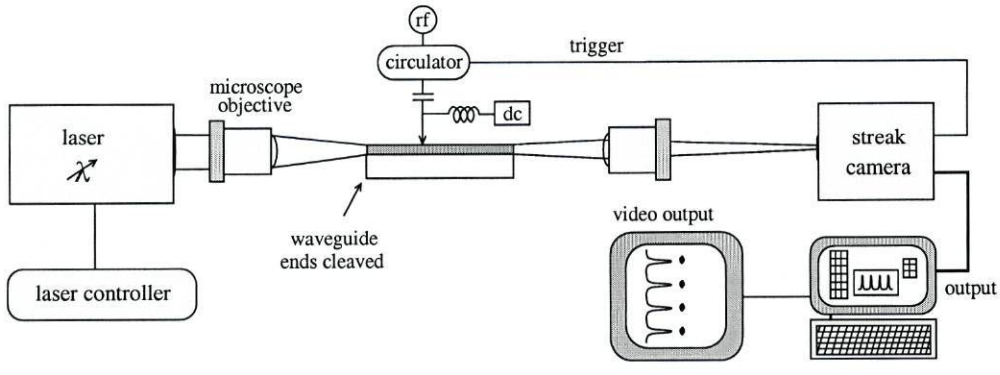


Figure 5.19: Scheme of the experimental set up for electro-absorption modulation characterization of GaAs based RTD-EAMs.

Figure 5.20 presents the optical modulation depth as a function of the operating wavelength, for 4 μm $\times$ 200 μm active area devices, due to the transition between the two positive differential conductance (PDC) regions, showing a maximum modulation depth of 13 dB at wavelengths around 908 nm.

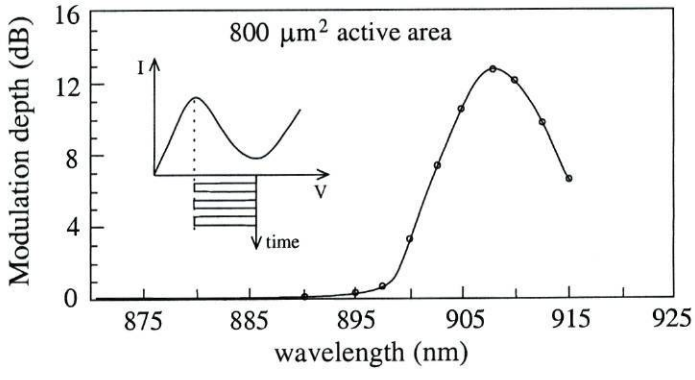


Figure 5.20: RTD-EAM modulation depth as a function of the wavelength.

From the knowledge of the magnitude of the electric field change, $\Delta\mathcal{E}\simeq$19 kV/cm, the mod-

ulation depth can be estimated, using Eqs. 3.64 and 2.113. The attenuation coefficient at the peak voltage is $\alpha(\hbar\omega, \mathcal{E}_p) \simeq 30$ cm^{-1} (estimated from the transmission spectrum presented above, Fig. 5.18, taking $\alpha(\hbar\omega, 0) \simeq 5.5$ cm^{-1}). Using the electric field change determined, 19 kV/cm, with $\hbar\omega \simeq 1.365$ eV (λ =908 nm) and $\hbar\omega_g' = hc/\lambda_g \simeq 1.388$ eV (893 nm), the aborption change $\Delta\alpha(\hbar\omega, \mathcal{E}_v)$ becomes $\simeq 200$ cm^{-1} (Eq. 3.65). The expected on/off modulation ratio is then, Eq. 2.113,

$$R_{\text{on/off}}(\text{dB}) = 4.343\gamma\Delta\alpha(\hbar\omega, \mathcal{E}_v)\, L, \tag{5.3}$$

with L=200 μm, and taking $\gamma \simeq 0.5$, gives a modulation depth around 9 dB. This result agrees relatively well with the measured 13 dB modulation depth, considering the approximations made.

Longer devices (1200 μm^2) show modulation depths up to 18 dB. Figure 5.21 presents the dependence of modulation depth results for RTD-EAMs on the device active area; waveguides were 4 μm wide and the devices were dc biased around the valley point, with a 100 MHz signal of 0.7 V amplitude being injected.

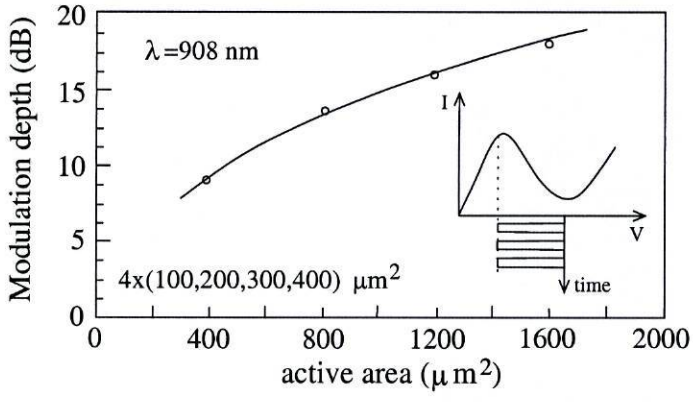


Figure 5.21: Modulation depth as function of the RTD-EAM active area (ridge width 4 μm).

5.4 AlGaAs RTD-EAM Modulation Characterization

The high frequency optical response of the modulator was characterized using the Hamamatsu C5680 streak camera, which offers a minimum time resolution of around 2 ps. A fraction of the generated/injected *rf* signal power was required to trigger the streak camera. Figure 5.19 schematically presents the experimental set-up. The dc and rf signals used to bias the devices were applied through a bias tee, with a circulator being employed to isolate the *rf* input signal (when provided) from the output *rf* signal of the RTD-EAM (which is used to drive the triggering circuit of the streak camera).

5.4.1 Relaxation Oscillation Optical Modulation

As discussed previously, section 5.2.2, the RTD-EAM is able to operate in the relaxation oscillation mode. Optical modulation associated with the relaxation oscillation operation is then to be

expected. To observe this behaviour, the implementation represented in Fig. 5.11 was employed, and no rf drive was directly applied. The RTD-EAM generated *rf* signal was used to feed the streak camera trigger circuit connected to the bias tee *rf* port, as indicated in Figs. 5.11 and 5.19.

The optical responses of RTD-EAMs with 4 μm$\times$ 200 μm active area, connected to transmission lines 1 m, 50, 15 and 10 cm long, operating in the relaxation oscillator mode, are presented in Fig. 5.22, showing modulation depth as high as 12 dB at $\lambda = 908$ nm.

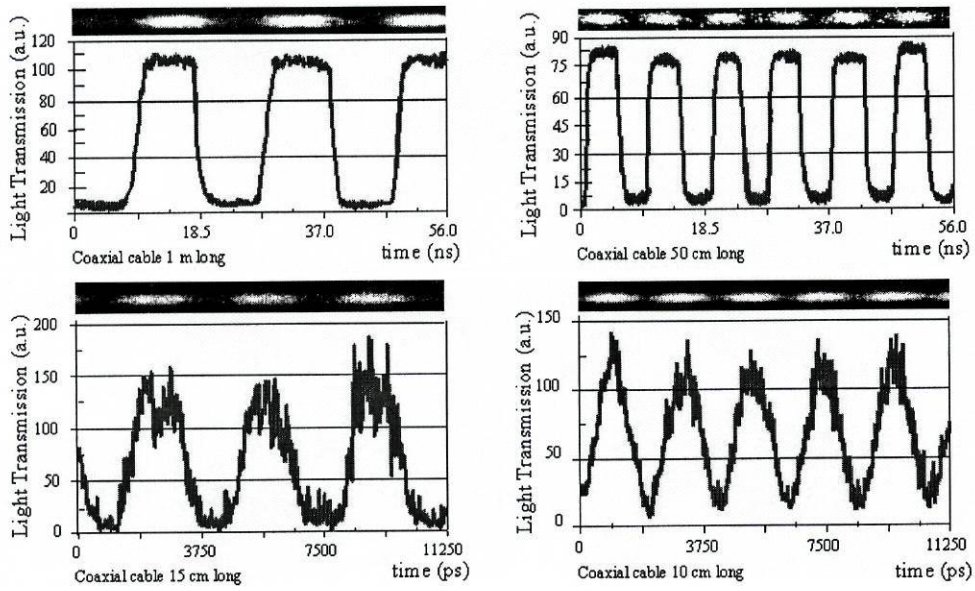


Figure 5.22: Optical responses of RTD-EAM transmission line relaxation oscillators, for lines 1 m, 50, 15 and 10 cm long, respectively.

This optical response is in good correlation with the electrical output displayed in Fig. 5.13. These results suggest that if a suitable cavity with a resonant frequency within the RTD-EAM electrical bandwidth is provided, a "relaxation oscillation modulator" can be implemented.

5.4.2 Direct Modulation

To study direct modulation, the RTD-EAMs were dc biased slightly above the valley point, and the modulating signal was injected through the *rf* connection of the bias tee, with a fraction of the signal used to trigger the streak camera (Fig. 5.19).

Direct modulation results, when 500 and 950 MHz voltage sine-waves were applied, are shown in Fig. 5.23. The drive signal amplitude was kept slightly larger than $\Delta V = V_v - V_p$ (0.6 V). The optical modulation is as high as 11 dB.

Direct modulation at around 16 GHz was characterized employing a InGaAs photodetector[5] connected to the *rf* input of the spectrum analyser. Figure 5.24 shows the modulator optical response to the 16 GHz *rf* signal with amplitude $\Delta V \simeq V_v - V_p = 0.4$ V.

[5]GEC Marconi photodiode, model Y-35-52-54: 18 GHz bandwidth, 0.5 A/W peak responsivity.

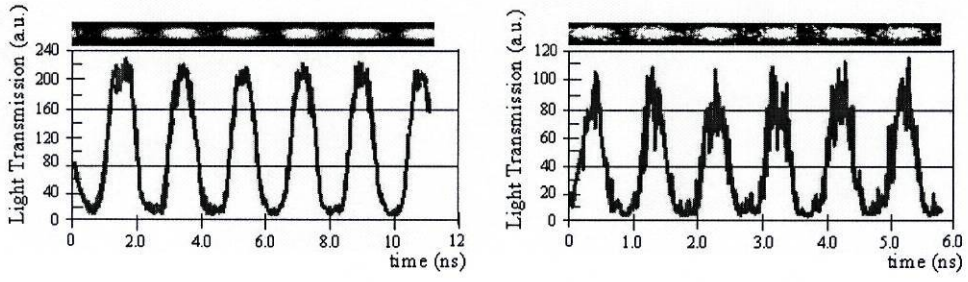


Figure 5.23: Direct optical modulation at around 500 and 950 MHz, showing modulation depth up to 11 dB ($\lambda = 908$ nm).

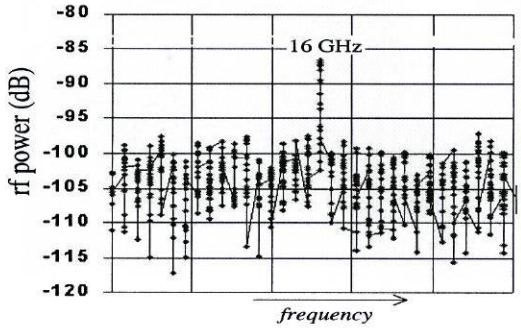


Figure 5.24: The HP radio frequency spectrum analyser trace of the 16 GHz optical response to a 16 GHz rf signal with 0.4 V amplitude, showing a modulation depth around 10 dB.

This result is a good estimate of the high frequency modulation depth, considering that $\Delta\mathcal{E}$ does not depend stringently on the operation frequency (instead, $\Delta\mathcal{E}$ depends on ΔV and ΔJ). The modulation efficiency, characterized by the bandwidth-to-drive-voltage ratio, defined as the ratio of the operation bandwidth to the operating voltage for 10 dB modulation depth, is higher than 40 GHz/V.

From the data presented in Figs. 5.23 and 5.24 it is reasonable to expect efficient high frequency optical modulation up to the RTD-EAM cut-off frequency.

5.4.3 Modulation Due to Injection Locking

The highly non-linear $I - V$ characteristic in the NDC region of the RTD-EAM makes possible the generation of many high-order harmonics of an injected signal [46]. Optical modulation at a higher harmonic corresponding to the fundamental mode of coplanar waveguide (CPW) transmission lines 5 and 7 mm long (length of the RTD-EAM packages CPW line) was observed.

With dc bias, V_B (Fig. 5.11), in the NDC region ($V_p < V_B < V_v$), and injecting a *rf* signal of 0.4 V amplitude and frequency around 957-960 MHz, optical modulation at around 14 GHz was observed, with devices connected through a 5 mm long CPW transmission line, but chirp is evident. Devices connected via a 7 mm long transmission line, again dc biased in the NDC region, showed optical modulation at 8 GHz ("average frequency") for an injected signal of 0.6 V

amplitude and frequency around 1.5-2.0 GHz.

The streak camera traces for the 14 GHz and 8 GHz optical responses are presented in Fig. 5.25, showing considerable chirp. In both cases, switching occurs between the extremes of the NDC regions, producing sinusoidal-like waveforms. This behaviour is dependent on the bias voltage, rf signal amplitude.

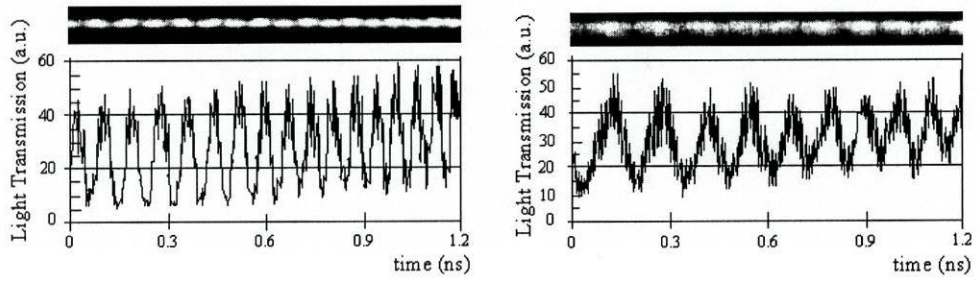


Figure 5.25: Streak camera traces of the optical response associated with locking of RTD-EAM relaxation oscillations to the fundamental mode of packages with CPW transmission lines $\simeq$5 mm and $\simeq$7 mm long, respectively. $\lambda = 908$ nm.

Frequency chirping is a consequence of the non-linear NDC section of the I-V characteristic. The instantaneous modulation frequency depends on the value of the negative dynamic conductance, G_d, which changes when the device operating point follows the low frequency injected signal. Figure 5.26 shows the instantaneous frequency of the 14 GHz optical response in one cicle of the injected signal, and the corresponding signal spectrum.

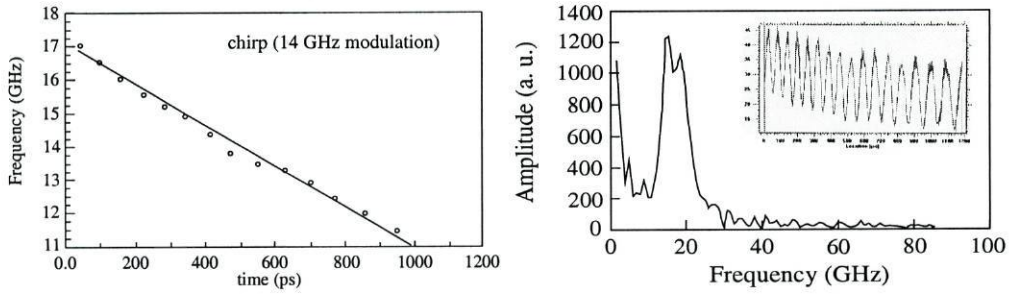


Figure 5.26: Instantaneous frequency of the optical response at $\simeq$14 GHz, and signal spectrum.

By inserting coaxial transmission lines between the packaged devices and the bias tee, optical modulation at the coaxial transmission line fundamental mode frequency was also observed, corresponding to harmonics of lower frequency injected signals.

When scanning the frequency of the injected signal, with amplitude in the range $0.8 - 1.4$ V, it was found that the packaged RTD-EAM, connected through a coaxial line 30 cm long, produced optical pulses with a full width at half maximum (FWHM) of approximately 31 ps, also as a result of device operation as a relaxation oscillator. Figure 5.27 displays the streak camera traces of two optical pulses with 33 and 31 ps FWHM, obtained at slightly different bias conditions. The traces show a maximum modulation depth around 18 dB, which suggests that the RTD-EAMs

now switch between two astable points well into the PDC regions.

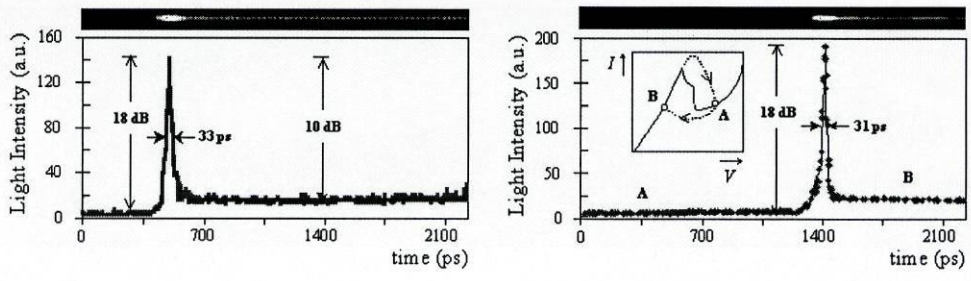


Figure 5.27: Streak camera measurements of the 33 and 31 ps pulse optical responses, showing modulation depth around 18 dB ($\lambda = 908$ nm).

The pulses are caused by the discharge of the RTD-EAM capacitance when the RTD-EAM switches from the second PDC region into the first PDC region [46]. Following the explanation of [46] and the inset of Fig. 5.27, at point A (the upper dwell zone of the relaxation oscillator, where the collector is fully depleted and the electric field is high), the transmission is low. A return pulse forces the trajectory down below the valley, turning up a few picoseconds later, towards point B (the lower dwell zone). During switching, the RTD-EAM capacitance first discharges, strongly decreasing the electric field across the depleted region, which gives rise to high transmission; then the RTD-EAM capacitance starts to recharge, causing the electric field to rise to the astable value associated with the first dwell zone (B). Transmission at point B is higher than at point A, because the electric field is considerably lower in B. Switching from B to A recharges the RTD-EAM capacitance (increasing the electric field across the collector), giving rise to a drop in the transmission. This transmission decrease is not evident in Fig. 5.27 because of the short time scale operation of the streak camera required to resolve the picosecond pulses. On a longer time scale, streak camera traces indicate a pulse repetition rate around 166 MHz, further confirmed by analysis of the optical signal with a 2 GHz bandwidth photo-detector. The relaxation oscillation waveforms are square-wave like.

Figure 5.28 presents typical streak camera traces of the optical response of RTD-EAMs connected to a 15 cm long coaxial line, in 2 ns and 6 ns time scales, showing pulse repetition rates around 340 MHz ($\lambda = 908$ nm). In the shorter time scale, some responses show residual oscillations.

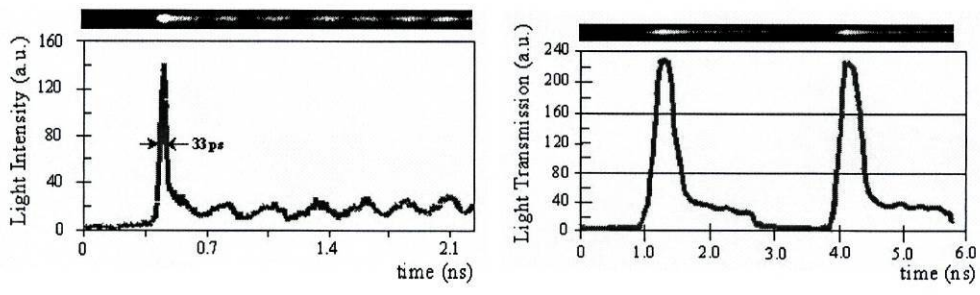


Figure 5.28: Streak camera traces of "33 ps" pulses with repetition rate around 340 MHz.

From this data, an extrapolation of the device optical bandwidth gives a value of about 30 GHz, possibly limited by the switching time of the RTD-EAMs ($t_R \simeq 8$ ps).

5.5 Final Comments

The AlGaAs based RTD-EAM was successfully implemented on GaAs substrates. AlGaAs/GaAs unipolar waveguide modulator structures, incorporating two different DBQW designs, and AlGaAs/GaAs unipolar waveguides bare of DBQW, were fabricated and tested. The waveguides embedding a 7 nm wide GaAs quantum well and 1.4 nm AlAs barriers showed typical RTD negative differential conductance behaviour; devices based on a 5 nm GaAs quantum well and 1.7 nm AlAs barriers presented unreproducible NDC characteristics. These three types of AlGaAs/GaAs unipolar waveguide devices were implemented to improve knowledge about the RTD-EAM through their intercomparison.

The RTD-EAM device layout was modified, resulting on ohmic contacts being applied directly on GaAs instead of AlGaAs, which provided higher quality electrical contacts. A high frequency test package, including chip and wire bonding, and connectorization, was designed and successfully implemented.

Characterized devices presented peak-to-valley current ratios around 1.6, peak current densities up to 13.5 kA/cm^2, and the negative dynamic conductance was typically 0.4 V wide. Simple AlGaAs/GaAs unipolar waveguides showed positive dynamic conductance behaviour. The AlGaAs/GaAs RTD-EAM obtained showed dc bias oscillation, making dc biasing on the NDC difficult to achieve. The relaxation oscillation behaviour demonstrated avoids the difficulties associated with dc biasing. Operation in the resonant tunnelling transmission line relaxation oscillator configuration produced non-sinusoidal waves with high harmonic content, at up to 16 GHz. From the RTD-EAM current-voltage characteristic, and the results obtained, a maximum frequency operation around 20 GHz was estimated (based on the model of the RTD-EAM implemented and on previous results [43]).

Modal characterization of the optical waveguide employing an end-fire microscope objective arrangement showed evidence of at least two optical modes, with propagation loss coefficients of 5.5 cm^{-1} and 6.7 cm^{-1}, determined using the sequential cleaving method for TE and TM polarizations, respectively. The low frequency electro-absorption analysis of 4 μm $\times$ 200 μm active area RTD-EAMs showed modulation depths around 13 dB at $\simeq$908 nm, when a voltage signal with amplitude slightly higher than the NDC voltage width was applied; devices 400 μm long showed a modulation ratio up to 18 dB. From the current-voltage characteristic and the spectral behaviour of the AlGaAs RTD-EAMs, an optical modulation depth of around 9 dB was to be expected.

The streak camera electro-absorption characterization of the same devices showed direct modulation higher than 10 dB at 16 GHz. Electro-absorption modulation associated to RTD-EAM self-oscillation was demonstrated, with a modulation depth up to 12 dB. Optical modulation

due to locking of higher-order harmonics of an injected signal was observed with an extinction ratio around 6 dB at 17 GHz, presenting, however, frequency chirping associated with the non-linear NDC section of the device I-V characteristic. Under determined dc bias conditions, and induced by an injected signal, the RTD-EAM produced optical pulses as short as 31 ps, showing modulation depths up to 18 dB with repetition rates related to the length of the coaxial cable transmission lines employed to feed the devices, for a range of amplitudes and frequencies of the injected signal. From the analysis of the operation in the relaxation oscillation mode, an optical bandwidth around 30 GHz is to be expected.

The AlGaAs RTD-EAM concept can be employed primarily in electrically controlled guided-wave optical systems for the 880 to 1100 nm wavelength range, such as in single waveguide intensity modulators, directional couplers and optical switches. The capability to operate in relaxation oscillation mode can be applied in clock extraction components, for optical pulse generation and demultiplexing in local optical time division multiplexed systems.

However, the AlGaAs RTD-EAM suitability as a practical device in LANs can be affected by the dc operation point instability due to the low frequency dc bias oscillations associated with RTDs, and the resulting frequency chirping can impose extra limitations.

The development of this device in the AlGaAs/GaAs material system was an extremely useful preparation towards the implementation of the RTD-EAM operating at 1550 nm, using a different semiconductor material system, capable of achieving the full potential of this device concept as a high speed intensity modulator in optical communication systems.

Chapter 6

InGaAlAs Based RTD-EAM

This chapter describes the design, fabrication, and characterization of InGaAlAs based RTD-EAMs designed to operate at wavelengths around 1550 nm, where standard single-mode optical fibres have minimum loss.

6.1 InGaAlAs RTD-EAM Design and Fabrication Procedures

The RTD-EAM waveguide core employing the $In_{1-x-y}Ga_xAl_yAs$ composition was designed to have an absorption bandedge below 1550 nm. The same device concept, however, can be implemented to operate around 1300 nm (where standard single-mode optical fibres have zero dispersion), just designing the $In_{1-x-y}Ga_xAl_yAs$ waveguide core with a bandedge absorption below 1300 nm, through an appropriate selection of the core material composition.

6.1.1 $In_{1-x-y}Ga_xAl_yAs$ System

The quaternary system $In_{1-x-y}Ga_xAl_yAs$ ($0.983y + x = 0.468$) can be grown lattice matched to InP, with a bandgap energy, E_g, at room temperature adjustable between that of $In_{0.53}Ga_{0.47}As$ (0.75 eV, 1.65 μm) and that of $In_{0.52}Al_{0.48}As$ (1.439 eV, 0.86 μm) [12]:

$$E_g(z) = 0.76 + 0.49z + 0.20z^2,$$

where $z = y/0.48$. This spectral range contains the region of lowest loss and lowest dispersion of standard single-mode optical fibres.

The $In_{0.53}Ga_xAl_yAs\Gamma$ conduction bandedge effective mass m is given by [12]

$$m = (0.0427 + 0.0328z)m_0,$$

which, for $z \ll 1$, is considerably smaller than the conduction band effective mass in the GaAs Γ-valley ($0.067m_0$); the light- and heavy-hole effective masses are $\simeq 0.051m_0$ and $\simeq 0.5m_0$, respectively.

The refractive index of the $In_{1-x-y}Ga_xAl_yAs$ alloy, as a function of the aluminium fraction y and of the wavelength λ, can be estimated using the expression [192]:

$$n(\lambda) = A + \frac{B\lambda^2}{\lambda^2 - C^2},$$

with $A(z) = 9.689 - 1.012z$, $B(z) = 1.590 - 0.376z$, and $C(z) = 1102.4 - 702.0z + 330.4z^2$, where $z = y/0.48$.

6.1.2 Ridge Waveguide

The InGaAlAs RTD-EAM waveguide core consists of a layer of $In_{0.53}Ga_{0.42}Al_{0.05}As$, 1 μm thick (refractive index $\simeq$3.56), with a bandgap energy around 0.826 eV (wavelength $\simeq$1500 nm), allowing operation at around 1550 nm. The upper cladding was implemented using a layer of $In_{0.52}Al_{0.48}As$, with a refractive index of $\simeq$ 3.24 [192]. Because the InP refractive index ($\simeq$ 3.17 [193]) is considerably smaller than the $In_{0.53}Ga_{0.42}Al_{0.05}As$ refractive index ($\simeq$ 3.56), the InP substrate acts as the lower cladding region. As previously discussed, the upper cladding layer thickness was made approximately twice the decay length of the fundamental optical mode in this region, typically 300 nm.

Figure 6.1 represents the cross section of the ridge waveguide to be defined by etching away the upper cladding and core layers, forming a 1.3 μm ridge.

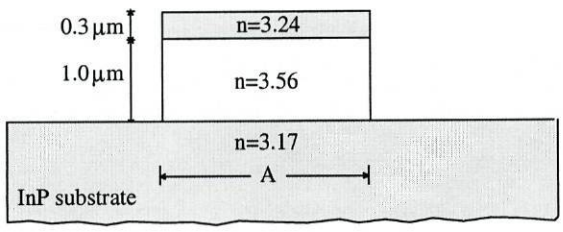


Figure 6.1: $In_{1-x-y}Ga_xAl_yAs$/InP ridge waveguide structure.

As the ridge is surrounded by air, giving a refractive index discontinuity in the transverse direction of $\Delta n \simeq 2.56$, the waveguide can support a higher number of modes in the transverse direction when compared with the vertical direction, in which index differences are around 0.34 – 0.39. Table 6.1 shows the values of the effective refractive index, n_{eff}, for the first three guided modes (for both polarizations, TE and TM), and mode profiles, for the case of a 4 μm wide ridge, obtained using the *FWave IV* software.[1]

6.1.3 Wafer Specification

The high frequency performance of the RTD-EAM is improved by increasing the differential negative conductance, G_n, or by decreasing the series resistance, R_s [5]. To obtain a larger value

[1] M. R. S. Taylor, "FWave IV: a vector E-M wave solver," user manual, University of Glasgow, Glasgow, United Kingdom (URL: http://www.elec.gla.ac.uk/~michaelt/software/fwave/).

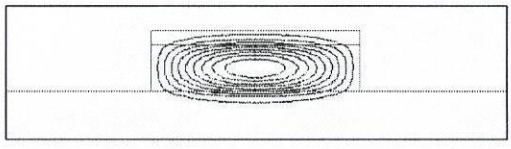
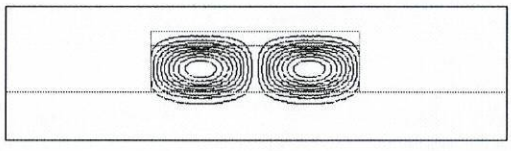
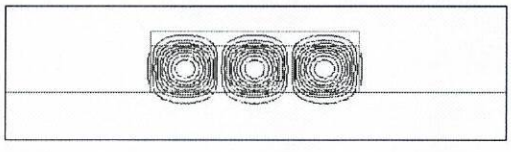
mode	$n_{\text{eff,TE}}$	$n_{\text{eff,TM}}$	*FWave IV* mode profile
1st	3.506	3.502	
2nd	3.487	3.490	
3rd	3.463	3.462	

Table 6.1: Ridge waveguide (4 μm wide) values of the effective refractive index and mode profile, for the first three guided modes obtained using *FWave* ($\lambda = 1550$ nm).

of G_n, it is necessary to achieve a high peak current density, J_p, and high peak-to-valley current ratio (PVCR), J_p/J_v.

The smaller electron effective mass in $In_{0.53}Ga_{0.47}As$ in relation to the GaAs case ($0.041m_0$, compared to $0.067m_0$ for GaAs), gives rise to higher peak current densities in InGaAs/AlAs RTDs due to larger broadening of the resonant levels, when compared with GaAs/AlAs RTDs; and because of the higher energy separation between the Γ and upper (X and L) satellite valleys (Γ_{InGaAs}-X_{AlAs} barrier height of 0.65 eV compared to 0.20 eV for GaAs/AlAs) non-resonant tunnelling parasitic Γ-X mediated transport current flow is reduced, resulting in smaller valley current [61][62]. This gives rise to a RTD-EAM with a larger peak-to-valley current ratio, resulting in higher NDC, when compared with that of GaAs based devices.

The conduction band discontinuities between $In_{0.52}Al_{0.48}As$ and the core material ($\simeq 0.47$ eV), and between the core and the InP substrate ($\simeq 0.26$ eV), can act as "ballistic launching ramps", injecting electrons into the lower bandgap material with high forward momentum, reducing the transit time considerably (the effect is much stronger than in GaAs devices because of the low electron effective mass and the fact that the energy separation between the Γ and the satellite valleys is larger) [61][62]. It is expected that $In_{0.53}Ga_{0.42}Al_{0.05}As$ presents higher electron saturation velocity when compared with GaAs, which will lead, together with the lower contact resistance achieved with InGaAs, to superior cut-off frequency compared to AlGaAs RTD-EAMs.

Because of the higher peak current density expected from InGaAs/AlAs RTDs as compared to similar GaAs devices, the DBQW structure was designed with thicker barriers. The waveguide core background doping concentration was also increased to around 5×10^{16} cm^{-3}, in order to safely avoid breakdown effects. The thicknesses chosen were 2 nm for the AlAs barriers, and the $In_{0.53}Ga_{0.47}As$ well was made 6 nm wide. The upper cladding doping concentration was kept at

$N_{d_{cl}} = 2 \times 10^{18}$ cm^{-3}. In the lower cladding, the InP substrate, the doping concentration was around 2×10^{18} cm^{-3}.

Wafer Structure

The RTD-EAM structure implemented using the InGaAlAs system was grown by Molecular Beam Epitaxy in a Varian Gen II system, on a InP substrate. The waveguide core was formed by two moderately doped (Si: 5×10^{16} cm^{-3}) $In_{0.53}Ga_{0.42}Al_{0.05}As$ layers 500 nm thick (absorption bandedge around 1500 nm and refractive index of 3.56) each side of the double barrier (2 nm thick AlAs and 6 nm wide $In_{0.53}Ga_{0.47}As$ quantum well). The upper cladding was implemented with a 300 nm thick heavily doped (Si: 2×10^{18} cm^{-3}) $In_{0.52}Al_{0.48}As$ layer (refractive index $\simeq$3.24). A 30 nm thick δ-doped $In_{0.53}Ga_{0.47}As$ cap layer was used for formation of non-alloyed Au-Ge-Ni ohmic contacts. Fig. 6.2 shows a diagram of the wafer structure, together with its conduction band Γ-valley and refractive index profiles.

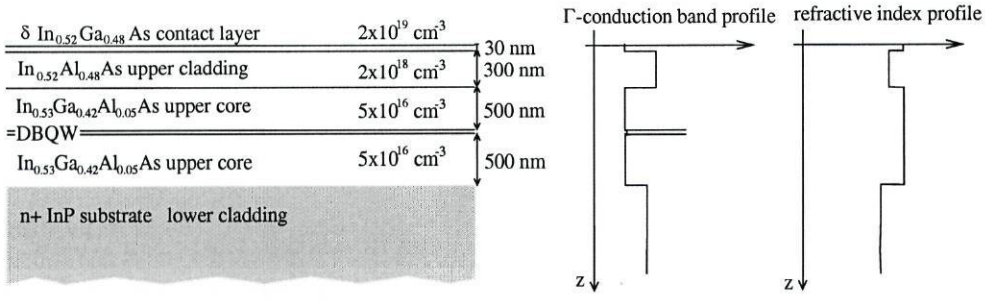


Figure 6.2: InGaAlAs/InP RTD-EAM structure, Γ-valley and refractive index profiles.

Following the design rules discussed above, wafer #A1304 was grown by the MBE group of the Department of Electronics and Electrical Engineering of the University of Glasgow. A wafer was previously grown with the top cladding layer made of $In_{0.53}Ga_{0.47}As$, highly doped (Si: 2×10^{18} cm^{-1}): wafer #B823. Although the devices fabricated from this wafer showed stable NDC, they were very lossy due to high absorption in the upper cladding (bandgap $\simeq$ 0.76 eV).

6.1.4 Fabrication and Testing Procedures

This subsection describes, step by step, the fabrication and testing procedures of InGaAlAs/InP RTD-EAMs. The general processes and technology were discussed in detail in chapter 4.

Phase 1: sample preparation

The wafer was scribed into pieces, approximately 10 mm $\times$ 10 mm. Care must be taken when handling InP substrates, as this material is somewhat weaker than GaAs and suffers higher breakage rates during processing.[2]

[2]Although breakage can be reduced by using thicker wafers, this only exacerbates the device cost issue.

The samples were cleaned in an ultrasonic bath, 5 min for each solvent: opticlear, methanol and acetone, followed by deionised water rinsing and nitrogen blow drying.

Phase 2: ohmic contacts formation

The approach described in section 5.1.3 was followed.

Due to the high solubility of the n-type dopants used in MBE grown $In_{0.53}Ga_{0.47}As$, allowing a dopant concentration much higher than typically 2×10^{18} cm^{-3}, and the relatively low Schottky barrier height for common metals (~0.2 eV compared to 0.8 eV for GaAs), $In_{0.53}Ga_{0.47}As$-AuGe ohmic contacts do not need annealing, showing ohmic behaviour [194]. Figure 6.3 presents the $In_{0.53}Ga_{0.47}As$-AuGe metallization I-V characteristic (a), together with the GaAs-AuGe metallization I-V curve (b) as term of comparison.

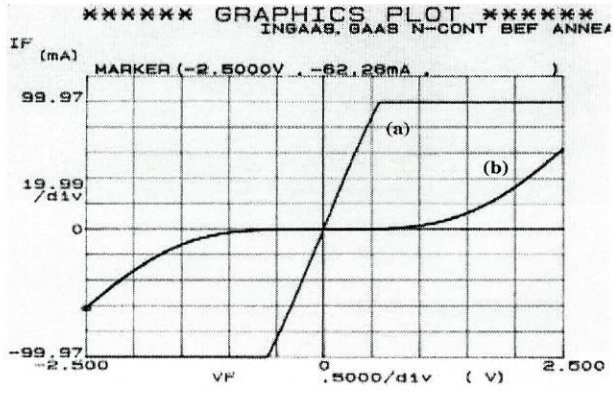


Figure 6.3: (a) $In_{0.53}Ga_{0.47}As$-AuGe ohmic contact I-V characteristic. (b) I-V characteristic of GaAs-AuGe ohmic contact, not annealed.

Phase 3: waveguide wet etching mask, and waveguide definition

The waveguide etching mask, consisting of stripes 4, 5, 6, 7, and 8 μm wide, and 8 mm long, was next formed.[3] The sample was coated with 1400-31 photoresist, spun at 4000 rpm for 30 s, followed by baking for 30 mins @ 90 ^{0}C. The waveguides mask and the sample were aligned, and then exposed for 7 s; photoresist development lasted for 120 s. The samples showing good sets of resist stripes were post-baked for 15 min @ 120 ^{0}C. The unprotected portions of the samples were then etched away, as shown in Fig. 6.4, using the InGaAlAs wet etching solution (3 orthophosphoric acid - 3 H_2O_2 - 200 H_2O, which gives an etching rate of $\simeq$60 nm/min, varying slightly from run to run). This etching solution is selective to InP.

The samples surface was examined under an optical microscope, and the etching depth was checked with a step profiler.

[3]The InGaAlAs waveguide etch mask stripes were slightly wider, to compensate for mask undercut associated with the wet etching waveguide patterning process.

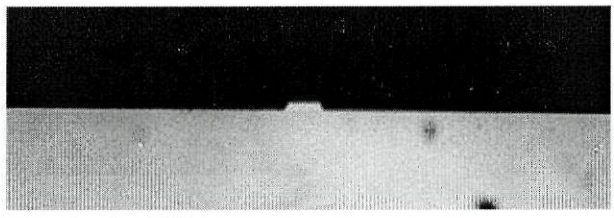


Figure 6.4: Side view of the InGaAlAs waveguide (ridge depth: 1.4 μm; width: 4 μm).

Phase 4: wet etching of the $n+$ InGaAs contact top layer

The waveguide etching mask was removed by soaking the samples in acetone, followed by wet etching of the $\delta-$doped InGaAs waveguide epilayer (uncovered by the metallization) using the alkaline solution of NH_4OH in combination with H_2O_2 and H_2O (1:1:200), for 30 s. This helps to reduce current spreading, present along the ridge axis beyond the length of the contact, and defines the active length of the device. The negative differential conductance characteristic of the RTD-EAM was then measured to evaluate the quality of the wafer structure and fabrication procedure. Figure 6.5 shows typical InGaAlAs/InP RTD-EAMs I-V characteristics from wafer #A1304.

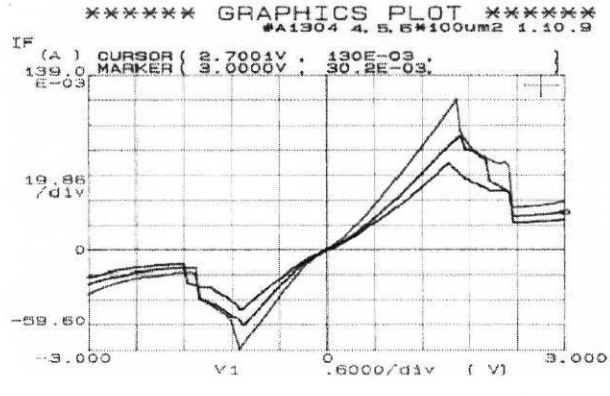


Figure 6.5: Typical InGaAlAs RTD-EAM current-voltage characteristics (from wafer #A1304).

The lower specific contact resistivity ($\sim 10^{-7}$ Ωcm^2) achieved in $In_{0.53}Ga_{0.47}As$ contributes to a smaller peak voltage when compared with GaAs based devices; the $In_{0.53}Ga_{0.47}As$ higher saturation velocity (above 10^7 cm/s) will lead to higher cut-off frequency [2]. (For GaAs/AlGaAs, typical metal to n^+-GaAs contacts have a specific contact resistivity of about 10^{-6} Ωcm^2, and the saturation velocity of electrons in GaAs layers is below 10^7 cm/s [2].)

Phase 5. SiO_2 deposition and contact windows

The procedure adopted has been described in section 5.1.3.

Phase 6: contact pad (CPW) formation

The approach described in section 5.1.3 was followed.

Phase 7: sample thinning and device separation

The InP substrate was hand lapped on a grinding plate, using a suspension of an abrasive powder, until its thickness was $\simeq$150 μm, as described in chapter 4. Great care must be taken not to break the sample or damage the fabricated RTD-EAMs. The devices were separated into 500 μm long by 500 μm wide dies with perfect crystal facets, which act as good end mirrors. The main difficulty associated with RTD-EAM device dicing results from the guiding region not being embedded into the semiconductor, being instead protuberant on the sample surface (Fig. 6.4).

Phase 8: Device packaging

Device packaging followed the steps described in section 5.1.3.

6.2 InGaAlAs RTD-EAM Electrical Characterization

The frequency cut-off and modulation depth characteristics of the RTD-EAM are a direct consequence of the carrier transport mechanisms across the DBQW and the waveguide region. They are closely related to the material system and the specific device structure. Because InGaAs/AlAs RTDs can present higher peak current and smaller valley current densities with a larger peak-to-valley voltage difference, ΔV_{v-p}, when compared with the GaAs system, higher frequency operation and larger modulation depth is to be expected.

6.2.1 dc Characterization

InGaAlAs RTD-EAM dc I-V characteristics were measured using the HP 4145 parametric analyser and showed typical RTD behaviour. Figure 6.6 presents the I-V characteristic of 2 μm $\times$ 100 μm active area RTD-EAMs, showing a peak current density around 18 kAcm^{-2} and PVCR about 7.

However, typical InGaAlAs RTD-EAM I-V characteristics show PVCRs around 4, with peak current density up to 18 kA/cm^2. The specially large PVCR of the I-V curves shown in Fig. 6.6 is mainly due to the high quality achieved in all fabrication procedure steps, specially ohmic contact metallization and wet etching processes of this particular sample.

Figure 6.7 presents the RTD-EAM I-V characteristics for several waveguide active length and width values. Contrary to the AlGaAs RTD-EAM, the InGaAlAs RTD-EAMs I-V characteristic is not strictly anti-symmetric, because the cladding/contact regions are not symmetric in relation to the core, mainly due to the different electrical characteristics of InP and $In_{0.52}Al_{0.48}As$ (as mobility and saturation velocity).

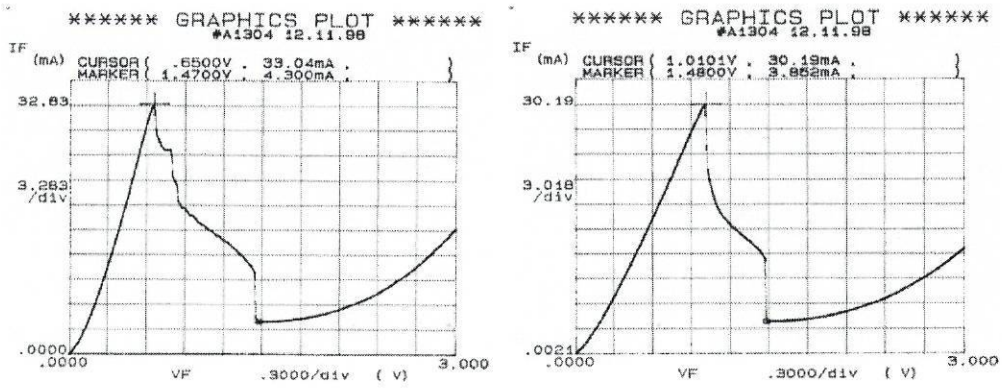


Figure 6.6: I-V characteristics of 2 μm $\times$ 100 μm active area RTD-EAMs.

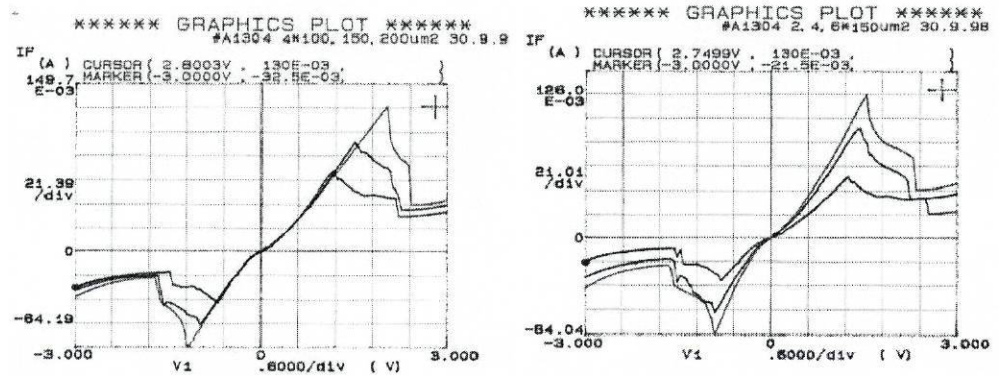


Figure 6.7: RTD-EAM I-V characteristics with waveguide active length (100, 150 and 200 μm; 4 μm wide) and width (2, 4 and 6 μm; 150 μm long), respectively, as parameter.

The peak voltage increases as the active length becomes longer; the same applies as the waveguide becomes wider; in this case, however, the peak voltage increment is much smaller. The associated increment in the peak conductance (I_p/V_p), specially for the wider waveguides, can be attributed to the decreasing importance of the leakage current.

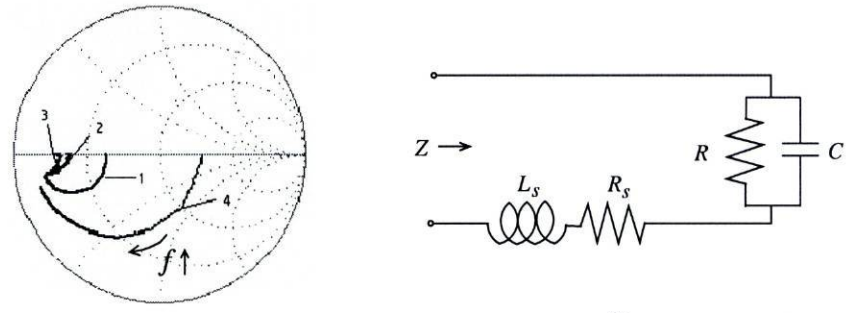
A typical 4 μm $\times$ 200 μm active area RTD-EAM presents a valley-to-peak voltage difference, ΔV_{v-p}, around 0.8 V, with peak-to-valley current density difference, $\Delta J_{p-v} = J_p\,(1 - PVCR^{-1})$, up to 13.5 kA/cm^2. (Typical GaAs/AlAs devices show a PVCR around 1.5, $J_p \simeq$13.5 kA/cm^2, $\Delta V_{v-p} \simeq$0.4 V, and $\Delta J_{p-v} \simeq$5 kA/cm^2.)

6.2.2 High Frequency Characterization[4]

RTD-EAM impedance in the 45 MHz to 18 GHz frequency range was investigated using the Hewlett-Packard 8510 Network Analyser with a Cascade probe station; the testing signal did not exceeded 200 μW. Figure 6.8(a) shows the Smith chart diagram of the RTD-EAM impedance at zero bias (1), -0.5 (2), -0.6 (3), and -2 V (4). RTD-EAM impedance versus frequency dependence for all values of bias voltage was found to correspond to the equivalent circuit consisting

[4]This part of the work was done by V. E. Lvubchenko and N. V. Alkeev at the Institute of Radioengineering & Electronics, Russian Academy of Sciences, Moskow, Russia, on devices supplied.

of a capacity C and a shunting resistor R, in series with a resistance R_s, as shown in Fig.6.8(b); the series inductance was found to be negligible.



(a) (b)

Figure 6.8: (a) Smith chart diagram of the RTD-EAM impedance at zero bias (1), −0.5 (2), −0.6 (3), and −2 V (4), (b) and corresponding equivalente circuit.

At low frequency (45 MHz) the capacitance impedance is high and the total resistance is equal to $R_s + R$; at high frequency (18 GHz) the capacitance shunts R, and the total resistance tends to R_s, as shown in the Smith chart diagram, Fig. 6.8(a).

Figure 6.9 presents the measured $R(V)$, $R_s(V)$ and $C(V)$ curves. The values of R, R_s and C show significant dependence on the bias voltage. The RTD-EAM impedance instability was noticed at bias voltages between −0.7 and −1.6 V, for the tested devices; at forward bias, the instability appears at around 1.16 V (the devices were forward biased below 1.8 V to avoid current beakdown.

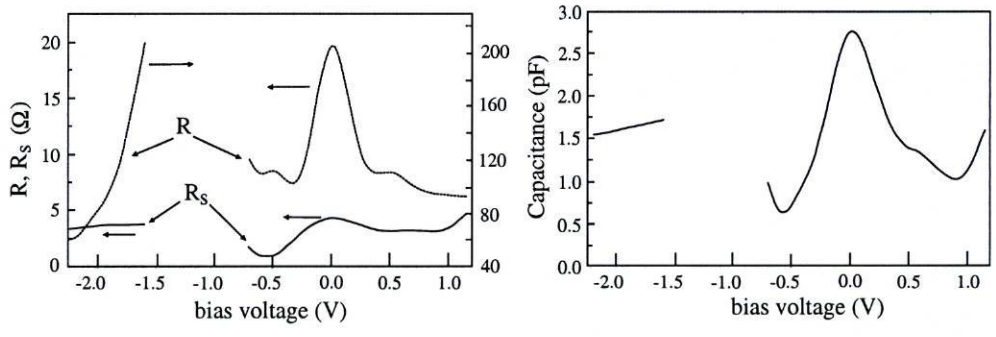


Figure 6.9: RTD-EAM R and R_s resistances, and capacitance dependence on the bias voltage.

The obtained results allow an estimate of the RTD-EAM cut-off frequency, f_{max} (frequency at which the real part of the diode impedance becomes zero, Eqs. 3.29 and 3.30). The negative resistance R is estimated directly from the I-V curve slope, and the capacitance C and the series resistance R_s are assumed to be equal to the RTD-EAM capacity at the bias voltage close to the NDC region. Samples showing $R \approx$-15 Ω, $C \approx$1.5 pF and $R_s \approx$3 Ω give a value of $f_{max} \approx$14.1 GHz.

The modulator bandwidth is related to the 10%-90% switching time, t_R, of the RTD-EAM, between the peak and valley points. The value of t_R can be estimated from $t_R = 4R_dC_d \simeq$

$4\,(\Delta V/\Delta J)\,C_v$ [152][157], where C_v is the capacitance at the valley point per unit area ($C_v = \epsilon/W$, where ϵ is the electrical permitivity, and W is the depletion region width). For the present devices, using $\Delta V_{v-p} \simeq 0.8$ V and $\Delta J_{p-v} \simeq 13.5$ kAcm^{-2}, with $W = 0.5$ μm and $\epsilon = 13\epsilon_0$, t_R is around 5.4 ps. Assuming $f_{max} = (4\pi t_R)^{-1}$, it gives f_{max} ≈14.6 GHz, which is close to the value estimated above.

It is expected that the frequency f_{max} will increase by reducing the waveguide core thickness which leads to a lower peak-to-valley voltage difference, ΔV_{v-p}. By moving at the same time the DBQW towards the emitter highly doped contact layer, a higher peak-to-valley current ratio will be achieved due to a larger effective charge injection because of a decrease in electron scattering in the shorter path from the emitter to the first barrier.

6.3 InGaAlAs RTD-EAM Optical Characterization

Characterization of InGaAlAs based RTD-EAMs employed a Tunics diode laser [186], tuneable in the wavelength region 1480-1580 nm. Optical characterization started with the determination of the $In_{0.53}Ga_{0.42}Al_{0.05}As$ waveguide core bandgap energy, employing photoluminescence techniques. The guided modes were then observed, and the propagation loss was measured using the Fabry-Pérot method. The electrical field induced bandedge broadening was estimated, followed by characterization, at low frequencies, of the modulation depth as a function of the wavelength.

6.3.1 InGaAlAs Waveguide Bandgap and Guided Modes

As pointed out previously, two InGaAlAs/InP wafers were grown, wafers #B823 and #A1304. The waveguide core region bandedge was estimated from their respective photoluminescence spectra.

Wafer Photoluminescence Spectra

Photoluminescence is a non-destructive technique that can be used to determine the bandgap energy, E_g, of a given material [55].[5] This technique was employed to estimate the bandgap energy of the $In_{0.53}Ga_{0.42}Al_{0.05}As$ waveguide. Infrared radiation from a laser with $\lambda < \lambda_g$ ($\lambda \simeq 1.06$ μm) was used to create a population of electronic excited states of the semiconductor material, generating electron-hole pairs which can recombine radiatively, producing the photoluminescence (PL) spectrum [55]. A schematic of the set-up used to measure the PL spectrum is illustrated in Fig. 6.10.[6]

The PL spectra from samples of wafers #B823 and #A1304, at 300 K, were analysed by a spectrometer with a cooled Ge photodetector. The detector output signal was fed to a lock-in

[5] It is also useful to check the presence of certain impurities in semiconductors [55].

[6] If necessary, the sample can be placed in a cryostat which allows low temperature measurements in order to obtain the fullest spectroscopic information by minimising thermally-actived non-radiative recombination processes and thermal line broadening.

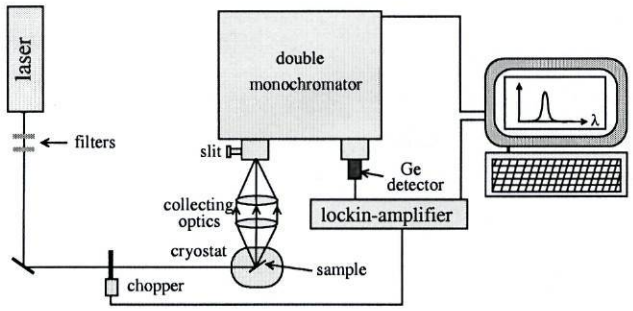


Figure 6.10: Set-up for photoluminescence measurements.

amplifier synchronised to a chopper in the laser beam path for enhanced signal-to-noise ratio. Figure 6.11 presents the photoluminescence data of the #B823 and #A1304 wafers.

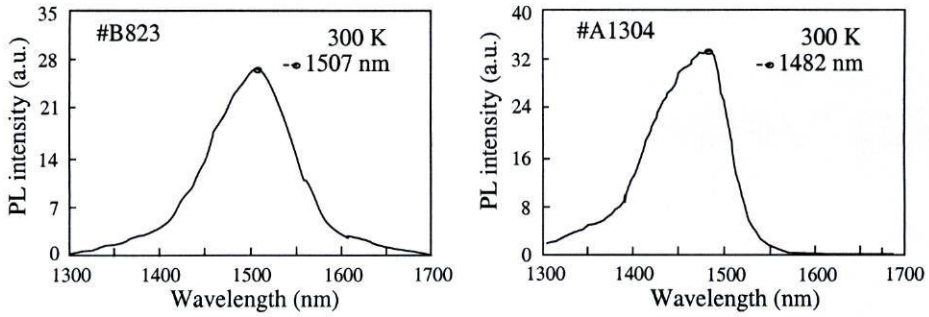


Figure 6.11: Wafers #B823 and #A1304 PL-spectra at 300 K.

The PL spectra present peaks at 1507 nm (B823) and at 1482 nm (#A1304). However, the waveguides fabricated from wafer #B823 were very lossy in the wavelength range 1500-1580 nm, mainly due to the 300 nm thick highly doped $In_{0.53}Ga_{0.47}As$ cladding layer. The wafer #A1304 was found to have reasonably good guiding characteristics.

Guided Modes

Using a single-mode fibre to launch light into the waveguide, it was possible to excite individual modes. Figure 6.12 shows the near field image of the three guided modes observed in a 4 μm wide ridge waveguide fabricated from wafer #A1304, as imaged with the Hamamatsu streak camera operating in the focus mode.

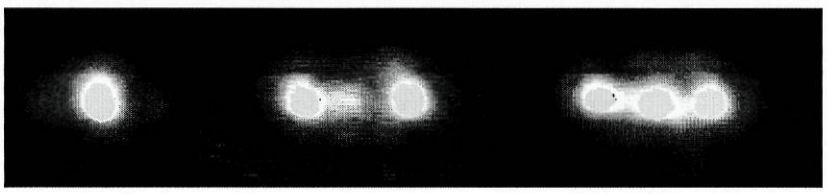


Figure 6.12: Images of the three optical modes observed in a InGaAlAs/InP ridge waveguide (ridge depth: 1.4 μm; width: 4 μm).

6.3.2 Propagation Loss

The Fabry-Pérot method was adopted to characterize the InGaAlAs waveguide guiding loss. This method is based on the fact that a single mode waveguide, with high reflection end facets, forms a Fabry-Pérot étalon, and the contrast of the étalon resonances is related to the combined loss-reflection factor [71]. When the waveguide is not single mode, the observed étalon response is a superposition of the resonance fringes for each mode excited; the same applies if the light source is not single frequency. However, if the propagation length is considerably short, the power transfer between the fundamental mode and the higher order modes can be assumed to be negligible.

Fabry-Pérot Method

The transmittance, T, of an étalon with identical front and back reflection coefficients ($R_1 = R_2 = R$)[7] for light incident normal to the waveguide facets, defined as the ratio of the transmitted P_L to the incident P_0 light powers, Fig. 2.23, is given by [55]

$$T = \frac{(1-R)^2 \exp(-\alpha L)}{1 + R^2 e^{-2\alpha L} - 2Re^{-\alpha L}\cos\phi}, \tag{6.1}$$

where $\phi = 4\pi n_{eff} L/\lambda$ and n_{eff} is the effective modal refractive index. The separation in frequency (wavelength) which leads to consecutive transmission maxima or minima, $\Delta\nu = \frac{c}{2n_{eff}L}$ $\left[\Delta\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda_1} - \frac{1}{\lambda_0} = \frac{1}{2n_{eff}L}\right]$, corresponds to the spectral range of the waveguide étalon [71]. If the resolution of the apparatus is sufficiently high ($\leq 1/2n_{eff}L$), plotting the transmittance as a function of the phase ϕ reveals an oscillatory transmittance curve [71]. From the knowledge of the étalon contrast, given by the ratio of the resonant and anti-resonant transmissions, $C_R = \frac{T_{max}}{T_{min}} = \frac{\left(1+Re^{-\alpha L}\right)^2}{\left(1-Re^{-\alpha L}\right)^2}$, the attenuation coefficient can be expressed as [177]

$$\alpha = -\frac{1}{L}\ln\left[\frac{1}{R}\left(\frac{\sqrt{C_R}-1}{\sqrt{C_R}+1}\right)\right]. \tag{6.2}$$

As a first approximation, the étalon method allows the determination of an upper limit to the attenuation coefficent [58][177]. However, it requires the use of tuneable single frequency lasers or the change of the étalon optical length.

Waveguide Loss

To measure the waveguide loss using the Fabry-Pérot étalon method, a heated soldering iron was placed close to the waveguide, in order to alter the refractive index/length of the waveguide by changing the waveguide temperature. As a consequence of waveguide warming, the light optical path varies and the fraction of light transmitted through the waveguide changes periodically

[7] $R \equiv \frac{(1-n)^2+\kappa^2}{(1+n)^2+\kappa^2}$, where κ is the extinction coefficient and n is the real refractive index [71].

between maxima and minima. The light source employed was a Tunics diode laser emitting at 1565 nm with a linewidth of 10^{-5} nm [186]. Light was launched into the waveguide using a standard single-mode fibre to ensure fundamental mode operation of the device waveguide; however, no information on the state of polarization was available. A fraction of the output power was measured by a photodetector and the contrast extracted, and α determined using Eq. 6.2. Figure 6.13 presents Fabry-Pérot experimental fringe data for a 4 μm wide, 600 μm long waveguide; from this data, the value of $\alpha \simeq 8.2$ cm^{-1} was obtained.

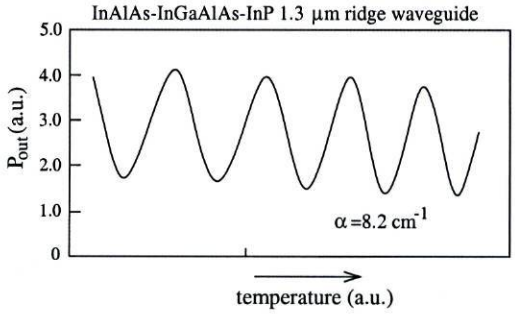


Figure 6.13: Fabry-Pérot fringe data for a 4 μm wide, 600 μm long waveguide.

Because the waveguide is not single mode, the observed response could be a superposition of the resonance fringes for each mode excited in the waveguide, if true operation in single mode is not satisfied. However, for a short propagation length, power transfer between the fundamental mode and the higher order modes is negligible. For this reason, the value of the loss obtained ($\alpha \simeq 8.2$ cm^{-1}) can be used as a good estimate of the propagation loss. The estimated value uses an assumed facet reflectivity value $R \simeq 0.33$. In general, however, the waveguide reflectivity differs from the simple Fresnel value, which can lead to inaccurate results. If a polarization maintaining fibre is used, the polarization can be controlled thus reducing measurement errors, and allowing the distinction between the loss of the two modal polarizations.

It is worth mentionning in this respect that the fibre-optic networks transmission capacity is no longer limited by loss because the signal strength can be boosted using fibre-amplifiers [14].

6.3.3 Waveguide Transmission Spectrum

RTD-EAM peak-to-valley switching induces, through the Franz-Keldysh effect, broadening of the waveguide absorption edge. The InGaAlAs/InP RTD-EAM spectrum change was determined using light from the Tunics diode laser, tuneable in the wavelength region 1480-1580 nm [186].

As previously, the waveguide transmission spectrum was measured, with no applied voltage, when dc biased at slightly below the peak voltage, and in the valley region; the device was not dc biased in the NDC region in order to avoid self-oscillation.

Figure 6.14 presents the InGaAlAs RTD-EAM transmission spectrum in the wavelength range 1500-1580 nm with the applied voltage as parameter.

The electric field enhancement across the depleted portion of the waveguide core due to the

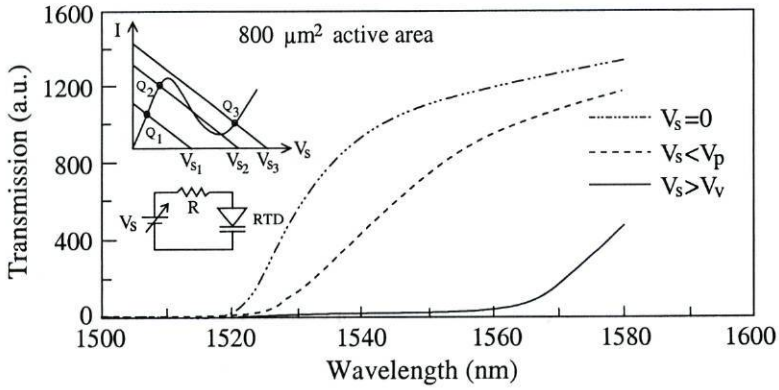


Figure 6.14: InGaAlAs RTD-EAM transmission spectrum at different dc bias voltages. The observed bandedge *shift* was approximately $\Delta\lambda_g \simeq 43$ nm.

RTD peak-to-valley transition, $\Delta\mathcal{E} \equiv \mathcal{E}_v - \mathcal{E}_p$, can be estimated from the dc I-V characteristic, using (Eq. 3.61):

$$\Delta\mathcal{E}_{p-v} \cong \frac{\Delta V_{v-p}}{W} + \frac{W}{2\epsilon v_{sat}}\Delta J_{p-v}. \tag{6.3}$$

Considering typical 4 μm $\times$ 200 μm active area devices with measured values for $\Delta V_{v-p} \simeq 0.8$ V and $\Delta J_{p-v} \simeq 13.5$ kAcm^{-2}, assuming $\epsilon = 14\epsilon_0$, v_{sat}=1x10^7 cm/s, and a depletion region 500 nm wide, $\Delta\mathcal{E}_{p-v}$ becomes 43 kV/cm (for GaAs based devices, $\Delta\mathcal{E}_{p-v} \simeq$ 19 kV/cm). Assuming that the electric field at the valley ($\mathcal{E}_v$) and peak ($\mathcal{E}_p$) points satisfies $\mathcal{E}_v \gg \mathcal{E}_p$, the "shift" in the InGaAlAs waveguide transmission bandedge, $\Delta\lambda_g$, due to electric field enhancement, $\Delta\mathcal{E} \cong \mathcal{E}_v$, as a result of the Franz-Keldysh effect, is given approximately by (Eq. 3.63)

$$\Delta\lambda_g \cong \frac{\lambda_g^2}{hc}\left(\frac{e^2h^2}{8\pi^2 m_r}\right)^{\frac{1}{3}} \mathcal{E}_v^{\frac{2}{3}}, \tag{6.4}$$

where λ_g represents the wavelength corresponding to the waveguide transmission edge at zero bias (around 1520 nm), and m_r is the effective reduced mass (the effective masses of the electrons and holes in the conduction and valence bands are, respectively, $m_e \simeq 0.046m_0$, $m_{lh} \simeq 0.05m_0$ and $m_{hh} \simeq 0.5m_0$).

From the above equation it comes that $\Delta\lambda_g \simeq$46 nm.[8] This estimated value $\Delta\lambda_g$=46 nm agrees reasonably well with the measured figure of 43 nm, regarding the uniform electric field approximation. (For the GaAs based device, the value $\Delta\lambda_g \simeq$ 9 nm was obtained from Eq. 6.4, and the observed spectrum shift was around 12 nm.)

This agreement, again, suggests that Eq. 6.3 can be used, as a first approximation, to determine the magnitude of the electric field enhancement due to peak-to-valley transition.

[8]Eq. 6.4 neglects any spectral shift due to thermal effects from the current flow, and the shift due to the applied peak voltage.

6.3.4 Electro-Absorption Modulation

To measure the electro-absorption modulation induced by the transition between the extremes of the NDC region, the RTD-EAMs were dc biased at slightly above the valley voltage (to minimize thermal effects), and *rf* signals (negative impulse square waves) up to 1 MHz, with amplitude, ≈ 1.0 V, greater than ΔV_{v-p}, were applied. A germanium *pin* photodetector, with a maximum responsivity of 0.7 A/W at around 1550 nm, was employed, as illustrated in Fig. 6.15.[9]

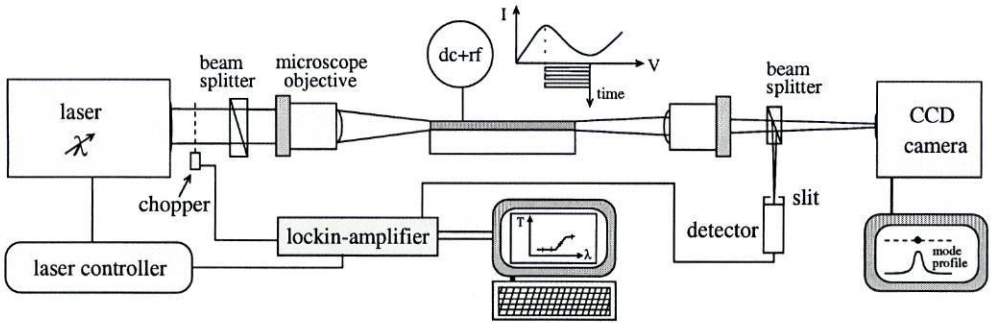


Figure 6.15: InGaAlAs RTD-EAM optoelectronic characterization apparatus.

Devices with 4 μm $\times$ 200 μm active area, showing the highest PVCR and largest ΔV_{v-p}, showed a maximum modulation depth of 28 dB at around 1565 nm. Figure 6.16 presents the modulation depth of these devices as a function of the operating wavelength.

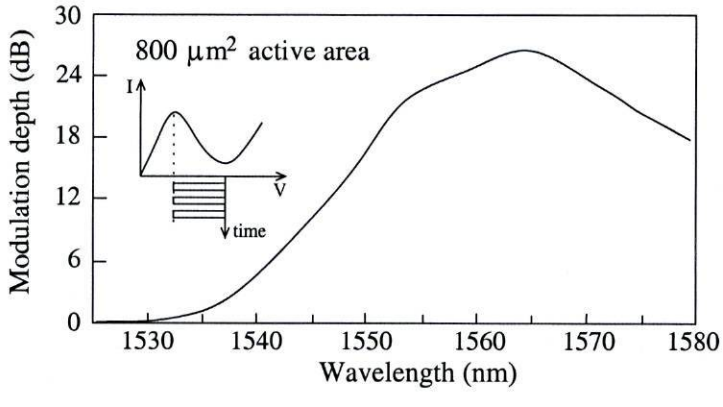


Figure 6.16: InGaAlAs RTD-EAM modulation depth as a function of the wavelength.

Typical 4 μm $\times$ 200 μm active area devices showed a maximum modulation depth around 20 dB in the wavelength range 1560 – 1567 nm (which is approximately 10 dB higher than the maximum value obtained with a GaAs device of equivalent active area).

Comparison of the experimental value of 28 dB for the modulation depth with that calculated from Eqs. 3.64 and 2.113, depends on the availability of the relevant material parameters for InGaAlAs (in particular, the values of the matrix elements $\vec{u} \cdot \vec{p_{cv}}$, section 2.3.4, which were not found in the literature).

[9]Thorlabs DET3-GE.

Dc Modulation Characterization[10]

The experimental set-up for measuring the dc electro-absorption response of the RTD-EAM consisted of an external cavity 1.5 μm tuneable diode laser, which was fibre coupled to the RTD-EAM, and the output from the RTD-EAM was fibre coupled to an optical power meter. The dc frequency characterization of the electro-absorption response of the modulator shows a total change in transmission of around 13 dB as the RTD-EAM biasing voltage was swept through the NDC region, as indicated in Fig. 6.17(a); the I-V curve of one of the RTD-EAMs characterized shows a PVCR around 4 and $\Delta V_{v-p} \simeq 0.5$ V.

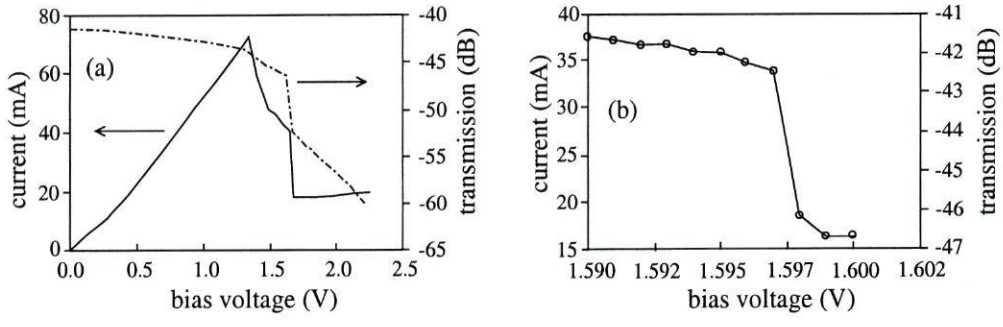


Figure 6.17: (a) Modulation change as the bias voltage is swept through the I-V characteristic. (b) Modulation variation as the bias is swept within the NDC.

Figure 6.17(b) presents the transmission change as the bias voltage is swept within the NDC region, showing a variation in absorption of 5 dB for a 1 mV change, an exceptionally high transmission change for this small drive voltage variation.

6.4 Final Comments

The RTD-EAM modulator concept was implemented, for operation at wavelengths around 1550 nm, with an unipolar $In_{0.52}Al_{0.48}As$-$In_{0.53}Ga_{0.42}Al_{0.05}As$-InP optical waveguide incorporating a $In_{0.53}Ga_{0.47}As$/AlAs double-barrier resonant tunnelling diode (RTD) on a InP substrate. The same material system can also be employed to fabricate a device functioning at wavelengths around 1300 nm (the optical communication window where the standard single-mode optical fibres have zero dispersion).

The RTD-EAM model implemented and the knowledge acquired from the AlGaAs RTD-EAM developed were crucial to the determination of design parameters of the waveguide, of the DBQW, and of the doping level across the structure. Basically, the structures successfully implemented consisted of a waveguide core 1000 nm thick with a doping concentration around 5×10^{16} cm^{-3}, incorporating a 6 nm wide $In_{0.53}Ga_{0.47}As$ quantum well and 1.7 nm AlAs barriers. A RTD-EAM structure employing a $In_{0.53}Ga_{0.47}As$ highly doped top-cladding layer showed reasonably

[10]These measurements were performed by Dr. S. G. McMeekin and co-workers from the Cardiff School of Engineering, University of Wales Cardiff, UK, and Dr. D. G. Moodie from British Telecom Laboratories, Ipswich, UK, on devices supplied.

good NDC characteristics, but very poor optical guiding properties. InGaAlAs-InP waveguide fabrication employed wet etching instead of dry etching, in order to avoid material damage associated to the dry etching gases generally used. The high frequency testing package, including chip and wire bonding, and connectorization, as implemented for characterization of the AlGaAs devices, was also employed.

The dc I-V characteristics of typical devices showed peak current densities up to 18 kA/cm^2, with a peak-to-valley current ratio (PVCR) around 4. (Typical AlGaAs devices showed a PVCR around 1.5, and peak current densities up to 13 kA/cm^2.) Contrary to the AlGaAs RTD-EAM, the InGaAlAs RTD-EAMs I-V characteristic is not strictly anti-symmetric, because the cladding/contact regions are not symmetric in relation to the core, mainly due to the different electrical characteristics of InP and $In_{0.52}Al_{0.48}As$. From the high frequency characterization and the RTD-EAM current-voltage characteristic, a maximum frequency of operation around 30 GHz was estimated, which agrees with the value calculated from the model.

The InGaAlAs RTD-EAMs optical guiding loss was determined using the Fabry-Pérot method. A propagation loss coefficient of 8.3 cm^{-1} was obtained. Because the input polarization state was not well controlled and since the waveguide is multi-mode, the propagation loss value is overestimated. RTD-EAMs with 4 μm $\times$ 200 μm active area, biased slightly below the peak voltage, showed an absorption change up to 28 dB at around 1565 nm (which is approximately 10 dB higher than the maximum absorption change obtained with a GaAs/AlAs device), due to peak-to-valley switching induced by a low-frequency (1 MHz) square wave signal with 1 V amplitude.

Independent dc frequency characterization of the electro-absorption response of a modulator with a PVCR around 4 and $\Delta V_{v-p} \simeq 0.5$ V showed a total change in absorption of around 13 dB as the RTD-EAM biasing voltage was swept through the NDC region; the modulation depth, as the bias voltage is varied within the NDC region, showed a variation in absorption of 5 dB for a 1 mV change, an exceptionally low drive voltage for this amount of modulation.[11] Due to schedule problems, it was not possible to fully characterize the high-speed optoelectronic properties of this modulator. However, from the results obtained and the ones achieved on the AlGaAs modulators, it is expected that the InGaAlAs devices will show significant improvements on high speed modulation characteristics when compared with AlGaAs RTD-EAMs.

[11] Measurements performed at British Telecom Laboratories, Ipswich, by Dr. S. G. McMeekin and co-workers from the Cardiff School of Engineering, University of Wales Cardiff, and Dr. D. G. Moodie from BT Labs, on devices supplied.

Chapter 7

Conclusions and Suggestions for Future Work

The purpose of this research was to investigate a novel optoelectronic device based on an unipolar optical waveguide, incorporating a double barrier quantum well resonant tunnelling diode (DBQW-RTD) in the moderately doped waveguide core region, for light modulation and detection at microwave and millimetre-wave rates: the resonant tunnelling diode optical waveguide (RTD-EAM). Potential for device applications on microwave and millimetre-wave optoelectronics provided the practical motivation for this study of the RTD-EAM properties.

Design, fabrication and characterization of resonant tunnelling diode electro-absorption modulators (RTD-EAMs) were described, in both the AlGaAs-GaAs and the InGaAlAs-InP material systems for shorter and longer wavelengths, respectively. Chapter 1 provided a short review of the state of the art in the domains of Optoelectronics and Optical Communications, followed by a brief description of the devices based on the resonant tunnelling effect. A broad summary of semiconductor properties and device blocks was presented in Chapter 2, emphasizing key electrical and optical phenomena that are employed in the RTD-EAM. Chapter 3 introduced the physics of RTDs, their applications in electronics and optoelectronics, and a discussion of the RTD-EAM concept and its model in some detail. RTD-EAM design principles, together with the general fabrication and characterization procedures were presented in Chapter 4. In Chapter 5, the specific AlGaAs/GaAs RTD-EAM waveguide design and fabrication were presented, and the respective optoelectronic results achieved were illustrated. Chapter 6 described the design, fabrication, and characterization of InGaAlAs based RTD-EAMs designed to operate at wavelengths around 1550 nm, where standard single-mode optical fibres have minimum loss.

7.1 Conclusions

The presence of a RTD within the optical waveguide introduces high non-linearities in the electric field distribution across the waveguide core, made evident by the negative differential conductance in the device I-V characteristic, which is strongly dependent on the bias voltage due to the non-

linear charge distribution in the emitter and collector regions of the RTD, respectively. A small high frequency ac signal (<1 V) can induce high speed switching between the first and the second positive differential conductance regions of the RTD-EAM current-voltage characteristic, which gives rise to electric field modulation across the depleted part of the waveguide core. Electric field enhancement broadens the waveguide absorption bandedge to longer wavelengths via the Franz-Keldysh effect, modulating the optical transmission characteristics of the semiconductor waveguide and, therefore, affecting the guided light at photon energies slightly below the core bandgap energy. Connection of this device to electrical resonators, such as coplanar transmission lines, takes advantage of the negative differential conductance to achieve different modes of operation, including relaxation oscillation, pulsed operation, subharmonic and harmonic injection at microwave frequencies.

The main achievements of this RTD-EAM research can be summarized as follows:

- AlGaAs/GaAs and InGaAlAs/InP based RTD-EAMs operating at 900 nm and 1550 nm, respectively, together with high frequency optoelectronic test packages, were designed and successfully implemented.

- AlGaAs/GaAs based RTD-EAMs operating at around 900 nm were first investigated. The RTD-EAM electrical properties were characterized at low and high frequencies. Device current-voltage characteristics showed peak-to-valley current ratios up to 2, with peak currents around 13.5 $kAcm^{-2}$ and an average valley-to-peak voltage difference of 0.4 V. The electric field amplitude across the depleted core was around 20 $kVcm^{-1}$. Dc and high frequency instability were analyzed and characterized. Operation as a relaxation oscillator, and high frequency (up to 16 GHz) harmonic generation, were both observed.

- Previous results on optical modulation (up to 7 dB at 900 MHz) obtained with the GaAs/AlGaAs system at around 900 nm were confirmed and improved. Direct high frequency (up to 16 GHz) electro-absorption modulation (as high as 18 dB) of light with 908 nm wavelength was demonstrated. Operation as a relaxation oscillation optical modulator was reported for the first time.

- A device model that helped to understand high frequency electrical and optical behaviour was implemented. Based on this model, the determined electric field magnitude across the depleted core was around 20 $kVcm^{-1}$, corresponding to an expected absorption change around 9 dB, which agrees relatively well with the amount of modulation observed considering the approximations made. This model was extremely useful in the design of the RTD-EAM operating at 1550 nm.

- First demonstration of a long wavelength RTD-EAM in the InGaAs-InAlAs material system, lattice matched to InP and operating at around 1550 nm was achieved. The use of a InGaAs/AlAs RTD in a InGaAlAs optical waveguide not only permits operation at optical

communication wavelengths, but also leads to a significant improvement in the optoelectronic characteristics of the device in comparison with the GaAs/AlAs system. Devices showed peak-to-valley current ratios up to 7, peak currents around 18 $kAcm^{-2}$ and an average valley-to-peak voltage difference of 0.8 V. Operation in the relaxation oscillation mode was also observed.

- Low frequency absorption change up to 28 dB, due to peak-to-valley switching, was attained, together with dc electro-absorption results showing a modulation change of 5 dB for a 1 mV variation on the voltage bias. From the optoelectronic device model, the magnitude of the electric field across the depleted core was estimated to be 45 $kVcm^{-1}$, giving an expected absorption change around 19 dB. A more detailed characterization of the optical properties of this modulator is needed; however, from the preliminary results obtained, it is expected that the InGaAlAs RTD-EAM will show significant improvements on high speed modulation characteristics when compared with the AlGaAs modulator.

Considering the performance attained with the devices investigated, some applications of the RTD-EAMs can be directly indicated, such as device operation as:

- low voltage, high bandwidth digital optical modulator;
- self-oscillating modulator for comb frequency generation;
- clock extractor;
- optically operated optoelectronic switch.

In conclusion, the integration of a RTD with an optical waveguide combines a wide bandwidth electrical amplifier with an electro-absorption modulator, which opens up the possibility for a variety of operation modes (such as modulation due to self-oscillation and relaxation oscillation). The device appears to offer a promising route towards a high speed, low power optoelectronic converter (rf-optical and optical-rf).

7.2 Suggestions for Future Work

Further work needs to be carried out to fully characterize the high speed optoelectronic properties of the InGaAlAs RTD-EAM, and verify its expected improved performance over that obtained from GaAs devices. Characterization of the device maximum electrical frequency of operation, and of the optical modulation depth at high frequency, are required. Optimization of the position of the RTD section inside the semiconductor ridge waveguide and waveguide cross-section geometry, in order to minimize the carriers transit time and decrease the required modulation voltage, should be performed. An important issue to be studied is the possibility of modulation induced by a radiating microwave signal coupled to the device through a microstrip antenna. The device optical insertion loss could be reduced using anti-reflection coatings on the waveguide facets and

by packaging in modules incorporating commercially available lens-ended optical fibres, fixed in position by laser welding.

Based on the successful demonstration of the device operating at 1550 nm, it would be interesting to demonstrate the same device concept at 1300nm, using the same material system but modifying the core material composition and, possibly, the optical waveguide cross-section geometry.

In this research, operation of the device as an electro-refraction modulator was not carried out, but it would be a topic worth of further efforts, possibly at 1300 nm and at 1550 nm, and employing either AlGaAs or InGaAlAs material systems.

If successful, this could lead to the implementation of practical integrated modulators based on interference effects.

Following the work of T. S. Moise *et al* [39], an important issue to be investigated is the device mode operation as high speed photo-detector (section 3.4.1). Recently, J. F. Martins-Filho *et al*, using a InGaAlAs RTD-EAM supplied, have demonstrated that a RTD-EAM works also as a photodetector: RTD optical waveguide photo-detector (RTD-OW-PD). The device behaves as a photoconductor integrated with an electrical amplifier (the RTD NDC provides gain on the photogenerated signal). Up to 40 dB relative gain was obtained, and a response time of 2 ns to a 100 ps optical pulse. It is worth noting that the RTD-OW-PD was originally designed as an optical modulator, and was not optimized as a photodetector.[1]

The successful demonstration of the RTD-EAM simultaneous operation both as photo-detector and optical modulator could be the route towards a high speed electro-absorption modulator transceiver, performing both transmission and reception of digital signals. This electro-absorption modulator transceiver could be used in local loop bi-directional fibre optic systems taking advantage of the RTD-EAM electrical gain, avoiding the need for the sophisticated current and temperature control circuitry of a conventional transceiver.

[1] "40 dB photodetection gain in a resonant tunneling diode optical waveguide," J. F. Martins-Filho, R. E. de Araújo, A. S. L. Gomes, J. M. L. Figueiredo, C. R. Stanley, and C. N. Ironside, to be submited. The RTD-OW-PD characterization was performed at Departamento de Electrónica e Sistemas, Universidade Federal de Pernambuco, Recife PE, Brazil.

Part I

References

References

[1] L. L. Chang, L. Esaki, and R. Tsu, "Resonant tunneling in semiconductor double barriers," Appl. Phys. Lett., 1974, **24**, 593.

[2] Hiroshi Mizuta and Tomonori Tanoue, "The physics and applications of resonant tunneling diodes," Cambridge University Press, Cambridge 1995.

[3] J. M. Chamberlain, and R. E. Miles, "New directions in terahertz technology," Nato ASI Series E Applied Sciences, **334**, Kluwer Academic Publishers, The Netherlands 1997.

[4] E. R. Brown, W. D. Goodhue, and T. C. L. G. Sollner, "Fundamental oscillations up to 200 GHz in resonant tunneling diodes and new estimates of their maximum oscillation frequency from stationary-state tunneling theory," J. Appl. Phys., 1988, **64**, 1519.

[5] E. E. R. Brown, T. C. L. G. Sollner, C. D. Parker, W. D. Goodhue, and C. L. Chen, "Oscillations up to 420 GHz in GaAs/AlAs resonant tunneling diodes," Appl. Phys. Lett., 1989, **55**, 1777.

[6] E. R. Brown, J. R. Söderström, C. D. Parker, L. J. Mahoney, K. M. Molvar and T. C. McGill, "Oscillations up to 712 GHz in InAs/AlSb resonant-tunneling diodes," Appl. Phys. Lett., 1991, **58**, 2291.

[7] T. C. L. G. Sollner, W. D. Goodhue, P. E. Tannenwald, C. D. Parker, and D. D. Peck, "Resonant tunneling through quantum wells at frequencies up to 2.5 THz," Appl. Phys. Lett., 1983, **43**, 588.

[8] E. Özbay, and David M. Bloom, "110 GHz monolithic resonant-tunneling diode trigger circuit," IEEE Electron Device Lett., 1991, **12**, 480.

[9] E. Özbay, David M. Bloom, D. H. Chow, and J. N. Schulman, "1.7 ps, microwave, integrated circuit compatible InAs/AlSb resonant tunneling diodes," IEEE Electron Device Lett., 1993, **14**, 400.

[10] M. Leonard Riaziat, "Introduction to high-speed electronics and optoelectronics," John Wiley & Sons Inc., New York 1996.

[11] S. M. Sze, "High-speed semiconductor devices," John Wiley & Sons Inc., New York 1990.

[12] Shun Lien Chuang, "Physics of optoelectronic devices," John Wiley & Sons Inc., New York 1995.

[13] Koichi Wakita, "Semiconductor optical modulators," Kluwer Academic Publishers, London 1998.

[14] Max Ming-Kang Liu, "Principles and applications of optical communications," Irwin Book, London 1996; Göran Einarsson, "Principles of lightwave communications," John Wiley & Sons, New York 1996.

[15] Mario Dagenais, Robert F. Leheny, and John Crow, "Integrated optoelectronics," Academic Press, London 1995.

[16] Fernando Agulló-López, José Manuel Cabrera, and Fernando Agulló-Rueda, "Electrooptics: phenomena, materials and applications," Academic Press, London 1994.

[17] Sadao Adachi, "Physical properties of III-V semiconductor compounds: InP, InAs, GaAs, GaP, InGaAs, and InGaAsP," John Wiley & Sons, New York 1992; Maurice Quillec, "Materials for optoelectronics," Kluwer Academic Publishers, London 1996.

[18] Govind P. Agrawal, Editor, "Semiconductor lasers: past, present, and future," American Institute of Physics, New York 1995; Peter Vasil'ev, "Ultrafast diode lasers: fundamentals and applications," Artech House Inc., London 1995.

[19] D. A B. Miller, "Optics for digital information processing," Semiconductor Quantum Optoelectronics: From Quantum Physics to Smart Devices, Edited by A. Miller, M. Ebrahimzadeh, and D. M. Finlayson, Proceedings of the 50th Scottish Universities, SUSSP Publications and Institute of Physics Publishing, London 1999.

[20] Ralph Williams, "Modern GaAs processing methods," Artech House, London 1990; S. J. Harrold, "An introduction to GaAs IC design," Prentice Hall, New York 1993.

[21] Avishay Katz, "Indium phosphide and related materials: processing, technology, and devices," Artech House, London 1992.

[22] Claude Weisbuch, and Borge Vinter, "Quantum semiconductor structures: fundamentals and applications," Academic Press Inc., London 1991; Gerald Bastard, "Wave mechanics applied to semiconductor heterostructures," Les Éditions de Physique, Les Ulis Cedex, France 1992.

[23] John H. Davies, "The physics of low-dimensional semiconductors: an introduction," Cambridge University Press, Cambridge 1998.

[24] Peter S. Zory, Jr., Editor, "Quantum well lasers," Academic Press, London 1993.

[25] M. Allovon, and M. Quillec, "Interest in AlGaInAs on InP for optoelectronic applications," IEE Proc.-J, 1992, **139**, 148; D. Olego, T. Y. Chang, E. Silberg, E. A. Caridi, and A. Pinczuk, "Compositional dependence of band-gap energy and conduction-band effective mass of InGaAlAs lattice matched to InP," Appl. Phys. Lett., 1982, **41**, 476.

[26] "Quantum Devices and Applications," Special Issue on Quantum Devices and Their Applications, Proceedings of the IEEE, **87**, 4, April 1999.

[27] R. Tsu, and L. Esaki, "Tunneling in a finite superlattice," Appl. Phys. Lett., 1973, **22**, 562; L. Esaki, and L. L. Chang, "New transport phenomenon in a semiconductor "superlattice"," Physical Review Letters, 1974, **33**, 495; L. L. Chang, L. Esaki, and R. Tsu, "Resonant tunneling in semiconductor double barriers," Appl. Phys. Lett., 1974, **24**, 593.

[28] Federico Capasso, Khalid Mohammed, Alfred Y. Cho, "Resonant tunneling through double barriers, perpendicular quantum transport phenomena in superlattices, and their device applications," IEEE J. Quantum Electron., 1986, **22**, 1853; A. Kastalsky, J. Goldman, and J. H. Abeles, "Possibility of infrared laser in a resonant tunneling structure," Appl. Phys. Lett. 1991, **59**, 2636.

[29] F. Capasso, "Band-structure engineering of new photonic materials: high performance mid-infrared quantum cascade lasers," Semiconductor Quantum Optoelectronics: From Quantum Physics to Smart Devices, Edited by A. Miller, M. Ebrahimzadeh, D. M. Finlayson, Proceedings of the 50th Scottish Universities, SUSSP Publications and Institute of Physics Publishing, London 1999.

[30] Supriyo Datta, "Electronic transport in mesoscopic systems," Cambridge University Press, Cambridge 1995.

[31] J. P. Sun, G. I. Haddad, P. Mazumder, and J. N. Schulman, "Resonant tunneling diodes: models and properties," Proceedings of the IEEE, 1998, **4**. 641.

[32] P. Mazumder, S. Kulkarni, M. Bhattacharya, J. P. Sun, and G. I. Haddad, "Digital circuit applications of resonant tunneling devices," Proceedings of the IEEE, 1998, **4**. 664.

[33] V. A. Wilkinson, M. J. Kelly, and M. Carr, "Tunnel devices are not yet manufacturable," Semicond. Sci. Technol., 1997, **12**, 91.

[34] Y. Kawamura, K. Wakita, and O. Mikami, "InGaAs/InAlAs bistable multiple quantum well lasers with large on/off ratio using the resonant tunneling effect," Appl. Phys. Lett., 1988, **53**, 1462; Yuichi Kawamura, Hiromitsu Asai, Shinji Matsuo, and Chikara Amaro, "InGaAs-InAlAs multiple quantum well optical bistable devices using the resonant tunneling effect," IEEE J. Quantum Electron., 1992, **28**, 308.

[35] Li Chen, R. M. Kapre, Kezhong Hu, and A. Madhukar, "High-contrast optically bistable optoelectronic switch based on InGaAs/GaAs (100) asymmetric Fabry-Perot modulator, detector, and resonant tunneling diode," Appl. Phys. Lett., 1991, **59**, 1523.

[36] I. Gravé, S. C. Kan, G. Griffel, S. W. Wu, A. Sa'ar, and A. Yariv, "Monolithic integration of a resonant tunneling diode and a quantum well semiconductor laser," Appl. Phys. Lett., 1991, **58**, 110.

[37] P. England, J. E. Golub, L. T. Florez, and J. P. Harbison, "Optical switching in a resonant tunneling structure," Appl. Phys. Lett., 1991, **58**, 887.

[38] T. S. Moise, Y.-C. Kao, L. D. Garrett, and J. C. Campbell, "Optically switched resonant tunneling diodes," Appl. Phys. Lett., 1995, **66**, 1104.

[39] T. S. Moise, Y.-C. Kao, C. L. Goldsmith, C. L. Schow, and J. C. Campbell, "High-speed resonant-tunneling photodetectors with low-switching energy," IEEE Photonics Technol. Lett., 1997, **9**, 803.

[40] A. F. Lann, E. Grumann, A. Gabai, J. E. Golub, and P. England, "Phase locking between light pulses and a resonant tunneling diode oscillator," Appl. Phys. Lett., 1993, **62**, 13.

[41] S. G. McMeekin, M. R. S. Taylor, B. Vögele, C. R. Stanley, and C. N. Ironside, "Franz-Keldysh effect in an optical waveguide containing a resonant tunneling diode," Appl. Phys. Lett., 1994, **65**, 1076.

[42] Jerome Faist, Federico Capasso, Deborah L. Sivco, Carlo Sirtori, Albert L. Hutchinson, and Alfred Y. Cho, "Quantum cascade laser," Science, 1994, **264**, 553.

[43] S. G. McMeekin, M. R. S. Taylor, C. R. Stanley, A. H. Kean, and C. N. Ironside, "A high speed resonant tunneling diode electro-absorption modulator," European Conference on Integrated Optics, Switzerland 1993, paper 2-38; S. G. McMeekin, M. R. S. Taylor, and C. N. Ironside, "Optical modulation with a resonant tunneling diode," IEE Proc.-Optoelectron., 1996, **143**, 12.

[44] J. M. L. Figueiredo, C. R. Stanley, A. R. Boyd, C. N. Ironside, S. G. McMeekin, and A. M. P. Leite, "Optical modulation in a resonant tunneling relaxation oscillator," Appl. Phys. Lett., 1999, **74**, 1197.

[45] J. M. L. Figueiredo, A. R. Boyd, C. R. Stanley, C. N. Ironside, S. G. McMeekin, and A. M. P. Leite, "Optical modulation at around 1550 nm in a InGaAlAs optical waveguide containing a InGaAs/AlAs resonant tunneling diode," Appl. Phys. Lett., 1999, **75**, 3443.

[46] E. R. Brown, C. D. Parker, S. Verghese, and M. W. Geis, "Resonant-tunneling transmission-line relaxation oscillator," Appl. Phys. Lett., 1997, **70**, 2787; S. Verghese, C. D. Parker, and E. R. Brown, "Phase noise of a resonant-tunneling relaxation oscillator," Appl. Phys. Lett., 1998, **72**, 2550;

[47] J. M. L. Figueiredo, A. R. Boyd, C. R. Stanley, C. N. Ironside, and A. M. P. Leite, "Optelectronic switching in semiconductor waveguides incorporating a resonant tunneling diode," Photonics in Switching 1999, TOPS Vol. 32, P. Prucnal and D. Blumenthal, Editors, Optical Society of America, Washington DC, 2000; J. M. L. Figueiredo, A. M. P. Leite, C. N. Ironside, C. R. Stanley, S. G. McMeekin, K. Bouris, D. G. Moodie, "Resonant tunneling diode electro-absorption waveguide modulator operating at around 1550 nm," Proceedings of Conference on Lasers and Electro-Optics (CLEO), Optical Society of America, 7-12 May 2000, San Francisco, USA.

[48] Sheng S. Li, "Semiconductor physical electronics," Plenum Press, New York 1993.

[49] S. M. Sze, "Physics of semiconductor devices," Second Edition, John Wiley & Sons, New York 1981.

[50] Charles M. Wolfe, Nick Holonyak, Jr., Gregory E. Stillman, "Physical properties of semiconductors," Prentice-Hall International Editions, London 1989.

[51] R. A. Smith, "Wave Mechanics of Crystalline Solids," John Wiley & Sons, New York 1961.

[52] Brian K. Tanner, "Introduction to the physics of electrons in solids," Cambridge University Press, Cambridge 1995.

[53] Jasprit Singh, "Semiconductor Optoelectronics: physics and technology," McGraw-Hill Inc., London 1995.

[54] Claude Cohen-Tannoudji, Bernard Diu, Franck Laloe, "Quantum mechanics," Volume I, John Wiley & Sons, London 1977.

[55] Dieter K. Schroder, "Semiconductor material and device characterization," John Wiley & Sons Inc., New York 1990.

[56] E. Fred Schubert, "Doping in III-V semiconductors," Cambridge University Press, New York 1993.

[57] Klaus K. Schuegraf, Editor, "Handbook of thin-film deposition processes and techniques," Noyes Publications, New Jersey 1988; Harold L. Grubin, David K. Ferry, and C. Jacoboni, Editors, "The physics of submicron semiconductor devices," Plenum Press, London 1988; R. A. Stradling, and P. C. Klipstein (Editors), "Growth and characterization of semiconductors," Bristol 1990.

[58] T. Tamir (Editor), "Integrated optics," Springer-Verlag, New York 1989; T. Tamir (Editor), "Guided -Wave Optoelectronics," Springer-Verlag, New York, 1988; Robert G. Hunsperger, "Integrated optics: theory and technology," Springer-Verlag, New York 1984; David Wood, "Optoelectronic semiconductor devices," Prentice Hall, New York 1994.

[59] Dietrich Marcuse, "Theory of Dielectric Optical Waveguides," 2nd Ed. Academic Press, New York 1991; M. J. Adams, "An Introduction to Optical Waveguides," John Wiley and Sons, New York 1981; A. W. Snyder and J. D. Love, "Optical Waveguide Theory," Chapman and Hall, New York 1983.

[60] New Focus Inc., 1998/99 Catalog, Vol. 9, Santa Clara, USA.

[61] J. Böhrer, A. Krost, and D. B. Bimberg, "Composition dependence of band gap and type of lineup in InGaAlAs/InP heterostructures," Appl. Phys. Lett., 1993, **63**, 1918; J. P. Praseuth, M. C. Joncour, J. M. Gérard, P. Hénoc, and M. Quillec, "Growth and characterization of AlGaInAs lattice matched to InP grown by molecular-beam epitaxy," J. Appl. Phys., 1988, **63**, 400; Yoshihiro Sugiyama, Tsuguo Inata, Toshio Fujii, Yoshiaki Nakata, Shunichi Muto, and Satoshi Hiyamizu, "Conduction band edge discontinuity of $In_{0.52}Ga_{0.48}As/In_{0.52}(Ga_{1-x}Al_x)_{0.48}As$ ($0 \leq x \leq 1$) heterostructures," Jpn. J. Appl. Phys., 1986, **25**, L648; Toshio Fujii, Yoshiaki Nakata, Yoshihiro Sugiyama and Satoshi Hiyamizu, "MBE growth of InGaAlAs lattice-matched to InP by pulsed molecular beam method," Jpn. J. Appl. Phys., 1986, **25**, L254; C. R. Stanley, D. Welch, G. W. Wicks, C. E. C. Wood, C. Palmstrom, F. H. Pollack, and P. Parayanthal, "$(Al_xGa_{1-x})_{0.47}In_{0.53}As$: growth and characterization, " Int. Phys. Conf., 1983, **65**, 173.

[62] D. Olego, T. Y. Chang, E. Silberg, E. A. Caridi, and A. Pinczuk, "Compositional dependence of band-gap energy and conduction-band effective mass of InGaAlAs lattice matched to inP," Appl. Phys. Lett., 1982, **41**, 476.

[63] V. K. Reddy, "Characterization of High Frequency Oscillator and Varactor Diodes Grown by Molecular Beam Epitaxy," Ph. D. Dissertation, The University of Texas at Austin, May 1994.

[64] B. F. Levine, "Quantum-well infrared photodetectors," J. Appl. Physics, 1993, **74**, R1; D. Ahn, and S. L. Chuang, "Calculation of linear and nonlinear intersubband optical absorption in a quantum well model with an applied electric field," IEEE J. Quantum Electron., 1987, **QE-23**, 2196.

[65] Brian R. Bennett, Richard A. Soref, Jesús A. Del Alamo, "Carrier-induced change in refractive index of InP, GaAs, and InGaAsP," IEEE J. Quant. Electron., 1990, **QE-26**, 113.

[66] J. M. Liu, Y. C. Chen, and M. Newkirk, "Carrier-induced phase shift and absorption in a semiconductor laser waveguide under current injection," Appl. Phys. Lett., 1987, **50**, 947.

[67] Sadao Adachi, "GaAs, AlAs, and $Al_xGa_{1-x}As$: material parameters for use in research and device applications," J. Appl. Phys., 1985, **58**, R1.

[68] A. Miller, M. Ebrahimzadeh, D. M. Finlayson, (Editors), "Semiconductor Quantum Optoelectronics: From Quantum Physics to Smart Devices," Proceedings of the 50th Scottish Universities, SUSSP Publications and Institute of Physics Publishing, London 1999.

[69] Robert J. Deri, and Eli Kapon, "Low-loss III-V semiconductor optical waveguides," IEEE J. Quantum Electron., 1991, **QE-27**, 626.

[70] Kazuto Noguchi, Osamu Mitomi, and Hiroshi Miyazawa, "Millimeter-wave Ti:$LiNbO_3$ optical modulators," J. ightwave Technology, 1998, **16**, 615.

[71] Max Born and Emil Wolf, "Principles of Optics," Sixth Edition, Pergamon Press, Oxford 1980.

[72] Kwok Wah Hui, Kin Seng Chiang, Boyu Wu, and Z. H. Zhang, "Electrode optimization for high-speed traveling-wave integrated optic modulators," J. ightwave Technology, 1998, **16**, 232; Weidong Wang, Robert Tavlykaev, and Ramu V. Ramaswamy, "Bandpass traveling-wave Mach-Zehnder modulator in $LiNbO_3$ with domain reversal," IEEE Photon. Technol. Lett., 1997, **9**, 610.

[73] Yi-Jen Chiu, Siegfried B. Fleischer, and John E. Bowers, "High-speed low-temperature-grown GaAs p-i-n traveling-wave photodetector," IEEE Photon. Technol. Lett., 1998, **10**, 1012.

[74] L. D. Westbrook, L. Noel, and D. G. Moodie, "Full-duplex, 25 km analogue fibre transmission at 120 Mbit/s with simultaneous modulation and detection in an electroabsorption modulator," Electron. Lett., 1997, **33**, 694; L. Noel, L. D. Westbrook, D. G. Moodie, and D. Nesset, "120 Mbit/s QPSK data and multi-channel TV transmission over 13 km fibre to a 60 GHz mobile radio link using an electroabsorption modulator as a transceiver," Electron. Lett., 1997, **33**, 1285; "Special issue on microwave-optical interactions," Optical and Quantum Electronics, 1998, **30**, numbers 11-12, December 1998.

[75] "Special issue on microwave-optical interactions," Optical and Quantum Electronics, 1998, **30**, numbers 11-12, December 1998.

[76] L. V. Keldysh, "The effect of a strong electric field on the optical properties of insulating crystals," Sov. Phys. JETP, 1958, **34**, 788; K. Tharmalingam, "Optical absorption in the presence of a uniform field," Phys. Review, 1963, **130**, 2204; David E. Aspnes, "Electric field effects on the dielectric constant of solids," Phys. Review, 1967, **153**, 972; H. Shen, and Fred H. Pollak, "Generalized Franz-Keldysh theory of electromodelation," Phys. Review B, 1990, **42**, 7097.

[77] B. O. Seraphin, and N. Bottka, "Franz-Keldysh effect of the refractive index in semiconductors," Phys. Review, 1965, **139**, 560; F. K. Reinhart, "Electroabsorption in $Al_yGa_{1-y}As$-$Al_xGa_{1-x}As$ double heterostructures," Appl. Phys. Lett., 1973, **22**, 372; S. E. Stillman, C. M. Wolfe, C. O. Bozler, and J. A. Rossi, "Electroabsorption in GaAs and its application to waveguide detectors and modulators," Appl. Phys. Lett., 1976, **28**, 544; T. E. Van Eck, L. M. Walpita, W. S. C. Chang, and H. H. Wieder, "Franz-Keldysh electrorefraction and

electroabsorption in bulk InP and GaAs," Appl. Phys. Lett., 1986, **48**, 451; Brain R. Bennett, and Richard A. Soref, "Electrorefraction and electroabsorption in InP, GaAs, GaSb, InAs, and InSb," IEEE J. Quant. Electron., 1987, **QE-23**, 2159.

[78] S. Wang, "Fundamentals of semiconductor theory and device physics," Prentice-Hall International Editions, London 1989; M. K. Chin, "Comparative analysis of the performance limits of Franz-Keldysh effect and quantum-confined Stark effect electroabsorption waveguide modulators," IEE Proc.-Optoelectron., 1995, **142**, 109; Mee Koy Chin, "An analysis of the performance of Frans-Keldysh Electroabsorption waveguide modulators," IEEE Photon. Technol. Lett., 1995, **7**, 309; P. Baumgartner, C. Engel, G. Bohm, and Abstreiter, "Franz-Keldysh effect in lateral GaAs/AlGaAs based *npn* structures," Appl. Phys. Lett., 1997, **70**, 2876.

[79] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and A. Burrus, "Electric field dependence of optical absorption near the band gap of quantum-well structures," Phys. Review B, 1985, **32**, 1043; John H. Davies, and John W. Wilkins, "Narrow electronic bands in high electric fields: static properties," Phys. Review B, 1988, **38**, 1667.

[80] T. H. Wood, C. A. Burrus, D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, and W. Wiegmann, "High-speed optical modulation with GaAs/GaAlAs quantum wells in a p-i-n diode structure," Appl. Phys. Lett., 1984, **44**, 16.

[81] D. A. B. Miller, D. S. Chemla, S. Schmitt-Rink, "Relation between electroabsorption in bulk semiconductors and in quantum wells: the quantum-confined Franz-Keldysh effect," Phys. Review B, 1986, **33**, 6976; M. K. Chin, "Comparative analysis of the performance limits of Franz-Keldysh effect and quantum-confined Stark effect electroabsorption waveguide modulators," IEE Proc.-Optoelectron., 1995, **142**, 109.

[82] M . Abramowitz, and I. A. Stegun, "Handbook of Mathematical Functions," Chapter 10, Dover, New York 1972.

[83] A. V. Dyadchenko, and E. D. Prokhorov, "Electron drift velocity in $In_xGa_{1-x}As$ compounds," Radio Eng. and Electron Phys. **21**, 151 (1976).

[84] David Bohm, "Quantum theory," Prentice-Hall Inc., New Jersey 1951 .

[85] L.V. Iogansen, "The possibility of resonance transmission of electrons in crystals through a system of barriers," Soviet Physics JETP, 1964, **18**, 146.

[86] E. E. Mendez, W. I. Wang, E. Calleja, and C. E. T. Gonçalves da Silva, "Resonant tunneling via X-point states in AlAs-GaAs-AlAs heterostructures," Appl. Phys. Lett., 1987, **50**, 1263; H. C. Liu, "Resonant tunneling through single layer heterostructures," Appl. Phys. Lett., 1987, **51**, 1019.

[87] H. Yamamoto, "Resonant tunneling condition and transmission coefficient in a symmetrical one-dimensional rectangular double-barrier system," Appl. Phys. A, 1987, **42**, 245.

[88] Masahiro Tsuchiya, Hiroyuki Sakaki, and Junji Yoshino, "Room temperature observation of differential negative resistance in an AlGaAs/GaAs/AlAs resonant tunneling diode," Jpn. J. Appl. Phys., 1985, **24**, L466.

[89] Noaki Harada, and Shigeru Kuroda, "Lifetime of resonant state in a resonant tunneling system," Jpn. J. Appl. Phys., 1986, **25**, L871.

[90] A. P. Stamp, and G. C. McIntosh, "A time-dependent study of resonant tunneling through a doubel barrier," A. J. Phys., 1996, **64**, 264; Daniela Dragoman, and Mircea Dragoman, "Tunneling-time asymmetry in resonant quantum structures," IEEE J. Quantum Electron., 1996, **32**, 1150; M. A. Talebian, and W. Potz, "Open boundary conditions for a time-dependent analysis of the resonant tunneling structure," Appl. Phys. Lett., 1996, **69**, 1148.

[91] M. Tsuchiya, T. Matsusue, and H. Sakaki, "Tunneling escape rate of electrons from quantum well in double-barrier heterostructures," Phys. Review Lett., 1987, **59**, 2356.

[92] M. O. Vassell, Johnson Lee, and H. F. Lockwood, "Multibarrier tunneling in $G_{1-x}Al_xAs/GaAs$ heterostructures," J. appl. Phys., 1983, **54**, 5206.

[93] B. Ricco, and M. Ya. Azbel, "Physics of resonant tunneling. The one-dimensional double barrier case," Physical Review B, 1984, **29**, 1970.

[94] J. N. Schulman, "$Ga_{1-x}Al_xAs$-$Ga_{1-y}Al_yAs$-GaAs double-barrier structures," J. appl. Phys., 1986, **69**, 3954; J. N. Schulman, "Extention of Tsu-Esaki model for effective mass effects in resonant tunneling," Appl. Phys. Lett., 1998, **72**, 2829.

[95] Serge Luryi, "Frequency limit of double-barrier resonant-tunneling oscillators," Appl. Phys. Lett., 1985, **47**, 490; D. D. Coon, and H. C. Liu, "Frequency limit of double barrier resonant tunneling oscillators," Appl. Phys. Lett., 1986, **49**, 94; Arvind S. Vengurlekar, and Federico Capasso, "The effect of incident momentum distribution function on the current-voltage characteristic of resonant tunneling double barriers," Superlattices and Microstructures, 1990, **7**, 89.

[96] H. C. Liu, and D. D. Coon, "Negative resistance of double barrier diodes at zero bias," Appl. Phys. Lett., 1987, **50**, 1669.

[97] J. N. Schulman, and M. Waldner, "Analysis of second level resonant tunneling diodes and transistors," J. Appl. Phys., 1988, **63**, 2859.

[98] E. Wolak, K. L. Lear, P. M. Pitner, E. S. Hellman, B. G. Park, T. Weil, J. S. Harris Jr., and D. Thomas, "Elastic scattering centers in resonant tunneling diodes," Appl. Phys. Lett., 1988, **53**, 201; Huaizhe Xu, and Guanghua Chen, "Effects of an impurity sheet in the well on resonant tunneling in double barrier structures," J. Appl. Phys., 1995, **77**, 5478.

[99] A. Douglas Stone, "Effect of inelastic processes on resonant tunneling in one dimension," Phys. Review Lett., 1985, **54**, 1196;

[100] V. J. Goldman, D. C. Tsui, and J. E. Cunningham, "Observation of intrinsic bistability in resonant-tunneling structures," Phys. Review Lett., 1987, **58**, 1256; T. C. L. G. Sollner, Comment on "Observation of intrinsic bistability in resonant-tunneling structures," Phys. Review Lett., 1987, **59**, 1622; reply: V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Goldman, Tsui, and Cunningham, Phys. Review Lett., 1987, **59**, 1623; N. C. Kluksdahl, A. M. Kriman, and D. K. Ferry, "Self-consistent study of resonant-tunneling diode," Phys. Review B, 1989, **39**, 7720.

[101] H. L. Berkowitz, and R. A. Lux, "Hysteresis predicted in I-V curve of heterojunction resonant resonant tunneling diodes simulated by a self-consistent quantum method," J. Vac. Sci. Technol. B, 1987, **5**, 967.

[102] William R. Frensley, "Wigner-function model of a resonant-tunneling semiconductor device," Phys. Rev. B, 1987, **36**, 1570; William R. Frensley, "Boundary conditions for open quantum systems driven far from equilibrium," Rev. of Modern Physics, 1990, **62**, 745; N. C. Kluksdahl, A. M. Kriman, and D. K. Ferry, "Self-consistent study of resonant-tunneling diode," Phys. Review B, 1989, **39**, 7720; Yuming Hu, and Shawn P. Stapleton, "Double-barrier resonant tunneling transport model," IEEE J. Quantum Electron., 1993, **29**, 327.

[103] Jih-Chen Chiang, Jiann-Shing Shyu, "L-electron effect in AlAs-GaAs-AlAs double-barrier structures," Appl. Phys. Lett., 1997, **70**, 2174.

[104] A. R. Bonnefoi, T. C. McGill, R. D. Burnham and G. B. Anderson, "Observation of resonant tunneling through GaAs quantum well states confined by AlAs X-point barriers," Appl. Phys. Lett., 1987, **50**, 344; E. E. Mendez, W. I. Wang, E. Calleja, and C. E. T. Gonçalves da Silva, "Resonant tunneling via X-point states in AlAs-GaAs-AlAs heterostructures," Appl. Phys. Lett., 1987, **50**, 1263.

[105] Baoquan Sun, Desheng Jiang, Zhenxing Liu, Yaohui Zhang, and Wei Liu, "Evidence of -X sequential resonant tunneling in GaAs/AlAs superlattices," Appl. Phys. Lett., 1996, **69**, 520.

[106] S. Vatannia, and G. Gildenblat, "Airy's functions impelementation of the transfer-matrix method for resonant tunneling in variably spaced finite superlattices," IEEE J. Quantum Electron., 1996, **32**, 1093.

[107] E. Wolak et al, "The design of GaAs/AlAs resonant tunneling diodes with peak current densities over 2×10^5 A/cm^2," J. Appl. Phys., 1991, **69**, 3345-3350.

[108] H. Ohnishi, T. Inata, S. Muto, N. Yokoyama, and A. Shibatomi, "Self-consistent analysis of resonant tunneling devices," Appl. Phys. Lett., 1986, **49**, 1248; M. Cahay, M. McLennan,

S. Datta, and M. S. Lundstrom, "Importance of space-charge effects in resonant tunneling devices," Appl. Phys. Lett., 1987, **50**, 612; J. P. Sun, R. K. Mains, W. L. Chen, J. R. East, and G. I. Haddad, "C-V and I-V characteristics quantum well varactors," J. Appl. Phys., 1992, **72**, 2340.

[109] T. Fiig, and A. P. Jauho, "Self-consistent model for two dimensional accumulation layer states in resonant tunneling devices," Appl. Phys. Lett., 1991, **59**, 2245.

[110] A. D. Stone, and P. A. Lee, "Effect of inelastic processes on resonant tunneling in one dimension," Phys. Rev. Lett., 1985, **54**, 1196; Y. Zohta, and T. Tanamoto, "Improved optical model for resonant tunneling diode," J. Appl. Phys., 1994, **74**, 6996.

[111] Serge Luryi, "Frequency limit of double-barrier resonant-tunneling oscillators," Appl. Phys. Lett., 1985, **47**, 490.

[112] T. Weil, and B. Vinter, "Equivalence between resonant tunneling and sequential tunneling in double-barrier diodes," Appl. Phys. Lett., 1987, **50**, 1281.

[113] S. Mohan, J. P. Sun, P. Mazumder, and G. I. Haddad, "Device and circuit simulation of quantum electronic devices," IEEE Trans. Computer-Aided Design, 1995, **14**, 653.

[114] E. R. Brown, O. B. McMahon, L. J. Mahoney, and K. M. Molvar, "SPICE model of the resonant-tunneling diode," Electronics Lett., 1996, **32**, 938; J. N. Schulman, H. J. De Los Santos, and D. H. Chow, "Physics-based RTD current-voltage equation," IEEE Electron Device Lett., 1996, **17**, 220.

[115] Ralph P. Santoro, "Piecewise-linear modeling of I-V characteristics with SPICE," IEEE Trans. on Education, 1995, **38**, 107.

[116] P. Huang, D. S. Pan, and N. C. Luhmann Jr., "A new method of determination of the I-V characteristic of negative differential conductance devices by microwave reflection coefficient measurements," IEEE Electron Device Lett., 1990, **11**, 570; P. Huang, D. S. Pan, and N. C. Luhmann Jr., "A microwave measurement technique for characterizating the I-V relationship for negative differential conductance devices," IEEE Trans. Microwave Theory Tech., 1993, **41**, 1455.

[117] K. Ismail, B. S. Meyerson, and P. J. Wang, "Electron resonant tunneling in SiGe double barrier diodes," Appl. Phys. Lett., 1991, **59**, 973.

[118] U. Konig, M. Kuisl, J. F. Luy, and F. Schaffler, "Si/SiGe resonant tunneling devices separated by surrounding polysilicon," Electron. Lett., 1990, **26**, 1904.

[119] C. I. Huang, M. J. Paulus, C. A. Bozada, S. C. Dudley, K. R. Evans, C. E. Stutz, R. L. Jones, and M. E. Cheney, "AlGaAs/GaAs double barrier diodes with high peak-to-valley current ratio," Appl. Phys. Lett., 1987, **51**, 121; Daryoosh Vakhshoori, and Shyh Wang,

"Resonant tunneling diodes with AlAs barrier: guides for improving room-temperature operation," J. Appl. Phys., 1987, **62**, 3474.

[120] M. A. Reed, J. W. Lee, and H-L. Tsai, "Resonant tunneling through a double GaAs/AlAs superlattice barrier, single quantum well heterostructure," Appl. Phys. Lett., 1986, **49**, 158.

[121] Tsuguo Inata, Shunichi Muto, Yoshiaki Nakata, Shigehiko Sasa, Toshio Fujii, and Satoshi Hiyamizu, "A pseudomorphic $In_{0.53}Ga_{0.47}As$/AlAs tunneling barrier with a peak-to-valley current ratio of 14 at room temperature," Jpn. J. Appl. Phys., 1987, **26**, L1332; Tom P. E. Broekaert, and Clifton Fonstad, "$In_{0.53}Ga_{0.47}As$/AlAs resonant tunneling diodes with current densities in excess of 450 kA/cm^2," J. Appl. Phys., 1990, **68**, 4310.

[122] Tom P. E. Broekaert, Wai Lee, and Clifton G. Fonstad, "Pseudomorphic $In_{0.53}Ga_{0.47}As$/AlAs/InAs resonant tunneling diodes with peak-to-valley current ratios of 30 at room temperature," Appl. Phys. Lett., 1988, **53**, 1545.

[123] Tsuguo Inata, Shunichi Muto, Yoshiaki Nakata, Toshio Fujii, Hiroaki Ohnishi, and Satoshi Hiyamizu, "Excellent negative differential resistance of InAlAs/InGaAs resonant tunneling barrier structures grown by MBE," Jpn. J. Appl. Phys., 1986, **25**, L983; S. Sen, F. Capasso, A. L. Hutchinson, A. Y. Cho, "Room-temperature operation of $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ resonant tunneling diodes," Electron. Lett., 1987, **23**, 1229.

[124] E. R. Brown, J. R. Soderstrom, C. D. Parker, L. J. Mahoney, K. M. Molvar and T. C. McGill, "Oscillations up to 712 GHz in InAs/AlSb resonant-tunneling diodes," Appl. Phys. Lett., 1991, **58**, 2291; Jih-Chen Chiang, "Tunneling current due to incident electrons derived from the valence bands in AlSb-InAs-AlSb double-barrier structures," Appl. Phys. Lett., 1994, **65**, 2821; A. Sigurdardottir, V. Krozer, and H. L. Hartnagel, "Modeling and design of InAs/AlSb-resonant tunneling diodes," Appl. Phys. Lett., 1995, **67**, 3313.

[125] Masahiro Tsuchiya, and Hiroyuki Sakaki, "Dependence of resonant tunneling current on well widths in AlAs/GaAs/AlAs double barrier diode structures," Appl. Phys. Lett., 1986, **49**, 88; Shunichi Muto, Tsuguo Inata, Yoshihiro Sugiyama, Yoshiaki Nakata, Toshio Fujii, Hiroaki Ohnishi, and Satoshi Hiyamizu, "Quantum well width dependence of negative differential resistance of $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ resonant tunneling barriers grown by MBE," Jpn. J. Appl. Phys., 1988, **26**, L220.

[126] Tom P. E. Broekaert, and Clifton Fonstad, "$In_{0.53}Ga_{0.47}As$/AlAs resonant tunneling diodes with current densities in excess of 450 kA/cm^2," J. Appl. Phys., 1990, **68**, 4310; D. H. Chow, J. N. Schulman, E. Ozbay, and D. M. Bloom, "Investigation of $In_{0.53}Ga_{0.47}As$/AlAs resonant tunneling diodes for high speed switching," Appl. Phys. Lett., 1992, **61**, 1685.

[127] Masahiro Tsuchiya, and Hiroyuki Sakaki, "Precise control of resonant tunneling current in AlAs/GaAs/AlAs double barrier diode with atomically-controlled barrier widths," Jpn. J.

Appl. Phys., 1986, **25**, L185; M. J. Paulus, C. A. Bozada, C. I. Huang, S. C. Dudley, K. R. Evans, C. E. Stutz, R. L. Jones, and M. E. Cheney, "Parametric study of AlAs/GaAS superlattice double-barrier diodes," Appl. Phys. Lett., 1988, **53**, 207.

[128] N. C. Kluksdahl, A. M. Kriman, and D. K. Ferry, "Self-consistent study of resonant-tunneling diode," Phys. Review B, 1989, **39**, 7720; M. Cahay, M. McLennan, S. Datta, and M. S. Lundstrom, "Importance of space-charge effects in resonant tunneling devices," Appl. Phys. Lett., 1987, **50**, 612.

[129] M. J. Paulus, E. T. Koenig, B. Jogai, C. A. Bozada, C. I. Huang, C. E. Stutz, and K. R. Evans, "Effects of a low-doped spacer layer in the emitter of a resonant tunneling diode," Superlattices and Microstructures, 1990, **7**, 135.

[130] N. Mingo, J. A. Porto, and J. Sánchez-Dehesa, "Doping-profile effects on the tunneling times of electrons confined in double-barrier heterostructures," Phys. Rev. B, 1994, **50**, 11884.

[131] Peng Cheng, and James S. Harris Jr., "Effect of Si doping in AlAs barrier layers of AlAs-GaAs-AlAs double-barrier resonant tunneling diodes," Appl. Phys. Lett., 1989, **55**, 572.

[132] J. Jo, K. Alt, and K. L. Wang, "Effect of doping density on capacitance of resonant tunneling diodes," J. Appl. Phys., 1997, **82**, 5206.

[133] Vijay P. Kesan, Dean P. Neikirk, Beg G. Streetman, and Peter A. Blakey, "A new transit-time device using quantum-well injection," IEEE Electron Device Lett., 1987, **8**, 129.

[134] V. P. Kesan, D. P. Neikirk, B. G. Streetman and P. A. Blakey, "The Influence of Transit-Time Effects on the Optimum Design and Maximum Oscillation Frequency of Quantum Well Oscillators," IEEE Trans. Electron Devices, 1988, **35**, 405; Inchae Song, and Dee-Son Pan, "Efficiency and power output of the quantum-well injection transit-time device," IEEE Electron Device Lett., 1987, **8**, 560.

[135] V. P. Kesan, A. Mortazawi, D. R. Miller, T. Itoh, B. G. Streetman, and D. P. Neikirk, "Microwave frequency operation of quantum-well injection transit time (QWITT) diode," Electron. Lett., 1988, **24**, 1473; V. P. Kesan, A. Mortazawi, D. R. Miller, V. K. Reddy, D. P. Neikirk and T. Itoh, "Microwave and Millimetre Wave QWITT Diode Oscillators," IEEE Trans. Microwave Theory Tech., 1989, **37**, 1933-1941; S. Javalagi, V. Reddy, K. Gullapalli, and D. Neikirk, "High efficiency microwave diode oscillators," Electron. Lett., 1992, **28**, 1699.

[136] E. L. Kollberg, "Superconducting devices and quantum well diodes used in low heterodyne receivers, part 2: quantum well diodes," Alta Frequenza, 1989, LVIII, 521.

[137] Sigfrid Yngvesson, "Microwave semiconductor devices," Kluwer Academic Publishers, London 1991.

[138] Runhua Sun, Dee-Son Pan, Tatsuo Itoh, "Simulation of a subharmonic excitation of series integrated resonant tunneling diodes," IEEE Microwave and Guided Wave Lett., 1995, **5**, 1051; Olga Boric-Lubecke, Dee-Son Pan, and Tatsuo Itoh, "RF excitation of an oscillator with several tunneling devices in series," IEEE Microwave and Guided Wave Lett., 1994, **4**, 364.

[139] E. R. Brown, C. D. Parker, and T. C. L. G. Sollner, "Effect of quasibound-state lifetime on the oscillation power of resonant tunneling diodes," Appl. Phys., 1989, **54**, 934.

[140] M. G. Boudreau, and H. C. Liu, "Effects of the quantum inductance on the equivalent crcuit model of a resonant tunneling diode," Superlattices and Microstructures, 1990, **8**, 429.

[141] C. kidner, I. Mehdi, J. R. East, and G. I. Haddad, "Power and stability limitations of resonant tunneling diodes," IEEE Trans. Microwave Theory and Techniq., 1990, **38**, 864; Chih Yuan Huang, James E. Morris, and Yan Kuin Su, "Generalized formula for the stability and instability criteria of current-voltage characteristics measurements in the negative differential conductance region of a resonant tunneling diode," J. Appl. Phys., 1997, **82**, 2690.

[142] C. V. Sammut, and N. J. Cronin, "Comparison of measured and computed conversion loss from a resonant tunneling device multiplier," IEEE Microwave and Guided Wave Lett., 1992, **2**, 486; G. Millington, R. E. Miles, R. D. Pollard, D. P. Steenson, J. M. Chamberlain, "A resonant tunneling diode self-oscillating mixer with conversion gain," IEEE Microwave and Guided Wave Lett., 1993, **1**, 320.

[143] E. R. Brown, and C. D. Parker, "Resonant tunnel diodes as submillimetre-wave sources," Phil. Trans. R. Soc. Lon., 1996, **A354**, 2365.

[144] T. C. L. G. Sollner, P. E. Tannenwald, D. D. Peck, and W. D. Goodhue, "Quantum well oscillators," Appl. Phys. Lett., 1984, **45**, 1319.

[145] E. Schomburg, S. Brandl, K. Hofbeck, T. Blomeier, J. Grenzer, A. A. Ignatov, D. G. Pavel'ev, Yu. Koschurinov, V. Ustinov, A. Zhukov, A. Kovsch, S. Ivanov, and P. S. Kop'ev, "Generation of millimeter waves with a GaAs/AlAs superlattice oscillator," Appl. Phys. Lett., 1998, **72**, 1498; E. Schomburg, K. Hofbeck, J. Grenzer, T. Blomeier, A. A. Ignatov, K. F. Renk, D. G. Pavel'ev, Yu. Koschurinov, V. Ustinov, A. Zhukov, S. Ivanov, and P. S. Kop'ev, "Millimeter waves oscillator based on a quasiplanar superlattice electronic device," Appl. Phys. Lett., 1997, **71**, 1498; K. Hofbeck, J. Grenzer, E. Schomburg, A. A. Ignatov, K. F. Renk, D. G. Pavel'ev, Yu. Koschurinov, B. Melzer, S. Ivanov, S. Schaposchnikov, and P. S. Kop'ev, "High-frequency self-sustained current oscillation in an Esaki-Tsu superlattice monitored via microwave emission," Phys. Lett. A, 1996, **218**, 349.

[146] H. Le Person, C. Minot, L. Boni, J. F. Palmier, and F. Mollot, "Gunn oscillations up to 20 GHz optically induced in GaAs/AlAs superlattice," Appl. Phys. Lett., 1992, **60**, 2396.

[147] A. Rydberg, M. Forsgren, H. Brugger, B. Landgraft, and M. Willander, "Multimode operation of antenna integrated microwave quantum-well-diode oscillator," Microwave and Optical Technology Lett., 1994, **7**, 43; A. Rydberg, and H. Gronqvist, "Quantum-well high-frequency millimeter-wave frequency tripler," Electon. Lett., 1989, **25**, 348; T. C. L. G. Sollner, E. R. Brown, W. D. Goodhue, and C. A. Correa, "Harmonic multiplification using resonant tunneling," J. Appl. Phys., 1988, **64**, 4248.

[148] R. J. Hwu, A. Djuandi, and S. C. Lee, "Negative differential resistance (NDR) frequency conversion with gain," IEEE Trans. Microwave Theory and Techniques, 1993, **41**, 890.

[149] E. R. Brown, C. D. Parker, A. R. Calawa, M. J. Manfra, and K. M. Molvar, "A quasioptical resonant-tunneling-diode oscillator operating above 200 GHz," IEEE Trans. on Microwave Theory and Techniques, 1993, **41**, 720; Elliott R. Brown, Christopher D. Parker, Karen M. Molvar, and Karl D. Stephan, "A quasioptically stabilized resonant-tunneling-diode oscillator for the millimeter- and submillimeter-wave regions," IEEE Trans. on Microwave Theory and Techniques, 1992, **40**, 846.

[150] K. D. Stephan, E. R. Brown, C. D. Parker, W. D. Goodhue, C. L. Chen, T. C. L. G. Sollner, "Resonant-tunneling diode oscillator using a slot-coupled quasioptical open resonator," Electron. Lett., 1991, **27**, 647.

[151] M. Reddy, S. C. Martin, A. C. Molnar, R. E. Muller, R. P. Smith, P. H. Siegel, M. J. Mondry, M. J. W. Rodwell, H. Kroemer, and S. J. Allen, Jr., "Monolithic Schottky-collector resonant tunnel diode oscillator arrays to 650 GHz," IEEE Electron Device Lett., 1997, **18**, 218.

[152] S. K. Diamond, E. Ozbay, M. J. W. Rodwell, and D. M. Bloom, "Resonant tunneling diodes for switching applications," Appl. Phys. Lett., 1989, **54**, 153.

[153] Naofumi Shimizu, Tadao Nagatsuma, Mitsuru Shinagawa, and Takao Waho, "Picosecond-switching time of $In_{0.53}Ga_{0.47}As$/AlAs resonant-tunneling diodes measured by electro-optic sampling technique," IEEE Electron Device Lett., 1995, **16**, 262.

[154] Atsushi Tackeuchi, Tsuguo Inata, Shunichi Muto, and Eizo Miyauchi, "Picosecond characterization of InGaAs/InAlAs resonant tunneling barrier diode by electro-optic sampling," Jpn. J. Appl. Phys., 1989, 28, L750.

[155] Yuming Hu, and Shawn Stapleton, "Quantum capacitance of resonant tunneling diodes," Appl. Phys. Lett., 1991, **58**, 167.

[156] H. C. Liu, and D. D. Coon, "Heterojunction double-barrier diodes for logic applications," Appl. Phys. Lett., 1987, **50**, 1246.

[157] E. R. Brown, C. D. Parker, S. Verghese, and M. W. Geis, "Resonant-tunneling transmission-line relaxation oscillator," Appl. Phys. Lett., 1997, **70**, 2787; S. Verghese, C. D. Parker, and E. R. Brown, "Phase noise of a resonant-tunneling relaxation oscillator," Appl. Phys. Lett., 1998, **72**, 2550.

[158] Woo F. Chow, "Principles of tunnel diode circuits," John Wiley & Sons Inc., London 1964; Herbert J. Reich, "Functional circuits and oscillators," D. Van Nostrand Company Inc., London 1961.

[159] H. Beneking, "High speed semiconductor devices: circuit aspects and fundamental behaviour," Chapman & Hall, Glasgow 1994.

[160] C. R. H. White, M. S. Skolnick, L. Eaves, and M. L. Leadbeater, "Electroluminescence and impact ionization phenomena in a double-barrier resonant tunneling structure," Appl. Phys. Lett., 1991, **58**, 1164.

[161] J. W. Cockburn, P. D. Buckle, M. S. Skolnick, D. M. Whittaker, W. I. E. Tagg, R. A. Hogg, R. Grey, G. Hill, and M. A. Pate, "Electroluminescence recombination from excited-state carrier populations in double-barrier resonant-tunneling structures," Phys. Rev. B, 1992, **45**, 13757; C. Van Hoof, J. Genoe, R. Mertens, G. Borghs, and E. Goovaerts, "Electroluminescence from bipolar resonant tunneling diodes," Appl. Phys. Lett., 1992, **60**, 77; A. Kindlihagen, K. A. Chao, and M. Willander, "Electroluminescence from double-barrier resonant-tunnelling structures," Semicond. Sci. Technol., 1997, **12**, 535; Hanyu Sheng, and Soo-Jin Chua, "Electroluminescence model of bipolar resonant tunnelling diode," Optical Quantum Electron., 1994, **26**, 397.

[162] J. Genoe, C. Van Hoof, K. Fobelets, R. Mertens, and G. Borghs, "*pnp* resonant tunneling light emitting transistor," Appl. Phys. Lett., 1992, **61**, 1051; H. Buhmann, L. Mansouri, J. Wang, P. H. Beton, L. Eaves, and M. Henini, "High efficiency submicron light-emitting resonant tunneling diodes," Appl. Phys. Lett., 1994, **65**, 3332.

[163] C. J. M. Langerak, J. W. Sakai, P. H. Beton, P. C. Main, and L. Eaves, "Photoluminescence of donor energy levels in resonant tunnelling devices," Semicond. Sci. Technol., 1994, **9**, 549; T. S. Turner, L. Eaves, C. R. H. White, M. Henini, and G. Hill, "Quantum well luminescence due to minority photoelectrons in *p*-type resonant tunnelling structures," Semicond. Sci. Technol., 1994, **9**, 552; H. B. Evans, L. Eaves, and M. Henini, "Quantum well luminescence at acceptors in *p-i-n* resonant tunnelling diodes," Semicond. Sci. Technol., 1994, **9**, 555; H. Buhmann, J. Wang, L. Mansouri, P. H. Beton, L. Eaves, M. Heath, and M. Menini, "Photohole-induced resonant tunneling of electrons in selectively etched small area GaAs/AlAs double barrier diodes," Solid-State Electronics, 1994, **37**, 973.

[164] V. Ryzhii, M. Willander, M. Ryzhii, and I. Khmyrova, "Heteostructure laser-transistors controlled by resonant-tunnelling electron extraction," Semicond. Sci. Technol., 1997, **12**, 431.

[165] Pyong Woon Park, Hye Yong Chu, Seon Gyu Han, Young Wan Choi, Gyungock Kim, and El-Hang Lee, "Optical switching mechanism based on charge accumulation effects in resonant tunneling diodes," Appl. Phys. Lett., 1995, **67**, 1241.

[166] H. S. Li, L. P. Chen, Y. W. Chen, K. L. Wang, D. S. Pan, and J. M. Liu, "Dominant photogenerated valley current in a double-barrier resonant-tunneling diode," Appl. Phys. Lett., 1994, **65**, 2999.

[167] S. C. Kan, S. Wu, S. Sanders, G. Griffel, and A. Yariv, "Optically controlled resonant tunneling in a double-barrier diode," J. Appl. Phys. 1991, **69**, 3384.

[168] I. Watanabe, M. Tsuji, K. Makita, and K. Tguchi, "A new planar-structure InAlGaAs-InAlAs superlattice avalanche photodiode with a Ti-implanted guard-ring," IEEE Photonics Technol. Lett., 1996, **8**, 827.

[169] A. F. Lann, E. Grumann, A. Gabai, and J. E. Golub, "Phase locking between light pulses and a resonant tunneling diode oscillator," Appl. Phys. Lett., 1993, **62**, 13.

[170] S. C. Kan, P. J. Harshman, K. Y. Lau, Y. Wang, and W. I. Wang, "Optical control of resonant tunneling diode monolithically integrated with *pin* photodiode," IEEE Photonics Technol. Lett., 1996, **8**, 641.

[171] K. Sano, K. Murata, T. Akeyoshi, N. Shimizu, T. Otsuji, M. Yamamoto, T. Ishibashi and E. Sano, "Ultra-fast optoelectronic circuit using resonant tunnelling diodes and uni-travelling-carrier photodiode," Electron. Lett., 1998, **34**, 215.

[172] Y. Kawamura, H. Asai, and H. Iwamura, "Fabrication of resonant-tunnelling optical bistable laser diodes," Electron. Lett., 1994, **30**, 225; Y. Kawamura, K. Wakita, H. Asai, and K. OE, "Optically bistable operation in InGaAs-InAlAs MQW laser diodes using resonant tunnelling effect," Electron. Lett., 1987, **23**, 719.

[173] O-Kyun Kwon, Kyu-Seok Lee, Hye Yong Chu, and El-Hang Lee, "Abnormal photocurrent-voltage behavior of GaAs/AlGaAs multiple shallow quantum well *p-i-n* diodes," Appl. Phys. Lett., 1998, **72**, 2586.

[174] H. Sakata, K. Utaka, and Y. Matsushima, "Novel bistable device; resonant-tunnelling triangular-barrier optoelectronic switch (R-TOPS)," Electron. Lett., 1994, **30**, 1714; H. Sakata, K. Utaka, and Y. Matsushima, "Diferent types of optoelectronic bistabilities in resonant-tunnelling triangular-barrier optoelectronic switch (R-TOPS)," Electron. Lett., 1995, **31**, 1098.

[175] C. Van Hoof, J. Genoe, R. P. Mertens, E. Goovaerts, and G. Borghs, "Optical switching characteristics of a bistable resonant tunnelling light emitting diode," Electron. Lett., 1992, **28**, 123; C. Van Hoof, J. Genoe, S. Raymond, and G. Borghs, "Giant optical bistable behavior using triple-barrier resonant tunnelling light-emitting diodes," Appl. Phys. Lett., 1993, **63**, 2390.

[176] A. Alping, and L. A. Coldren, "Electrorefraction in GaAs and InGaAsP and its application to phase modulators," J. Appl. Phys., 1987, **61**, 2430; G. E. Stillman, C. M. Wolfe, C. O. Bozler, and J. A. Rossi, "Electroabsorption in GaAs and its application to waveguide detectors and modulators," Appl. Phys. Lett., 1976, **28**, 544; K. Tharmalingam, "Optical absorption in the presence of a uniform field," Phys. Review, 1963, **130**, 2204; Joseph Callaway, "Optical absorption in an electric field," Phys. Rev., 1963, **130**, 549; Joseph Callaway, "Optical absorption in an electric field," Phys. Rev., 1964, **134**, A998.

[177] G. Tittelbach, B. Richter, and W. Karthe, "Comparison of three transmission methods for integrated optical waveguide propagation loss measurement," Pure Applied Optics, 1993, **2**, 683-706; Thomas Feuchter, and Carsten Thirstrup, "High precision planar waveguide propagation loss measurement technique using a Fabry-Perot cavity," IEEE Photonics Technol. Lett., 1994, **6**, 1244; R. Regener, and W. Sohler, "Loss in low-finess Ti:$LiNbO_3$ optical waveguide resonators," Applied Physics B, 1985, **36**, 143; K. H. Park, M. W. Kim, Y. T. Byun, D. Woo, S. H. Kim, S. S. Choi, Y. Chung, W. R. Cho, S. H. Park, and U. Kim, "Nondestructive propagation loss and facet reflectance measurements of GaAs/AlGaAs strip-loaded waveguides," J. Appl. Phys., 1995, **78**, 6318.

[178] Robert A. Pucel, "Design considerations for monolithic microwave circuits," IEEE Trans. on Microwave Theory and Techniq., 1981, **MTT-29**, 513;

[179] Stephen C. Harsany, "Principles of Microwave Technology," Prentice Hall, London 1997; John A. Seeger, "Microwave theory, components, and devices," Prentice-Hall, London ;

[180] K. C. Gupta, Ramesh Garg, I. J. Bahl, "Microstrip Lines and Slotlines," Artech House, Inc., Dedham 1979; Terry Edwards, "Foundations for microstrip circuit design," John Wiley & Sons, London 1992; J. S. Izadian, and S. M. Izadian, "Microwave transition design," Artech House, Dedham 1988.

[181] Cheng P. Wen, "Coplanar waveguide: a surface strip transmission line suitable for nonreciprocal gyromagnetic device applications," IEEE Trans. on Microwave Theory and Techniq., 1969, **MTT-17**, 1087.

[182] Kevin R. Nary, Kiran G. Belare, and Stephen I. Long, "A model for coplanar waveguide transmission line structures on semiconductor substrates," IEEE Trans. on Microwave Theory and Techniq., 1993, **MTT-41**, 2034; Wolfgang Heinrich, "Quasi-TEM description of MMIC coplanar lines including conductor-loss effects," IEEE Trans. on Microwave Theory

and Techniq., 1993, **MTT-41**, 45; A. Gopinath, "Losses in coplanar waveguides," IEEE Trans. on Microwave Theory and Techniq., 1982, **MTT-30**, 1101.

[183] G. Ghione, and C. Naldi, "Analytical formulas for coplanar lines in hybrid and monolithic MICs," Electron. Lett., 1984, **20**, 179; G. Ghione, and C. Naldi, "Parametrs of coplanar waveguides with lower ground plane," Electron. Lett., 1983, **19**, 734; M. Riaziat, I. J. Feng, R. Majidi-Ahy, and B. A. Auld, "Single-mode operation of coplanar waveguides," Electron. Lett., 1987, **23**, 1281; Y. C Shih, and T. Itoh, "Analysis of conductor-backed coplanar waveguide," Electron. Lett., 1982, **18**, 538; Michael Y. Frankel, Shantanu Gupta, Janis A. Valdmanis, and Gerard A. Mourou, "Terahertz attenuation and dispersion characteristics of coplanar transmission lines," IEEE Trans. on Microwave Theory and Techniq., 1991, **MTT-39**, 910.

[184] Alessandro Stano, "Chemical Etching Characteritics of InGaAs/InP and InAlAs/InP Heterostructures," J. Electrochem. Soc.: Solid-tate Science and Technology, 1987, 448.

[185] Model 3900 CW Ti:Sapphire Laser Instruction Manual, Spectra-Physics, Laser Products Division, Mountain View, CA, USA 1989.

[186] "Tunics-PR-1550 wavelength tunable laser diode source: user's guide," Photonetics, Marly-Le-Roi, France 1998.

[187] Yi Chen, Steven Williamson, Tim Brock, F. W. Smith and A. R. Calawa, "375-GHz-bandwidth photoconductive detector," Appl. Phys. Lett., 1991, **59**, 1984.

[188] Picometrix, Ann Arbor, MI, and New Port Corp., 1999 Catalog, Fountain Valley, CA, USA.

[189] "Instruction manual for universal streak camera C5680," Hamamatsu Photonics K. K., Hamamatsu City, Japan 1993.

[190] Kenneth W. DeLong, and Rick Trebino, "Improved ultrashort pulse-retrieval algorithm for frequency-resolved optical gating," JOSA A, 1993, **11**, 2429; Greg Taft, Andy Rundquist, Margaret M. Murname, Henry C. Kapteyn, Kenneth W. DeLong, Rick Trebino, and Ivan P. Christov, "Ultrashort optical waveform measurements using frequency-resolved optical gating," Opt. Lett., 1995, **20**, 743.

[191] R. J. Deri, and M. A. Emanuel, "Consistent formula for the refractive index of AlGaAs below the band edge," J. Appl. Phys. 1995, **77**, 4667; Sadao Adachi (Editor), "Properties of Aluminium Gallium Arsenide," IEE Inspec publication 1993.

[192] M. J. Mondry, D. I. Babic, J. E. Bowers, and L. A. Coldren, "Refractive indexes of (Al, Ga, In)As Epilayers on InP for Optoelectronic Applications," IEEE Photonics Technol. Lett., 1992, **4**, 627; C. De Bernardi, M. Meliga, S. Morasca, C. Rigo, B. Sordo, and A. Stano,

"Refractive indices of InGaAlAs grown by molecular beam epitaxy," J. Appl. Phys., 1990, **68**, 6513; S. Nojima, and H. Asahi, "Refractive index of InGaAs/InAlAs multiquantum-well layers grown by molecular-beam epitaxy," J. Appl. Phys., 1988, **63**, 479.

[193] G. D. Pettit, and W. J. Turner, "Refractive index of InP," J. Appl. Phys., 1965, **36,** 2081.

[194] Pallab Bhattacharya (Editor), "Properties of lattice-matched and strained Indium Gallium Arsenide," EMIS Datareviews Series, No. 8, IEE Inspec publication 1993.