Two and three terminal double barrier resonant tunneling devices

Citation for published version (APA):
Lepsa, M. I. (1997). *Two and three terminal double barrier resonant tunneling devices*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Electrical Engineering]. Technische Universiteit Eindhoven. https://doi.org/10.6100/IR500621

DOI:
10.6100/IR500621

Document status and date:
Published: 01/01/1997

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

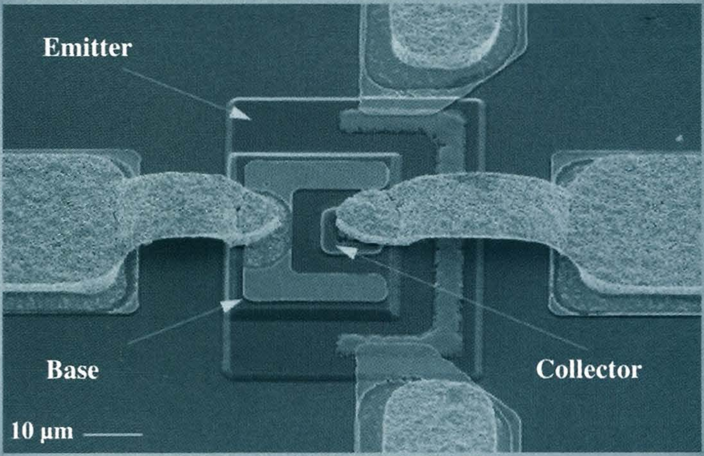
Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 16. Nov. 2023

Two and Three Terminal Double Barrier Resonant Tunneling Devices



M.I. Lepsa

Two and Three Terminal Double Barrier Resonant Tunneling Devices

Front cover: SEM picture of a three terminal double barrier tunneling device.

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Lepsa, Mihail Ion

Two and three terminal double barrier resonant tunneling devices / by Mihail Ion Lepsa. - Eindhoven: Technische Universiteit Eindhoven, 1997.
Proefschrift. - ISBN 90 - 386 - 0390 - 8
NUGI 832
Trefw.: halfgeleiders; tunneleffect / halfgeleidertechnologie / microgolfmetingen.
Subject headings: resonant tunnelling devices / semiconductor technology / microwave measurement.

druk: Universiteitsdrukkerij

Two and Three Terminal Double Barrier Resonant Tunneling Devices

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
Rector Magnificus, prof.dr. M. Rem, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op woensdag 12 november 1997 om 16.00 uur

door

Mihail Ion Lepsa

geboren te Brasov, Roemenië

Dit proefschrift is goedgekeurd door de promotoren:

prof.Dr.-Ing. L.M.F. Kaufmann
en
prof.dr. D. Lenstra

Copromotor:
dr.ir. Th.G. van de Roer

to Anca and our parents

Acknowledgements

This work has been carried out with the contribution and collaboration of many people. Gratefully I acknowledge support and friendship from the members of the Electronic Devices Group of Eindhoven University of Technology. I would like to thank prof. Dr. L. M. F. Kaufmann for giving me the opportunity to realize this work and for permanent help. I thank dr. ir. Th. G. Van de Roer, my copromotor, who has been alongside of me from the very beginning till the end. I acknowledge the critical reviewing of this work and fruitful discussions by prof. dr. D. Lenstra, my second promotor. I thank also drs. E. Smalbrugge for wonderful collaboration during device processing and ing. J. J. M. Kwaspen unstinted help in microwave measurements, from which I learned a lot and I still have to do.

I acknowledge the colaboration of W. C. van der Vleuten from Solid State Department of Eindhoven University for having grown the epitaxial structures. I wish to thank prof. Dr. rer. nat. F. J. Tegude for getting me support in partial processing of the three terminal DBRT devices at Duisburg University, Electronic Devices Department.

Acknowledgements are due to dr. C. Popescu from Research Institute of Material Physics, Bucharest, for permanent encouragement. I also thank dr. T. Stoica from the same institute for help in temperature dependent measurements and von Humboldt foundation for the equipment donation.

Finally I will not forget my family for their permanent support and especially Anca, Simina and Silvana for their patience during the work and preparation of this thesis.

Contents

1. Introduction

2. Resonant Tunneling Device Physics

3. Technology

4. Two Terminal DBRT Devices

5. Three Terminal DBRT Devices

1

Introduction

1.1 Tunneling in Solids

Tunneling is a quantum process in which a particle penetrates into and traverses a barrier region where the potential energy exceeds its initial, kinetic plus potential, energy. According to classical physics, conservation of the energy would not permit this to occur, the particle being totally reflected by the barrier. It is possible in quantum physics, where the particle wave function, and hence its probability density, is spread out over both the classically allowed and classically forbidden parts of the path. Wave functions penetrate from allowed regions into adjoining forbidden regions, for a limited distance, and this allows non-classical motions of the particle that can cross a sufficiently narrow barrier.

Historically the definition of the concept of tunneling and the observation of its experimental manifestations represented one of the early triumphs of the Schrödinger wave mechanics. After Oppenheimer in 1928 [1], used this prediction to describe the autoionization of excited states of atomic hydrogen in a strong electric field, other tunneling processes were conceived and investigated with important consequences in the fields of atomic, nuclear and solid state physics. They included *the field emission* or the escape of an electron from the surface of a metal into the vacuum in an externally high electric field (Fowler and Nordheim 1928 [2], Nordheim 1929 [3]); the penetration of a metal surface barrier by electrons in *thermionic emission* (Frenkel 1929 [4], Nordheim 1929 [3]); and *the alpha decay* of an atomic nucleus in which the alpha particle escapes by tunneling through a barrier of an atomic nucleus (Gamow 1928 [5], Gurney and Condon 1929 [6]). Soon afterwards it was proposed that the current flow between two conductors in electrical contact involved electron tunneling through a potential barrier at the interface (Frenkel 1930 [7]). A short time later the advent of the quantum theory of the (extrinsic) semiconductors (1931) permitted several tries to explain *the rectification* by metal semiconductor junctions (Schottky 1931 [8], Frenkel and Joffe 1932 [9], Wilson 1932 [10]) and construction of a model of "internal field emission" or interband tunneling as an explanation for the dielectric breakdown (Zener 1934 [11]). Unfortunately in both situations other mechanisms proved to have prevalence, the thermally activated charge transfer over the top of the metal semiconductor barrier in the first case and the hot electron effects such as impact ionization, avalanche, etc. in the second.

All these offered the theoretical basis for many phenomena characterized as being "tunneling in solids". Nevertheless the experimental observation of most of these phenomena and their study using simple models occurred in the postwar era, when the material technology developed intensively. As a result of this technological progress another two important events were registered in the field, the invention of the p-n tunnel diode (Esaki 1957 [12]) and the discovery of the tunneling into superconductors (Giaever 1960 [13]). The first one represented the most convincing experimental evidence for tunneling in homogeneous, single-crystal material. The later led to the prediction and observation of weak coupling between superconductors (Josephson 1962 [14], 1963 [15]).

In general, in solids two categories of tunneling phenomena can be involved. Thus, one can speak about the elementary tunneling processes in atomic field, with "microscopic barrier", invoked to explain certain transport phenomena in the bulk semiconductors and also in molecular crystals. On the other hand there are the tunneling processes through macroscopic potential barriers or junctions (the interface of different materials). Concerning this aspect it is interesting to point out that most of the people in the field consider only the last category as referring to "tunneling phenomena in solids" [16, 17]. An explanation could be the great importance which the junction tunneling processes have for the electronic devices.

In their turn the junction tunneling processes also can be divided into two classes according to either the tunneling being in heterojunction systems, e.g. metal-insulator-metal, metal-insulator-semiconductor, supraconductor-insulator-supraconductor, Schottky junctions, etc., or in the same material (semiconductor) as interband tunneling.

A fundamental and difficult question to be asked about junction tunneling in a given system is "*what is the physical origin of the barrier potential?*". Although the answer to this question is obvious in the case of field ionization of atomic hydrogen (microscopic barrier), it ceases to be self-evident for external or internal field emission from a solid (macroscopic barrier). For systems that are in a self-binding state, the physical mechanisms responsible for the barrier potential are primarily the multitude of microscopic interactions that cause the self-binding, rather than a simple one-electron external potential like that due to a proton (ie. hydrogen nucleus). The simplest model of an interface or a tunnel junction is the "independent-electron" model in which the junction region is replaced by an external (one electron) potential. Prior to the discussion of tunneling into superconductors that involves collective-phenomena and thus a many-body approach, this model was exclusively considered in literature. The early applications of tunneling through junctions considered the free-electron model in which "de Broglie waves" were associated to the electrons.

After the development of the quantum theory of solids and in particular that of semiconductors, it was evident that the electron tunneling through junctions involves the relatively complex Bloch states. But to solve the Schrödinger equation even in the adiabatic approximation, in which the motion of the ions and electrons are separated, proved to be practically impossible. Two qualitatively different model simplifications of the Hamiltonian describing the junction have been studied. In the first an adequate treatment of the Coulomb electron-electron interactions is attempted at the expense of neglecting the electron-ion potential altogether by replacing it with a uniform positive background in each part of the junction. This is the so called "jellium" model and has been applied to evaluate the charge density at a metal-vacuum interface (Bardeen [18], Bennett and Duke [19]) and to describe bimetallic and metal-semiconductor contacts (Bennett and Duke [20]). In the second model simplification, an attempt is made to include the periodic component of electron-ion potential *ab initio* by using one electron states appropriate for the relevant periodic potential on each side of the interface, but permitting the charge distribution near the junction to introduce a slowly varying average potential determined by Poisson equation. This average potential is said "to bend the bands" near the interface. In its initial form this approach was developed to describe situations in which the local perturbing non-periodic potential is caused by impurities, electric fields, etc. and is known as the envelope function representation. Supplementary approximations in combination with this are the neglect of band bending in one part (usually the metal) of the junction and the use of WKB (Wentzel, Kramers, Brillouin) or semiclassical approximation, or the effective-mass approximation, depending on whether the junction is gradual or abrupt. The value at the interface of the shift of the energy zero (usually conduction band edge) in one electrode relative to another (which is primarily a consequence of the electron Coulomb interactions) is treated as a phenomenological parameter to be determined from experiment. The effective mass method was introduced by Ben Daniel and Duke [21] for the study of the metal-oxide-semimetal tunnel junctions and abrupt p-n tunnel diode. It consists of considering the description of the electron eigenstates by the envelope functions rather than using the full wave functions and using of appropriate boundary conditions for these to guarantee the probability current conservation. The abrupt-junction effective-mass and the gradual-junction WKB methods do not lead to identical expressions for the tunneling current. It is important to note the distinction between the two methods: the accurate treatment of the external potential at the expense of a quadratic approximation for the band energies which characterizes the effective-mass method and the accurate treatment of the band energies (in the classically forbidden region) at the expense of an approximate treatment of the external potential which characterizes the WKB method.

1.2 Resonant Tunneling in Quantum Structures

Resonant tunneling is distinguished from the simple tunneling process by the presence within the classically forbidden region of the quasibound, or metastable, states of the tunneling particle.

The most typical situation is that one of the double barrier resonant tunneling which is involved in both nuclear and solid state phenomena. In the quantum theory textbook written by Bohm [22] in 1951, the double barrier problem was solved in the WKB approximation. He showed that, at certain energies, unity transmission resonances (resonant tunneling) occur for particles incident upon the structure. These energies correspond to quasibound states between the two barriers. Also the subject was treated by Kane [23] in a general presentation of the basic concepts of tunneling. In his approach Kane describes two methods for the determination of the tunneling probability in square barriers and square wells: the transfer matrix method that starts from the Ben Daniel and Duke's effective mass approximation and the transfer or tunneling Hamiltonian method, well known in the field of superconductivity but originated from the first tunneling approach of Oppenheimer. Later these became useful tools in the treatment of resonant tunneling in quantum heterostructures.

The extension of the double barrier resonant tunneling problem to the more general case of a periodic rectangular potential (Kronig-Penney), also presented in quatum mechanics textbooks [24], offered a simple model of the behaviour of electrons in a crystal lattice. The electron energy spectrum along the periodic potential consists of continuous *bands* separated by forbidden gaps.

The advance of the epitaxial growth techniques, namely Molecular Beam Epitaxy (MBE) and Metal Organic Vapour Phase Epitaxy (MOVPE), made possible the growth of very pure and thin semiconductor layers (especially from III-V compounds) with accuracy up to one atomic layer. In this way it became possible to design and fabricate a large variety of structures characterized by manifestation of the quantum phenomena. Among them, naturally one can say, it was also the double barrier resonant tunneling (DBRT) structure. In conjunction with the interest manifested for the possible high speed microwave applications the DBRT diode became the most studied resonant tunneling device in the last decade.

The DBRT structure consists primary of two layers of semiconductor material, such as AlAs, embedded in another semiconductor having a smaller band gap, such as GaAs. The band diagram of this layer sequence represents a quantum well between two potential barriers. Two electron reservoir layers clad the basic structure resulting the DBRT diode.

In the study of resonant tunneling properties of the DBRT diode two main approaches have been developed. The first one, known as the coherent tunneling model, is based on the calculation of the transmission coefficient for the full structure as one coherent wave propagation, using the Kane's transfer-matrix method. The coherent tunneling description is a true wave-mechanical approach to tunneling: the resonance is due to multiple reflections of the electron waves in the well, whence the analogy with the Fabry-Perot interferometer in optics.

The second approach, the sequential-tunneling model, considers the tunneling as two separate and subsequent processes, from emitter to well and from well to collector. No coherence of the wave function is assumed in the description. The transmission coefficient is obtained using the transfer Hamiltonian method. In this method the Hamiltonian for the full double barrier structure is replaced by three Hamiltonians for each allowed region, separately. For each barrier a "transfer Hamiltonian" is considered which is treated as a first order perturbation in the perturbation theory and from this the transmission coefficient is calculated using the Fermi Golden Rule.

In addition to the two previous models, a third approach, using Wigner distribution functions has to be mentioned [25]. The Wigner distribution function is the Fourier-transformed density matrix in phase space. It is the closest parallel to the classical distribution function, that quantum mechanics can offer. Consequently, many results of classical transport theory can be transferred to a Wigner function based quantum transport theory. The usefulness of the Wigner representation consists in the fact that it permits the evaluation of the time evolution of the resonant tunneling phenomenon as well as the steady state situation.

The practical transposition of the physical situation described by the Kronig-Penney potential has been also possible due to the advanced epitaxial technics. The structures, named *superlattices,* have become both means for fundamental research in solid state physics and component parts of electronic devices.

1.3 Materials for Resonant Tunneling Structures

As we have seen, the resonant tunneling process requires stacked thin semiconductor layers with different band gaps which, energetically speaking, give rise to at least a double barrier structure. The diversity of band alignments between III-V semiconductors suggests, at least in principle, that there are considerable number of possibilities to realize resonant tunneling structures from these materials. To illustrate this, Fig. 1.1 shows the best estimate of the band alignments for eight lattice- or nearly lattice-matched III-V compound heterojunctions. We specify that for a general heterojuction the band gap discontinuity is distributed between the valence and conduction bands. As seen in Fig. 1.1a the corresponding band "*offsets*", ΔE_V and ΔE_C, are abrupt discontinuities in the band edges at the heterojunction interface. Additionally, electrostatic band bending also occurs because of charge redistribution near the interface. The band bending changes (increasing usually) if the heterojunction is doped, but the band offsets is presumably to remain constant. In general, one speaks about an *isotype* junction if the same conductivity type is achieved on both sides of the junction (*n-n* or *p-p*) and an *anisotype* junction if the type of conductivity is opposite (*n-p* or *p-n*). The diagrams in Fig. 1.1, illustrating the band alignments for the different systems, show the behaviour of the E_C and E_V in the immediate vicinity of the heterojunction, that is they figure only the band offsets. According to the relative ordering of the band edge energies we can distinguish several types of band alignments [26]. The most common (and generally considered to be the "normal") alignment is the *straddling* configuration illustrated in the cases (b) [27], (c) [27], (d) [28], (e)[28] and (g)[29] of Fig. 3.1, in which the bandgaps of the two materials entirely overlap. Another situation is when the bandgaps are partial overlapped i.e., the conduction band of the smaller-gap material might lie above that of the larger-gap material, or its valence band might lie below that of the larger-gap material. Such a band alignment is called *staggered* and is known to occur in the cases (f) [30] and (h) [31] presented in Fig. 1.1. The staggering might become so extreme that the bandgaps cease to overlap. This situation is known as a *broken gap* alignment, being illustrated in Fig. 1.1i [29]. Another nomenclature, employed usually in describing superlattices, divides the band alignments in two main categories: type I and type II alignments. If the extrema of both the conduction and valence bands lie both in the same layers, the alignment is referred as type I, whereas if the band extrema are in different layers the alignment is type II. Comparing the two notations, one can easily observe that the type I alignments correspond to the straddling case and the type II alignments include both staggered and broken-gap situations.

Grown on a lattice-matched substrate, as GaAs, InP or GaSb, each of the above systems have been used to demonstrate the resonant tunneling phenomena in double barrier structures (see next section). This case, in which all the layers in and around the double barrier structure are made of materials having the same lattice constant as the substrate, is the most desirable from the point of view of the epitaxial growth of high quality materials. But, some other possibilities, using also not lattice-matched materials, have been successfully demonstrated. One of these is to lattice match the quantum well and cladding layers with the substrate but use a material having a different lattice constant in the barriers. If the barriers are sufficiently thin, their lattice will conform by compression or tension to surrounding layers in the lateral plane, and will deform in the opposite sense in the longitudinal axis. This is called a pseudomorphic match and the resulting layers have few defects (dislocations). Another possibility is to lattice match the barriers but use material of different lattice constant in all or part of the quantum well, or some portion of the cladding layers. Again, if the layers are sufficiently thin, a pseudomorphic match results.

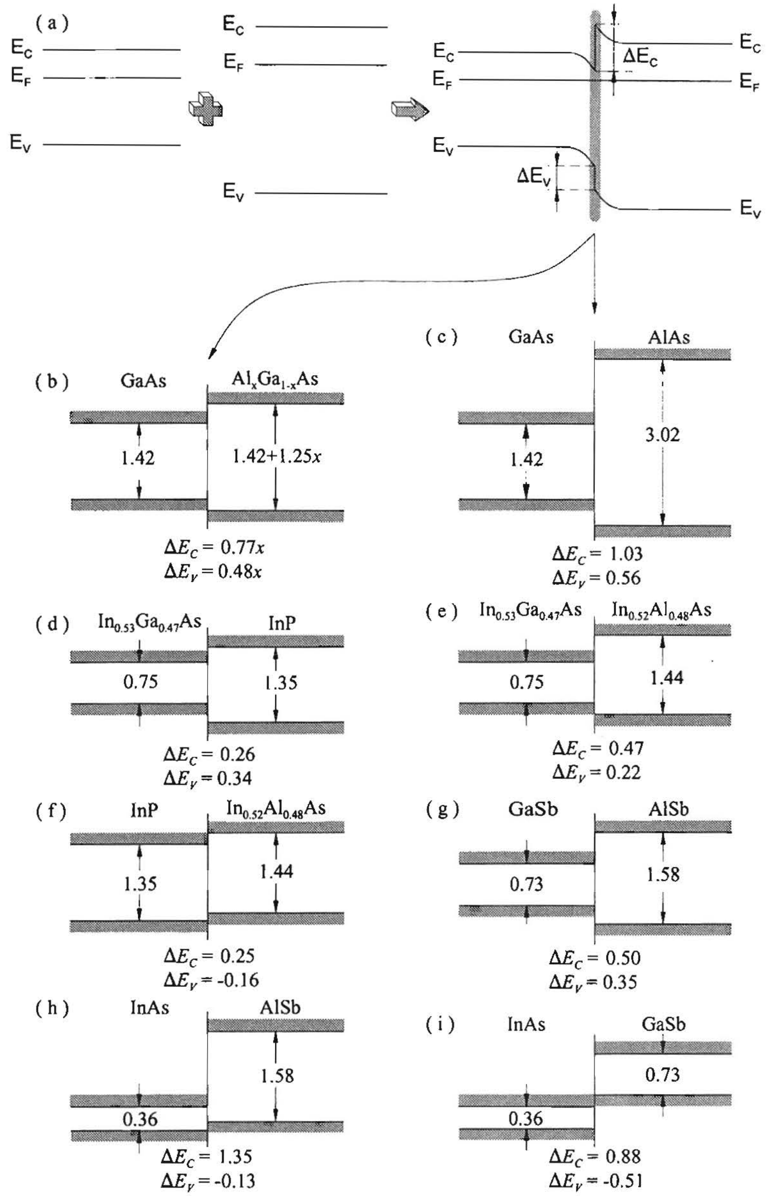


Fig. 1.1 Conduction- and valence- band offsets in a semiconductor heterojunction (a) and experimentally determined band alignments for lattice-matched heterojunctins between group III-V semiconductors (b-i) [27]. Energies are indicated in electron volts. For GaAs/$\mathrm{Al_xGa_{1-x}As}$ the case (b) illustrates the direct gap range and (c) the indirect gap corresponding to Γ minimum. Cases (b), (c), (d), (e), and (g) show straddling alignments, cases (f) and (h) are typical for staggered alignments, and case (i) corresponds to the broken-gap alignment.

1.4 Evolution of the Two and Three Terminal Resonant Tunneling Devices

The history of the double barrier resonant tunneling (DBRT) diode or two terminal (2T) resonant tunneling device, identifies with the progress of the molecular beam epitaxy (MBE) technique. In 1974 Chang et al. [32] fabricated the first MBE grown DBRT structure that exhibited evidence of negative differential resistance (NDR) at low temperature. Following this pioneering result, most of the work in heterostructures centred around the development of the MBE, since it was clear that the quality of the materials in these structures would have to be improved in order to observe the NDR or any other quantum transport at room temperature. The primary structures studied during this time were single-heterojunction devices or quantum wells rather than resonant tunneling structures.

In 1983 Sollner et al. [33] used a cooled GaAs/AlGaAs double barrier structure to rectify 2.5 THz laser radiation. This pivotal experiment demonstrated that resonant tunneling was a very fast process and it spurred the development of improved materials. Within 3 years several groups have achieved a relative large room-temperature NDR effect, first in the GaAs/$Al_{0.25}Ga_{0.75}As$ material [34] and shortly after that in the GaAs/AlAs system [35, 36]. After these first successful examples an intense research activity developed to increase the performances of the resonant tunneling structures and at the same time to study the physical phenomenon itself. Before other MBE material systems became available for quantum devices the field of resonant tunneling structures was already matured using almost exclusively the GaAs/$Al_xGa_{1-x}As$ system.

Fig. 1.2 illustrates schematically a typical electron resonant tunneling structure, where a GaAs quantum well is clad by two thin barrier layers of a material with a larger gap (e.g. AlGaAs). The contact layers are doped n-type for an electron tunneling device. Two other kinds of layers are inserted between the proper structure and the contacts: the undoped "spacer" layers and the lowly doped layers.

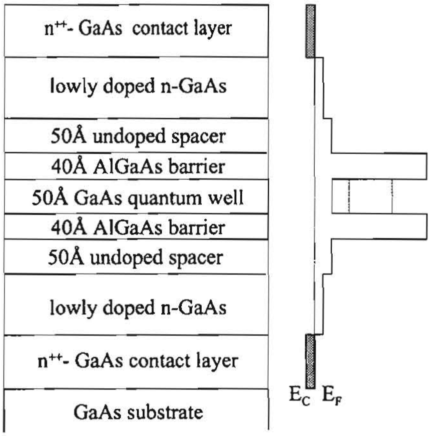


Fig. 1.2 Schematic illustration of the MBE-grown structure and the corresponding conduction band diagram for a typical n-type DBRT structure.

From the application point of view the quality of a DBRT diode is usually given by the peak current density J_p and the peak to valley current density ratio PVCR $\equiv J_p/J_v$, where J_p and J_v are the peak and valley current densities, respectively, associated with the NDR region of the current-voltage (I-V) characteristics (see Fig. 1.3). Other quantities of importance are the peak to valley voltage ratio PVVR $\equiv V_p/V_v$, where V_p and V_v are the peak and valley voltages, respectively, and the available current density $\Delta J \equiv J_p - J_v$. For high speed operation the current density should be large ($J_p \gtrsim 10^5$ A/cm^2) for the same reasons as in any other fast speed electronic device: high current density is required for fast charging and discharging of the device and circuit capacitance. ΔJ should be also as large as possible and $\Delta V = V_v - V_p$ small. The PVCR of a double barrier structure is sensitive to relative magnitudes of a number of resonant, nonresonant, and parasitic currents. Some of the major contributions to the overall carrier transport in a DBRT diode include (a) resonant tunneling, (b) tunneling through a higher resonance level, (c) thermionic emission, (d) field-assisted tunneling, (e) tunneling through evanescent states, (f) inelastic tunneling, and (g) surface leakage current [37]. These processes are illustrated in Fig. 1.4.

The process of achieving a high PVCR represented in fact an "exercise" in enhancing the resonant tunneling current density while minimizing all nonresonant and parasitic contributions to the valley current. A number of factors that modify both resonant and nonresonant currents are due solely to specific MBE growth conditions. They refer to the quality of the layers and especially to the sharpness of the interfaces expressed by the so called interface roughness. Theoretical studies [38,39] have shown that scattering from interface roughness of the quantum well and barriers broadens the resonance leading to the decrease of the peak current. Unfortunately for the DBRT diode there is no technique to measure quantitatively the degree of interface roughness and to correlate it with the observed transport characteristics of the device. The "standard" quantum well characterization technique - the observation of the excitonic spectrum using photoluminescence (PL) - is not possible in this case because the high doping of the contact layers effectively masks the PL emission from the energy levels in the quantum well [40]. In this way the most sensitive characterization tool for the analysis of the structure of materials used in resonant tunneling applications has appeared to be the tunneling device itself.

Besides the optimization of the growth conditions, DBRT devices with improved performances resulted by altering the design characteristics of the structure. Among them the increase of the barrier height proved to be very important, modifying the Al concentration towards larger values, and the optimization of their width. The introduction of undoped "spacer" layers on the outer sides of the barrier layers was also beneficial for the device quality. Finally, the same positive influence had the use of lowly doped layers, interposed between the contact and spacer layers.

The first DBRT structure functioning at room-temperature had a GaAs quantum well and $Al_{0.25}Ga_{0.75}As$ barriers without spacer and lowly doped layers [34].With a J_p of 1.6×10^3 A cm^{-2} and a PVCR of 1.5 (8) at 300 K (77 K) its performances were relative modest. Increasing the Al mole fraction in the barriers to $x \approx 0.4$ and introducing the spacer layers, useful PVCR at values of J_p up to about 1×10^4 A cm^{-2} resulted [41]. At higher values of J_p the PVCR is usually degraded by large thermionic current over the top of the barriers as well as the nonresonant tunneling current through the next quasilevel in the quantum well. Consequently, much better PVCR at high J_p offered the GaAs/AlAs structures [36]. In GaAs/AlAs DBRT diodes having J_p in the range from 10^4 to 10^5 A cm^{-2}, the best observed room-temperature PVCRs are about 4.

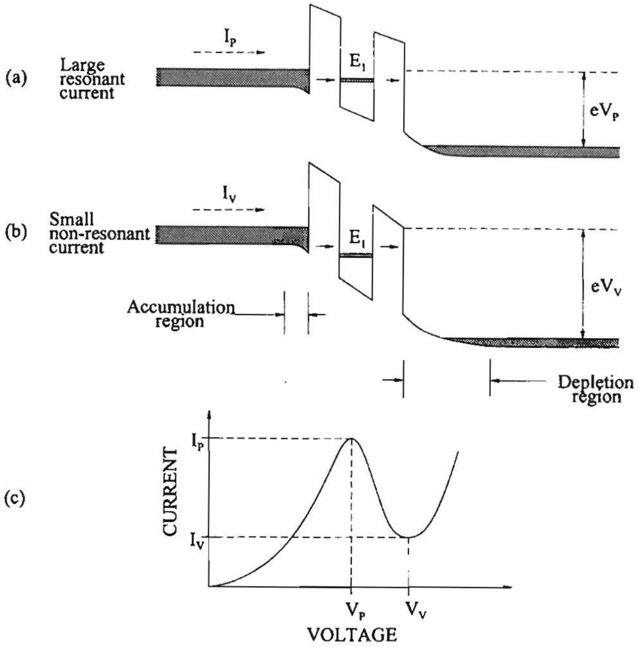


Fig. 1.3 Schematic diagrams of the current flow through a DBRT diode at bias voltage near peak (a) and valley (b) points, and representative I-V characteristics showing the NDR region (c).

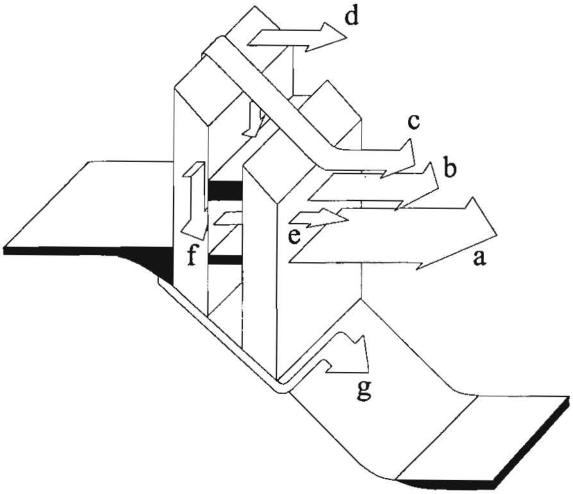


Fig. 1.4 Contributions to the resonant and nonresonant currents in a DBRT diode: (a) resonant tunneling, (b) tunneling through a higher resonance level, (c) thermionic emission, (d) field-assisted tunneling, (e) tunneling through evanescent states, (f) inelastic tunneling, and (g) surface leakage current [37].

These results followed after systematic studies of the effects of the barrier thickness and composition in Al [41, 42] as well as of the spacer layer width [43, 44]. At higher peak density currents, the PVCR degrades significantly [45, 46], the best reported PVCR with $J_p \gtrsim 10^5$ A cm^{-2} being 3.0 [47].

Several clever structures have been implemented to improve the PVCR of the GaAs DBRT diodes by lowering the valley current. In a first case, the spacer side AlAs barriers have been grown with $Al_xGa_{1-x}As$ layers just outside to form "chair barriers" [48].This has resulted in a DBRT diode having PVCR of 6.3 at room temperature [49]. Other structures contained an $In_xGa_{1-x}As$ layer either as a prewell on the cathode side [50] or in place of the GaAs quantum well [44]. An improved PVCR resulted in both cases.

New material systems successful grown by MBE have been used also for DBRT structures. The indium based systems, InGaAs/InAlAs and InAs/AlSb are the most important but others like InAs/GaSb/AlSb arouse more and more interest.

The first useful DBRT diodes in the In containing system have been made from $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ materials [51], because both of these ternary alloys are lattice matched to InP substrates. The structure yielded a PVCR of 2.3 at room temperature. This was subsequently improved to about 7 at room temperature and 40 at 77K [52, 53]. Superior results have been obtained by replacing the $In_{0.52}Al_{0.48}As$ barriers with pseudomorphic AlAs. The resulting DBRT diodes demonstrated an unprecedent PVCR of 14 at room temperature [54]. Further improvement followed with the achievement of a PVCR of 30 at room temperature in a DBRT structure having a thin InAs layer embedded in the $In_{0.53}Ga_{0.47}As$ quantum well [55].

The primary reason for the superiority of the $In_{0.53}Ga_{0.47}As$/AlAs diodes over the GaAs/AlAs diodes has been attributed to the large separation between the Γ point conduction band minimum in the $In_{0.53}Ga_{0.47}As$ and the X point conduction band point in the AlAs barrier [54]. In the GaAs/AlAs system this separation is much smaller and has been correlated with the contributions to the valley current in these diodes [56, 57].

High quality DBRT diodes have been fabricated from the InAs/AlSb material system. The system has a staggered band offset in which the valence band edge of the AlSb lines up in the band gap of the InAs. This determines a significantly smaller barrier attenuation coefficient than is realized in GaAs/AlAs or $In_{0.53}Ga_{0.47}As$/AlAs structures having the same barrier thickness. The values of $J_p \approx 3.7 \times 10^5$ A cm^{-2} and PVCR ≈ 3 obtained for InAs/AlSb structures [58] are comparable with the best results achieved in the $In_{0.53}Ga_{0.47}As$/AlAs system and are considerably better than those of GaAs/AlAs DBRT diodes. For comparison Fig. 1.5 shows I-V curves of high-J_p DBRT diodes made from different material systems [59].

From the beginning the long range application of the resonant tunneling structures was their implementation in three terminal (3T) devices. A particularly powerful concept for achieving this has been to incorporate a DBRT structure into one or another junction or transport region of a bipolar, hot electron, or field effect transistor. In this implementation the primary operational parameters are those of the "traditional" device, with the tunneling features added to alter these characteristics. The most successful solutions in this approach proved to be the resonant tunneling hot electron transistor (RHET) and the resonant tunneling heterostructure bipolar transistor (RBT), with applications in logic circuits. The band structures and operation principle of an RHET and an RBT are shown in Fig. 1.6 [60].

The RHET evolved from a transistor with a wide base and no confinement [61, 62] to a

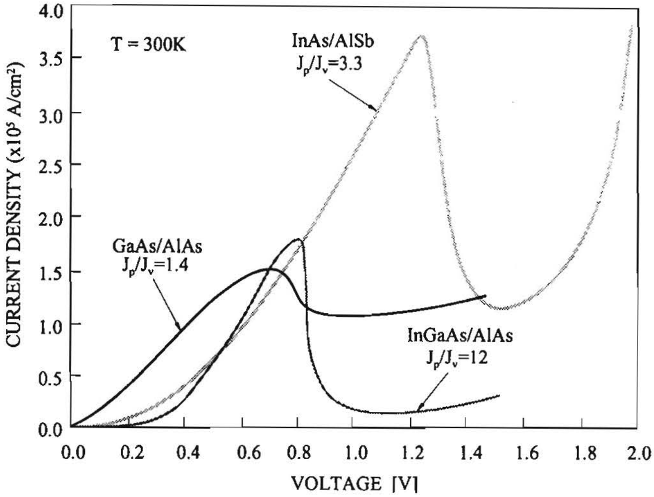


Fig. 1.5. Room temperature current density versus voltage curves for DBRT diodes made from three different material systems [60].

device with thin base (10nm) and hence confinement [63]. In this way a room temperature gain as high as 12 at a current density of 1.3 × 10^5 A cm^{-2} has been reached.

As an alternative at RHET several solutions of RBTs have been examined. The differences between them consists in the region of the bipolar transistor where the DBRT structure has been incorporated and the material system used. The best performances of such a transistor were obtained using InGaAs/InAlAs/InP system with the DBRT structure integrated in the emitter[64]: PVCR ≈ 70, $J_p > 10^4$ A cm^{-2} and the differential current gain at resonance of 19.

Another approach has been to make direct contact to the quantum well of a DBRT structure

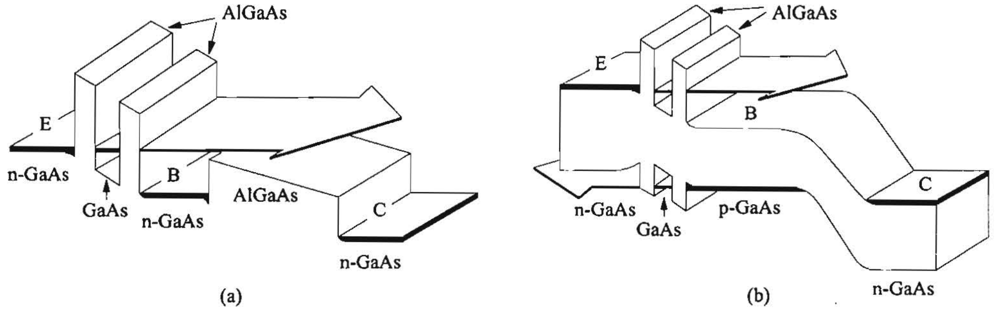


Fig. 1.6 Band Structure and principle of operation illustration of an RHET (a) and RBT (b) [61].

and thereby form a 3T device in which the quantum well base is two-dimensional (2D), with transport proceeding from three dimensional (3D) emitter to 2D base and then to 3D collector. In this way a separate control of the quantum well potential can be achieved. The primary challenge in this approach consists in the difficulty of exposing and making contact to such a thin layer (~ 5.0 nm) as the quantum well base. This technological obstacle conducted only to partial results in different laboratories [65-69].

After more than two decades, from the first DBRT structure, the resonant tunneling devices are a well defined and important part in the fast growing field of quantum devices.

1.5 Outline of the Dissertation

The goal of this dissertation is to study some physical aspects of 2T resonant tunneling devices and to propose a solution for a 3T resonant tunneling device with the base contact to the quantum well. To realize this, a comprehensive approach of the problems was developed, from the technological aspects of the device processing to the electrical measurements and then to modeling, with possible and inherent feed-back interconnections between them.

The second chapter gives an overview of the fundamental aspects of the resonant tunneling device physics. The actual understanding of the resonant tunneling process, with the two main approaches, coherent and incoherent tunneling, is exposed in detail, with the intention to give a unitary description of the phenomenon. The expression for the current density is derived and some important factors, which describe its behaviour with applied voltage, are discussed. The section offers the theoretical base for measurement interpretation and modeling presented in the next chapters.

The technological processes involved in the fabrication of 2T and 3T resonant tunneling devices are presented in detail within chapter 3. Special attention is paid to the photolithography using both direct and dual image resists, wet chemical etching, selective etching of GaAs and AlGaAs, and gold plating for airbridge interconnections, which have been optimized for the specific design characteristics of the devices. The strategy of the mask design is also described.

The next chapters present the devices studied in this thesis. The epitaxial structures necessary for device fabrication have been grown by MBE at the Solid State Physics department of Eindhoven University of Technology (EUT). Both 2T and 3T resonant tunneling devices described in the dissertation have been processed from structures using the available GaAs/AlGaAs material system. The processing has been accomplished using the clean room facilities of the Electronic Devices department of EUT. Chapter 4 deals with the 2T resonant tunneling devices. After the device fabrication description, the electrical measurements are presented and explained in detail, along several paragraphs. First, a comparative analysis of the room temperature DC results corresponding to the two types of processed devices, based on $GaAs/Al_{0.4}Ga_{0.6}As$ and GaAs/AlAs materials, is carried out. Then, DC measurements versus temperature are interpreted in connection with possible tunneling mechanisms. Finally, accurate microwave measurements performed over the whole bias range of the device I-V characteristic are used to establish the proper equivalent-circuit of our DBRT structures and to extract the bias behaviour of the parameters associated to the circuit elements. The last part of the chapter is devoted to the device physical modeling of the I-V characteristic. Using an existing model, developed within the Electronic Devices Group in the last periode [70, 71], the DC behaviour of the 2T DBRT devices is described both qualitatively and quantitatively. Chapter 5 describes the fabrication and characterisation of the 3T DBRT devices. First, the technological solution to realize the device is exposed, with all the difficulties generated by the effort to contact such a thin layer as the quantum well. The mask layout designed especially for this aim is also described. Finally the DC and AC measurement results, that demonstrate the successful contacting of the quantum well, are discussed in connection with possible applications of this new device.

The last part summarises the work and gives an outlook on subsequent developments in the field of resonant tunneling devices described in this dissertation.

References

1. J. R. Oppenheimer, *Phys. Rev.* **31**, 66 (1928).
2. R. H. Fowler and L. W. Nordheim, *Proc. Roy. Soc.* **A119**, 173 (1928).
3. L. W. Nordheim, *Phys. Z.* **30**, 177 (1929).
4. J. Frenkel, *Einführug in die Wellenmechanik* ,Springer, Berlin, 1929.
5. G. Gamow, *Z. Phys* **51**, 204 (1928).
6. R. W. Gurney and E. U. Condon, *Phys. Rev.* **33**, 127 (1929).
7. J. Frenkel, *Phys. Rev.* **36**, 1604 (1930).
8. W. Schottky, *Phys. Z.* **32**, 833 (1931).
9. J. Frenkel and A. Joffe, *Phys. Z. Sowjetunion* **1**, 60 (1932).
10. A. H. Wilson, *Proc. Roy. Soc.* **A137**, 487 (1932).
11. C. Zener, *Proc. Roy. Soc.* **A145**, 523 (1934).
12. L. Esaki, *Phys. Rev.* **109**, 603 (1957).
13. I. Giavaer, *Phys. Rev. Lett.* **5**, 147 (1960); **5**, 464 (1960).
14. B. D. Josephson, *Phys. Lett.* **1**, 251 (1962).
15. B. D. Josephson, *Advan.Phys.* **14**, 419 (1963)
16. C. B. Duke, "*Tunneling in Solids*" Academic Press, New York, 1969.
17. E. Burnstein and S. Lundqvist eds., "*Tunneling Phenomena in Solids*" Plenum Press, 1969.
18. J. Bardeen, *Phys. Rev.* **49**, 653 (1936).
19. A. J. Benett and C. B. Duke, *Phys. Rev.* **160**, 541 (1967).
20. A. J. Benett and C. B. Duke, *Phys. Rev.* **162**, 578 (1967).
21. D. J. Ben Daniel and C. B. Duke, *Phys. Rev.* **152**, 683 (1966).
22. D. Bohm, "*Quantum Theory*" , p. 283, Prentice-Hall, Engelwood Cliffs,N.Y., 1951.
23. E. O. Kane, "Basic Concepts of Tunneling", in "*Tunneling Phenomena in Solids*", eds. E. Burnstein and S. Lundqvist, Plenum Press, New York,1969.
24. E. Merzbacher, "*Quantum Mechanics*", p. 100, John Willey & Sons, New York, 1970.
25. W.R. Frensley, *Phys. Rev. B* **36**, 1570 (1987).
26. H. Kroemer, *Surface Sci.* **132**, 543 (1983).
27. E. T. Yu, J. O. McCaldin, and T. C. Mc.Gill, in "*Solid State Physics, Advance in Research and Applications*", eds. Ehrenreich and D. Turnbull, Academic Press, San Diego, CA,

Vol.46, p. 2, 1992.
28. J. R. Waldrop, E. A. Kraut, C. W. Farley, and R. W. Grant, *J. Appl. Phys.* **69**, 372 (1991).
29. U. Cebulla, G. Tränkle, U. Ziem, A. Forchel, G. Griffiths, H. Kroemer, and S. Subbanna, *Phys. Rev. B* **37**, 6278 (1988).
30. J. R. Waldrop, E. A. Kraut, C. W. Farley, and R. W. Grant, *J. Vac. Sci. Technol.* **B8**, 768 (1990).
31. A. Nakagawa, H. Kroemer, and J. H. English, *Appl. Phys. Lett.* **54**, 1893 (1989).
32. L. L. Chang, L. Esaki, and R. Tsu, *Appl. Phys. Lett.* **24**, 593 (1974).
33. T. C. L. G. Sollner, W. D. Goodhue, P. E. Tannenwald, C. D. Parker, and D. D. Peck, *Appl. Phys. Lett.* **43**, 588 (1983).
34. T. J. Shewchuk, P. C. Chapin, P. D. Coleman, W. Kopp, R. Fisher, and H. Morkoç, *Appl. Phys. Lett.* **46**, 508 (1985).
35. M. Tsuchiya, H. Sakaki, and J. Yoshimo, *Jpn. J. Appl. Phys.* **24**, L466 (1985).
36. W. D. Goodhue, T. C. L. G. Sollner, H. Q. Lee, E. R. Brown, and B. A. Vojak, *Appl. Phys. Lett.* **49**, 1086 (1986).
37. R. D. Schnell, H. Tews, R. Neumann, A. Mitwalsky, R. Treichler, and G. Packeiser *Inst. Phys. Conf. Ser.* **106**, 825, (1990).
38. H. C. Liu and D. D. Coon, *J. Appl. Phys.* **64**, 6785 (1988).
39. H. C. Liu, *J. Appl. Phys.* **67**, 593 (1990).
40. R. L. Tober, J. Pamulapati, J. E. Oh, and P. K. Bhattacharya, *Appl. Phys. Lett.* **53**, 883 (1988).
41. 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, *Appl. Phys. Lett.* **51**, 121 (1987).
42. J. Söderström and T. G. Anderson, *Superlatt. Microstruct.* **5**, 109 (1989).
43. S. Muto, T. Inata, H. Onishi, N. Yokoyama, and S. Hiyamizu, *Jpn. J. Appl. Phys.* **25**, L577 (1986).
44. H. M. Yoo, S. M. Goodnick, and J. R. Arthur, *Appl. Phys. Lett.* **56**, 84 (1990).
45. E. R. Brown, T. C. L. G. Sollner, C. D. Parker, W. D. Goodhue, and C. L. Chen, *Appl. Phys. Lett.* **55**, 1777 (1989).
46. S. K. Diamond, E. Ozbay, M. J. W. Rodwell, D. M. Bloom, Y.C. Pao, E. Wolak, and J. Harris, Jr., *IEEE Eletron Device Lett.* **10**, 104 (1989).
47. E. Ozbay,and D.M. Bloom, *IEEE Eletron Device Lett.* **12**, 480 (1991).
48. P. Cheng and J. S. Harris, *Appl. Phys. Lett.* **56**, 1676 (1990).
49. V. K. Reddy, A. J. Tso, and D. P. Neikirk, *Electron. Lett.* **26**, 1742 (1990).
50. K. L. Lear, W. S. Lee, and J. S. Harris, *IEEE Trans. Electron Devices* **36**, 2619 (1989).
51. T. Inata, S. Muto, Y. Nakata, T. Fujii, H. Ohnishi, and S. Hiyamizu, *Jpn. J. Appl. Phys.* **15**, L983 (1986).
52. S. Sen, F. Capasso, A.L. Hutchinson, and A. Y. Cho, *Electron. Lett.* **23**, 1229 (1987).
53. A. A. Lakhani, R. C. Potter, D. Beyea, H. H. Hier, E. Hempfling, L. Aina, and J. M. O'Connor, *Electron. Lett.* **24**, 153 (1988).
54. T. Inata, S. Muto, Y. Nakata, T. Fujii, H. Ohnishi, and S. Hiyamizu, *Jpn. J. Appl. Phys.* **26**, L1332 (1987).
55. T. P. E. Broekaert, W. Lee, and C. G. Fonstand, *Appl. Phys. Lett.* **53**, 1545 (1988).
57. E. E. Mendez, W. I. Wang, E. Calleja, and C. E. T. Goncalves da Silva, *Appl. Phys. Lett.*

50, 1263 (1987).
57. A. R. Bonnefoi and T. C. McGill, *Phys. Rev. B* **37**, 8754 (1988).
58. J. Söderström, E. R. Brown, C. D. Parker, L. J. Mahoney, and T. C. McGill, *Appl. Phys. Lett.* **58**, 275 (1991).
59. E. R. Brown, "High-Speed Resonant-Tunneling Diodes", in *Heterostructures and Quantum Devices*, eds. N. G. Einspruch and W. R. Fresnley, Academic Press, San Diego, 1994.
60. D. Lippens, L. De Saint Pol, R. Bouregba, P. Mounaix, and T. Vinchon, *Revue Phys. Appl.* **24**, 17 (1989).
61. T. Mori, H. Ohnishi, K. Imamura, S. Muto, and N. Yokoyama, *Appl. Phys. Lett.* **49**, 1779 (1986).
62. N. Yokoyama, K. Imamura, H. Ohnishi, T. Mori, S. Muto, and A. Shibatomi, *Solid-State Electronics* **31**, 577 (1988).
63. A. C. Seabaugh, Y.-C. Kao, J. Randall, W. Frensley, and A. Khatibzadeh, *Jpn. J. Appl. Phys.* **30**, 921 (1991).
64. A. C. Seabaugh, E. A. Beam III, A. H. Taddiken, J. N.Randall, and Y.-C. Kao, *IEEE Electron Device Lett.* **14**, 472 (1993).
65. U. K. Reddy, I. Mehdi, R. K. Mains and G.I. Haddad, *Solid-State Electronics* **32,** 1377 (1989).
66. G. I. Haddad, R. K. Mains, U. K. Reddy, and J. R. East, *Superlatt. Microstruct.* **5**, 437 (1989).
67. M. A. Reed, W. R. Frensley, R. J. Matyi, J. N. Randall, and A. C. Seabaugh, *Appl. Phys. Lett.* **54**, 1034 (1989).
68. A. C. Seabaugh, W. R. Frensley, J. N. Randall, M. A. Reed, D. L. Farrington, and R. J. Matyi, *IEEE Trans. Electron Devices* **36**, 2328 (1989).
69. J. Genoe, *Ph. D. Thesis*, Leuven, 1994.
70. Th. G. van de Roer, "*Modeling of Double Barrier Resonant Tunneling Diodes: D.C. and Noise Model*", EUT Report 95-E-285, ISBN 90-6144-285-0, 1995.
71. A. M. P. J. Hendriks, W. Magnus, and T. G. van de Roer, *Solid-State Electronics* **39**, 703 (1996).

2

Physics of Resonant Tunneling

2.1 Introduction

In general, the specific properties of any electronic device are determined by one or several (not many) basic physical phenomena. As seen in the previous chapter the DBRT device characteristics originate in a quantum phenomenon, namely resonant tunneling. Consequently, the study and explanation of the behaviour of such devices need, at least for the active region, a quantum analysis. In this chapter main aspects of the DBRT device physics are introduced. The presentation is intended as a general picture of the actual understanding of the resonant tunneling phenomenon used in the context of a real device, which may constitute a theoretical base for the measurement interpretation and modeling from the next chapters. In this sense, first the physical principles of the DBRT are discussed from a point of view used in quantum mechanics to describe the resonant scattering phenomena [1]. This approach, usually named the Breit-Wigner formulation of DBRT, situates this phenomenon, specific for semiconductor heterostructure, in a general context and correlates the transmission probabilities with measurable physical quantities, namely the energy and life-time of the quasi-levels inside the structure and scattering times. Then, the most accessible quantity for the DBRT characterization, the current density, is introduced. At last, factors, such as band bending under bias conditions, quantum well charge storage, band nonparabolicity, scattering mechanisms, which influence the behaviour of the current density in a real DBRT device are presented and discussed.

2.2 Principles of Resonant Tunneling

The DBRT devices operate on physical principles that are shared by all resonant-tunneling structures. As we know the process of resonant tunneling implies the existence of metastable states in the potential barrier region. Thus, one can say that the fundamental requirement for this process to occur is quantization by spatial confinement. For the DBRT structures the presence of the quantum well between the two barriers ensures this condition. Each quasi-stationary state from the quantum well is characterized by an energy E_n and a lifetime τ_n [1]. In the real case, when different scattering processes are involved, the condition for spatial quantization is

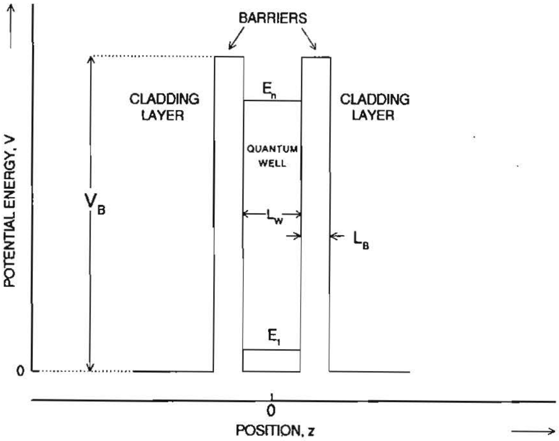


Fig. 2.1 Schematic diagram of the double barrier structure

E_{n+1}-E_n » $\hbar/\tau_s$ where τ_s is the scattering time for an electron in the two dimensional subband associated with the n^{th} quasi-level. This is similar to the condition that applies to other types of quantization in solids. For example, magnetic quantization requires $\omega_c\tau_s$ » 1, where ω_c is the cyclotron angular frequency.

The spatial quantization ensures that the quasi-two dimensional states exist in the quantum well, but it does not specify the nature of the resonant tunneling that occurs through these states. The nature of the resonant tunneling is usually defined by the relative magnitudes of τ_n and τ_s [2, 3]. If τ_n « τ_s the resonant tunneling through the n^{th} state is said coherent, since under this condition the electron does not scatter during the process and thus the phase of the wave function at each point in space is continuous in time. If τ_n » τ_s, the resonant tunneling is said to be incoherent, or sequential. In this case the wave function phase is partially randomized by the scattering events that occur in the double barrier structure.

The basic principles of resonant tunneling are readily illustrated if one neglects both the complicating effects of compositional grading in the heterojunction and the electrostatic band bending. This means that we assume that the heterojunctions of the double barrier structure are perfectly abrupt and no externally voltage is applied. Such a situation is shown in Fig. 2.1. The potential energy difference, or band offset V_B, at the interfaces is some fraction of the difference in the bandgaps between the two materials. For thin barriers the band gaps for both quantum well and barrier materials are usually taken at the minimum Γ-point of the Brillouin zone [4].

2.2.1 Electronic States in a Double Barrier Structure

The energies corresponding to the quasi-stationary states of the DBRT structure shown in Fig. 2.1 can be found based on the description of similar kind of states given in [1]. Within this description an electron placed in the well, in a quasi-stationary state moves inside the well for a considerable period of time τ_n, called the lifetime of the almost stationary state. The energy of this state is not discrete but broadened, whose "width" is related to the lifetime by $\Gamma_n \sim \hbar/\tau_n$ (uncertainty principle for energy). In this conditions we have to look for solutions of the Schrödinger equation which represent outgoing waves outside the barrier region (at infinity). This corresponds to an electron finally leaving the quantum well by tunneling. Solving Schrödinger equation with such boundary conditions, a set of complex values of the energy are obtained, which can be written in the form $\overline{E_n} = E_n - i\Gamma_n / 2$.The physical significance of the complex energy values is in connection with the limited time spent by the electron inside the quantum well during the tunneling process. If we write the time factor of the wave function of the quasi-stationary state in the form $e^{-(i/\hbar)\overline{E_n}t} = e^{-(i/\hbar)E_n t} e^{-[\Gamma_n/(2\hbar)]t}$, then the probability of finding the electron "inside the system" (quantum well) decreases with time as $e^{-(\Gamma_n/\hbar)t}$. Thus Γ_n determines the lifetime τ_n of the state.

In our case, we consider that the electron wave functions are solutions of the one-dimensional effective-mass Schrödinger equation [5, 6]:

$$\left[-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m^*(z)}\frac{\partial}{\partial z} + V(z)\right]\psi(z) = E\psi(z) \tag{2.1}$$

where $\psi(z)$ is the envelope function, $m^*(z)$ is the effective mass of the electron and $V(z)$ is the potential energy at the conduction band minimum. As shown in Fig. 2.1, the origin of the coordinates $z = 0$ is taken as the centre of the quantum well and the structure is symmetric, $V(z) = V(-z)$ and $m^*(z) = m^*(-z)$. We can restrict ourselves to the half-space $z > 0$ and consider for the quantum well the even eigenfunction. In this situation the wave functions for the different regions of the structure can be written as:

$$\psi(z) = \begin{cases} A\cos(\bar{k}z) & 0 < z < \dfrac{L_W}{2} \\ B\exp(\bar{\alpha}z) + C\exp(-\bar{\alpha}z) & \dfrac{L_W}{2} < z < \dfrac{L_W}{2} + L_B \,, \\ D\exp(i\bar{k}z) & z > \dfrac{L_W}{2} + L_B \end{cases} \tag{2.2}$$

where

$$\bar{k} = \left(\frac{2m_W^*}{\hbar^2}\overline{E}\right)^{1/2},$$

and

$$\bar{\alpha} = \left[\frac{2m_B^*}{\hbar^2}\left(V_B - \overline{E}\right)\right]^{1/2},$$

are the wave vector in the quantum well and cladding layer regions and the barrier attenuation coefficient, respectively, m_W^* and m_B^* denote the corresponding electron effective masses, and L_W and L_B represent the quantum well and barrier widths, respectively. The energy $\overline{E}$ as well as the wave vector $\bar{k}$ and the attenuation coefficient $\bar{\alpha}$ are complex.

In the effective-mass approximation the boundary conditions at the interfaces between different regions are: the continuity of $\psi(z)$ and $(1/m^*)d\psi/dz$ [5]. These connection rules express the conservation of the probability density and current. Using the boudary conditions at $z = L_W/2$ and $z = L_W/2 + L_B$ we obtain a system of four homogenous equations for the constants A, B, C and D. These equations will have a nontrivial solution if we require the determinant of the coefficient matrix to vanish. This condition gives the following equation for the allowed wave vectors:

$$\left[\tan\left(\frac{\bar{k}L_W}{2}\right)+\bar{\gamma}\right][i+\bar{\gamma}]\exp(-2\bar{\alpha}L_B)=\left[\tan\left(\frac{\bar{k}L_W}{2}\right)-\bar{\gamma}\right][i-\bar{\gamma}], \qquad (2.3)$$

where $\bar{\gamma} = \frac{m_W^* \bar{\alpha}}{m_B^* \bar{k}}$.

We recognize in (2.3) a transcendental equation, whose general solution requires numerical techniques. The roots $\bar{k}_n$ have negative imaginary parts and the magnitudes of both real and imaginary parts increase with n. For each $\bar{k}_n$ one can calculate the corresponding complex energy $\bar{E}_n$ and extract from this the real energy, $E_n = (\hbar^2/2m_W^*)\mathrm{Re}(\bar{k}_n^2)$ and the width, $\Gamma_n = -(\hbar^2/m_W^*)\mathrm{Im}(\bar{k}_n^2)$ of the resonant level.

Fig. 2.2 shows the solution associated with the ground state for a number of GaAs-AlAs double barrier structures. It can be seen that for each case the real part E_1 of the solution approaches the asymptotic value E_1^∞ for $L_B \gg 1.5$ nm. The imaginary part Γ_1 is exponential over the same range of L_B. From both curves it is possible thus to establish the cross-over from "thin-barrier" to "thick-barrier" behaviour. In the "thick-barrier" limit , when $\alpha L_B \gg 1$, with $\alpha = \mathrm{Re}\,\bar{\alpha}$, Eq. (2.3) reduces to the familiar bound-state eigen-value equation for an isolated quantum well, given as example in any quantum mechanics textbook, $\tan(kL_W/2) = m_W^*\alpha/(m_B^*k)$, where $k = \left[(2m_W^*/\hbar^2)E\right]^{1/2}$ and $\alpha = \left[(2m_B^*/\hbar^2)(V_B - E)\right]^{1/2}$.

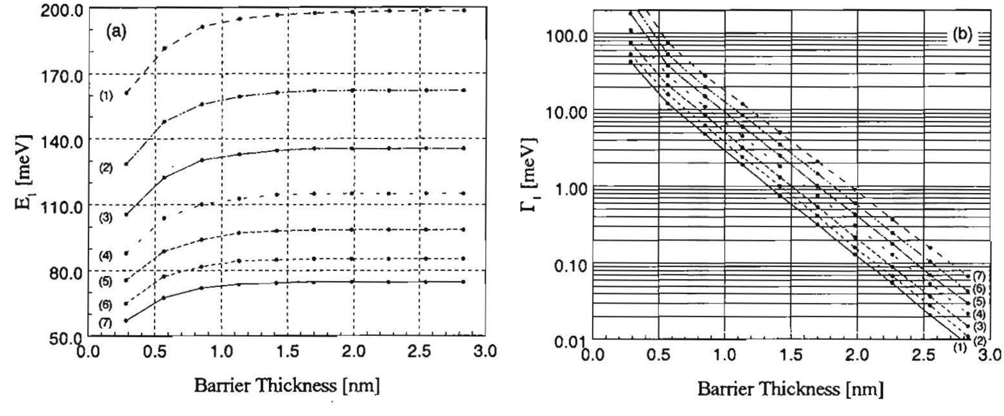


Fig. 2.2 (a) Quasi-bound-state energy E_1 (real part of the complex energy $\bar{E}_1$) in a GaAs quantum well as a function of the AlAs barrier thickness with quantum well width (integral number of monolayers ML of GaAs lattice at room temperature, ML = 0.565/2 nm) as a parameter: (1) L_W = 12 ML, (2) L_W = 14 ML, (3) L_W = 16 ML, (4) L_W = 18 ML, (5) L_W = 20 ML, (6) L_W = 22 ML, (7) L_W = 24 ML. Each point on a curve represents an integral number of ML of AlAs lattice at room temperature (ML = 0.566/2 nm). (b) Quasibound state width Γ_1 (imaginary part of the complex energy $\bar{E}_1$) for the same parameters as in (a) [38].

2.2.2 Coherent Resonant Tunneling

The interaction of the electron with the double barrier structure during the resonant tunneling process is described entirely by the transmission probability $T(E)$. This quantity is the ratio of the outgoing probability current density on one side of the structure to the incoming probability current density on the opposite side. Within the effective-mass model $T(E)$ is found assuming the same trial envelope functions as in the case of the quasi-bound states except that one cladding layer has an inward-going wave in addition to the outward wave. The most handy way to solve for $T(E)$, under these conditions, is by considering the transmission (T) matrix approach. Appropriate for multiple heterostructures the procedure consists, in principal, in the following: for each interface of the structure, using the specific boundary conditions, the amplitudes of the waves on both sides are related to each other by a T-matrix; after this the T-matrix of the whole structure, defined in the same way, is obtained as a product of the previous elementary matrices; finally the transmission and reflection probabilities can be calculated using the elements of the total T-matrix. In the Appendix A, section A.1, $T(E)$ is calculated in this way for a general double barrier structure, with the assumption of a piecewise constant potential. The result is:

$$T = \frac{T_{B_1} T_{B_2}}{1 + R_{B_1} R_{B_2} - 2\sqrt{R_{B_1} R_{B_2}} \cos\phi} \tag{2.4}$$

where T_{B_1,B_2} and R_{B_1,B_2} are the transmission and reflection probabilities, respectively, of the two barriers, and ϕ is a phase angle.

The transmission probability has the same form as the optical transmission of a Fabry-Perot resonator, with sharp peaks at longitudinal energies corresponding to each quasi-bound level in the quantum well. Physically the behaviour of $T(E)$ is the result of interference of the coherent electron wave function with itself, in the quantum well, by way of multiple reflections off the two barriers. For off-resonance energies the interference is destructive and the $T(E)$ is very small, but for energies equal to the quantum well quasi-levels the interference is constructive resulting in the cancellation of the reflected wave outside the structure and enhancing of the transmitted one.

During resonant tunneling through a double barrier structure the electrons reside in one of the decaying quasi-stable states of the quantum well. This suggests that the process also can be described in the frame of the quantum scattering theory [1]. In such an approach at energies near the resonances the transmission probability $T(E)$ can be expressed in the Breit-Wigner form permitting a direct correlation of the life time of the quasi-bound states, τ_n, with the full energetic width at half maximum of the transmission probability, Γ_n ($\tau_n = \hbar/\Gamma_n$).

It can be shown that the transmission probability obtained using T-matrix formalism is reduced to the Breit-Wigner form at energies near the resonances [7]. This derivation offers an interesting insight into the physics of resonant tunneling and confirms the success of a Breit-Wigner formulation of the process [8]. Following [7], we establish this correspondence for the simple case of the symmetric double barrier structure (see Fig. 2.1).

The transmission probability $T(E)$ is proportional with the square modulus of the transmission coefficient $t(E)$ (see Appendix A, section A.1), which represents the ratio between the amplitude of the transmitted and incident wave. The transmission coefficient for a double barrier structure can be written as (see Appendix A, section A.1):

$$t(E)=\left[M_{11}^{B_1}M_{11}^{B_2}+M_{12}^{B_1}M_{21}^{B_2}\right]^{-1}, \tag{2.5}$$

where $M_{11}^{B_1,B_2}$, $M_{12}^{B_1}$ and $M_{21}^{B_2}$ are elements of the T-matrices corresponding to the two barriers. Making explicit the matrix elements, the above expression can be worked out and the transmission coeficient put in the following form (see Appendix A, sectionA.2):

$$t(E)=\left\{Q\frac{1}{\gamma}(1+i\gamma)(1-\frac{i}{\gamma})\left[L+2\frac{1-i\gamma}{1+i\gamma}P\exp(-2\alpha L_B)+\left(\frac{1-i\gamma}{1+i\gamma}\right)^2 N\exp(-4\alpha L_B)\right]\right\}^{-1}, \tag{2.6}$$

where

$$\begin{aligned}
Q&=\frac{\exp[ik(2\alpha L_B+L_W)]\exp(2\alpha L_B)}{8},\\
L&=(\gamma^2-1)\sin(kL_W)+2\gamma\cos(kL_W),\\
P&=(\gamma^2+1)\sin(kL_W),\\
N&=(\gamma^2-1)\sin(kL_W)-2\gamma\cos(kL_W),
\end{aligned} \tag{2.7}$$

and $\gamma=\frac{m_W^*\alpha}{m_B^*k}$; α, k, L_W have the same meaning as in the previous section. We see that each $E=\overline{E_n^0}$ where the part in square brackets of Eq. (2.6) vanishes is a pole of $t(E)$. Because this expression contains imaginary terms, the poles E_n^0 of $t(E)$ generally are complex, i.e. lie off the real axis in the complex E-plane. From the foregoing section we know that the energy values corresponding to quasi-stationary states in the quantum well are also complex, so we can associate them with the complex poles E_n^0 of $t(E)$ and write:

$$\overline{E_n^0}=E_n-i\frac{\Gamma_n}{2} \tag{2.8}$$

where again E_n and Γ_n are positive.

The behaviour of t(E) at real incident energies E in the vicinity of $\overline{E_n^0}$ describes quantitatively the resonant tunneling via the quasi-stable states associated with E_n^0. This can be conveniently determined from the behaviour of the right side of Eq.(2.6) around its poles. As shown in the Appendix A, section A.2, such an analysis gives for the transmission coefficient, at real energies E near E_n, the expression:

$$t(E)=\mp i\exp[-ik(2L_B+L_W)]\exp(-i\eta)\frac{\frac{1}{2}\Gamma_n}{E-E_n+i\frac{\Gamma_n}{2}} \tag{2.9}$$

where η =arctgγ. Further on , the transmission probability is obained as:

$$T(E)=|t(E)|^2=\frac{\left(\frac{1}{2}\Gamma_n\right)^2}{(E-E_n)^2+\left(\frac{\Gamma_n}{2}\right)^2}. \tag{2.10}$$

Eq. (2.10) shows that the transmission probability through the double barrier structure has the accepted Breit-Wigner form [9] at energies near those corresponding to resonant tunneling. Also we can clearly see that Γ_n represents in fact the full width at half maximum of the transmission probability. It can be calculated using the folowing expression (see Appendix A, section A.2):

$$\Gamma_n = \left. \frac{16\hbar^2 k\gamma^2 \exp(-2\alpha L_B)}{m_W(1+\gamma^2)\left[(1+\gamma^2)L_W + \left(1+\frac{m_B}{m_W}\gamma^2\right)2\alpha^{-1}\right]} \right|_{E=E_n} . \qquad (2.11)$$

From the equivalence between the width of the quasi-bound states and the full width at half maximum of the corresponding transmission probability results the direct correlation of the life time of the quasi-bound states, τ_n, with the latter, $\tau_n = \hbar/\Gamma_n$.

For a general double barrier structure as that described in the Appendix A, section A1 the transmission probability can be also expressed, near resonances, in a Breit-Wigner form [7]:

$$T = \frac{\Gamma_n^{B_1}\Gamma_n^{B_2}}{(E-E_n)^2 + \frac{\Gamma_n^2}{4}}, \qquad (2.12)$$

where $\Gamma_n = \Gamma_n^{B_1} + \Gamma_n^{B_2}$, and $\Gamma_n^{B_1}$, $\Gamma_n^{B_2}$ are partial widths or transparencies associated with the transmission properties of the left and right barriers, respectively. For the case of constant effective mass along the double barrier structure the partial widths are estimated using the expression [7]:

$$\Gamma_n^{B_1,B_2} = \left[8\frac{\hbar^2 k_W^2}{m^*} \frac{k_{L,R}\alpha^2_{B_1,B_2}}{(L_W + \alpha_{B_1}^{-1} + \alpha_{B_2}^{-1})} \frac{\exp(-2\alpha_{B_1,B_2}L_{B_1,B_2})}{K_{B_1,B_2}} \right]_{E=E_n}, \qquad (2.13)$$

where k_W is the wave vector in the quantum well, α_{B_1,B_2} are the attenuation coefficients in the barriers, $k_{L,R}$ are the wave vectors in the left and right cladding layers, $K_{B_1,B_2} = \sqrt{2m^* V_{B_1,B_2}} / \hbar$, and V_{B_1,B_2} are the heights of the two barriers. In the WKB approximation they take the simple form $\Gamma_n^{B_i} = \hbar \nu_n T_{B_i}$ [3], where T_{B_i} are the transmission probabilities through barriers and $\nu_n = 2L_W / \mathrm{v}_n$ with L_W the width of the quantum well and v_n the velocity of carrier in the well at the resonant energy E_n.

The Breit-Wigner form of the transmission probability can be used in the estimating of the traversal time τ_T for resonant tunneling. The expression for the traversal time of an electron wave packet through a barrier within flat-band conditions is given by [10]:

$$\tau_T(E^0) = \left(\frac{D}{\mathrm{v}_g} + \hbar\frac{d\varphi}{dE} \right)_{E^0}, \qquad (2.14)$$

where D is the width of the barrier, $\mathrm{v}_g = \hbar^{-1}\mathrm{d}E/\mathrm{d}k$ is the group velocity in the region surrounding the barrier, φ is the phase angle of the transmission coefficient, and $E^0 = (\hbar k^0)^2/2m^*$, where k^0 is the centre of the wave packet at the initial time. The first term in Eq. (2.14) is the transit time, and the second term is the quantum-mechanical phase time. This expression is derived under the assumption that the transmission coefficient varies only slightly over the range E that corresponds to the width of the wave packet in k space. In double barrier structures this assumption is not usually valid. The traversal time in these structures is the sum of Eq. (2.14) (with $D = 2L_B+L_W$) and correction terms that depend on the shape of the wave packet [11]. For a resonant wave packet (i.e., $E^0 = E_n$) the corrections vanish, yielding:

$$\tau_T^n = \left(\frac{2L_B + L_W}{\mathrm{v}_g} + \hbar \frac{\frac{\Gamma_n}{2}}{(E-E_n)^2 + \frac{\Gamma_n^2}{4}} \right)_{E^0=E_n} = \frac{2L_B + L_w}{\mathrm{v}_g} + \frac{2\hbar}{\Gamma_n}. \tag{2.15}$$

We observe that the phase-time component is just twice τ_n, as might be expected intuitively. As an electron traverses the double barrier structure, the probability density must first increase from zero to some maximum value and than decrease back to zero. Both the increase and the decrease occur with a characteristic time τ_n.

In double barrier structures having thin barriers, the phase time usually exceeds the transit time by a large factor. For example, a GaAs-AlAs double barrier structure having 6ML-thick barriers and 16ML-thick quantum well has $E_1 \approx 0.14$ eV, $\mathrm{v}_g = \hbar k_1/m^* \approx 8\mathrm{x}10^7$ cm s^{-1} and $\Gamma_1 \approx$ 1 meV. Using Eq. (2.15) these values lead to a semiclassical transit time of 10 fs and a phase time of 1.3 ps. This estimate is comparable to the traversal times determined more accurately by numerical studies of the transmission of Gaussian wave packets through double barrier structures [12, 13].

2.2.3 Incoherent Resonant Tunneling

As in any physical system, inelastic scattering that occurs in the double barrier structure changes the phase of the wave function and its coherence is lost. This should have an important effect on the transmission probability, since the sharp resonances that are characteristic for typical quasi-bound levels require a multiple-pass self-interference of the wave function in the quantum well. The most common models of the effect make use of the correspondence between the resonant tunneling in double barrier structures and the general theory of collisions already pointed out for the elastic case in the previous section [3, 14, 15].

We illustrate the influence of scattering on the transmission probability for the simple case of the symmetric double barrier structure. The presence of collision phenomena inside the quantum well can be taken into account assuming that the scattering potential has a small constant imaginary part $i\Gamma_s/2$ [1, 14, 16] which "absorbs" part of the incident electron flux and causes the breakdown of the unitarity of the total coherent probability (transmission plus reflexion). The "lost electrons" into the quantum well will tunnel out, forward or backward relative to the direction of incident flux, in a separate incoherent process. The probability of each of these events is correlated with the relative transparency of the corresponding barrier. For symmetric barriers these probabilities are equal.

In this picture the coherent transmission probability T_{coh} is expressed as [14, 15]:

$$T_{coh} = \frac{\left(\frac{1}{2}\Gamma_n\right)^2}{(E-E_n)^2 + \left(\frac{1}{2}\Gamma_T\right)^2}, \tag{2.16}$$

where $\Gamma_T = \Gamma_n + \Gamma_s$. We observe that this expression for T_{coh} is similar with Eq. (2.10) for the fully coherent case in which Γ_n in the denominator is replaced by the larger quantity Γ_T.

Alongside with the T_{coh} an incoherent transmission probability T_{incoh} can be calculated [14, 15]:

$$T_{incoh} = \frac{\frac{1}{4}\Gamma_n\Gamma_s}{(E-E_n)^2 + \left(\frac{1}{2}\Gamma_T\right)^2}. \tag{2.17}$$

Using Eqs. (2.16) and (2.17) the total transmission probability in the presence of scattering is given by:

$$T = T_{coh} + T_{incoh} = \frac{\frac{1}{4}\Gamma_n\Gamma_T}{(E-E_n)^2 + \left(\frac{1}{2}\Gamma_T\right)^2}. \tag{2.18}$$

In the case of a general double barrier structure the Breit-Wigner form of the total transmission probability, including different collisions, has the expression [3]:

$$T = \frac{\Gamma_n^{B_1}\Gamma_n^{B_2}}{\Gamma_n^{B_1}+\Gamma_n^{B_2}} \frac{\Gamma_T}{(E-E_n)^2 + \left(\frac{1}{2}\Gamma_T\right)^2}, \tag{2.19}$$

where $\Gamma_T = (\Gamma_n^{B_1} + \Gamma_n^{B_2}) + \Gamma_s$, with $\Gamma_n^{B_1}$ and $\Gamma_n^{B_2}$ being the partial width from the coherent resonant tunneling.

For any finite value of Γ_s, a given transmission resonance is broader than in the coherent limit, and the peak value of the transmission is less than unity. Γ_s can be related to the scattering time through the uncertainty principle, $\Gamma_s \approx \hbar/\tau_s$.

The Breit-Wigner approach does not apply to processes in which the scattering occurs in quanta of energy comparable to E_n in the quantum well. One of such processes is the scattering with longitudinal optical (LO) phonons. The LO phonons shift the elastic resonance and introduce inelastic transmission side bands separated by $\hbar\omega_{LO}$ [17,18]. The transmission peak is reduced compared to the coherent case but with no additional broadening.

In both the Breit-Wigner and LO-phonon-mediated models the integrated transmission probability does not depend on the scattering parameter, i.e., Γ_s and $\hbar\omega_{LO}$, respectively.

The study of the incoherent resonant tunneling necessitates the estimate of τ_s. Several authors have argued that τ_s has approximatively the same value as the momentum relaxation time in the lateral plane of the quantum well [2, 19]. Thus one can estimate τ_s in double-barrier structures from the lateral mobility in the channels of the modulation-doped field-effect transistors (MODFETs) whose channel widths are comparable to L_W. The room-temperature GaAs MODFET electron mobility is typically of 7000 $\mathrm{cm^{-2}V^{-1}s^{-1}}$, which corresponds to a momentum relaxation time of $\tau_s = m^*\mu/e = 0.3$ ps. Setting τ_l equal to this τ_s, one finds a coherent width $\Gamma_1 = 2.0$ meV. From Fig. 2.2(b) we see that for $L_B \lesssim 1.7$ nm (6ML) in the GaAs-AlAs double barrier structures the resonant-tunneling process may be primarily coherent at room temperature.

2.3 Current Density

The physical principles underlying the general shape of the I-V curve of a double barrier structure can be understood by referring to the diagram in Fig. 2.3, in which we use the simple picture introduced in [20]. One assumes that the structure has a quantum well containing only one quasi-bound state energy level E_1. If the total energy and momentum of the electrons in the plane of the heterojunctions are conserved, then only those electrons having E_z on the cathode side approximately equal to E_1 can traverse the structure with any significant probability. As the bias voltage is increased from zero, the quasi-bound level drops relative to the electron energy in the cathode side. The number of electrons in the cathode Fermi sphere that have $E_z = E_1$ increases and this will determine a corresponding increasing of the current. This eventually approaches a peak at a voltage close to that which aligns the resonant level with the conduction band edge in the cathode side. At higher voltages there are no electrons with $E_z = E_1$, so that the current decreases precipitously and a negative differential resistance (NDR) occurs in the I-V curve.

To calculate the current density one usually considers that on the cathode and anode sides of the double barrier structure there are two independent 'electron reservoirs', in the thermodynamic sense, within which the electrons are all in equilibrium with a constant temperature and Fermi

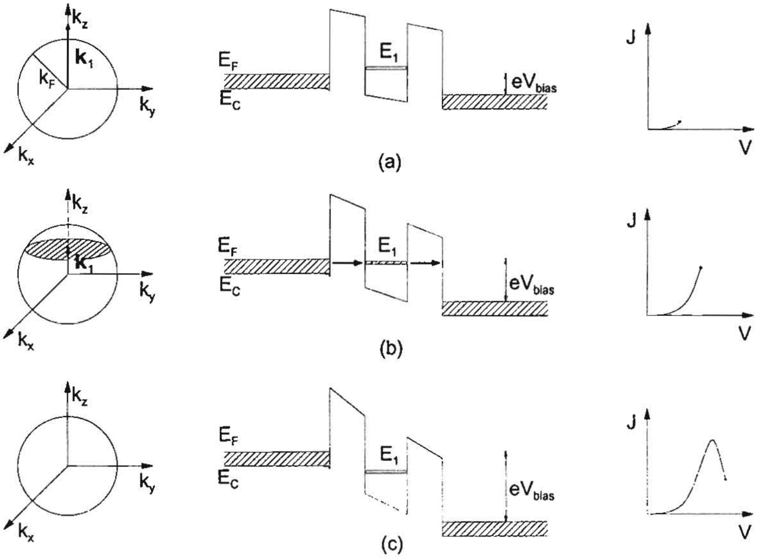


Fig. 2.3 Fermi sphere (FS), band-edge profile, and I-V curve for a double-barrier structure in three different bias conditions: (a) the momentum k_1 in the cathode side, corresponding to the quasi-level E_1 is situated outside the FS, only few electrons fulfiling the resonant condition; (b) the k_1 = constant plane intersects the FS, the area of the resulting disk being proportional with the states in the cathode side which can tunnel through E_1 to the anode; (c) the quasi-level is situated bellow the catode edge conduction band, no electrons are in the resonant condition.

level, and which are so large that the current flow into or out of the structure represents a negligible perturbation. The electrons inside the two reservoirs are treated in the independent-electron approximation. Also, the corresponding left and right wave functions do not overlap [21], so no terms which account for the non-orthogonality of these will appear in the expression of the density current, as proposed in [22]. Quantitatively, the net current through the double barrier structure is taken to be the difference between the current from the left to right and that from right to left. We measure the electron energy E from the bottom of the conduction band of the left-hand electrode. The energy conservation requires that an energy, E_r, measured to the bottom of the conduction band in the right-hand electrode be related to its counterpart $E_l \equiv E$ in the left hand electrode by $E = E + E_F^r - E_F^l + eV_{bias}$ with $E_F^{r,l}$ the Fermi levels in the reservoirs, e the electron charge and V_{bias} the applied bias. The electron effective mass in the electrodes is supposed to be the same, as usually happens in practice. Within this picture each of the two components of the current density is obtained by summing the expectation values of current operator impinging on the structure, multiplied by the corresponding transmission probability T_{tot}, over the occupied electronic states. Then the current density due to electron transport from left to right is given by

$$J^{\rightarrow} = \frac{2e\hbar}{(2\pi)^3 m^*} \int_0^{\infty} dk_z k_z T_{tot}(k_z) \int d^2\mathbf{k}_{\parallel} f_l(E)[1 - f_r(E)] \tag{2.20}$$

where 2 is the electron spinfactor, k_z and $\mathbf{k}_{\parallel}$ are the components of its momentum perpendicular and parallel, respectively to the plane of the heterojunctions and

$$\begin{gathered} f_l(E) = \frac{1}{1+\exp\left[(E - E_F^l)/k_B T\right]}, \\ f_r(E) = \frac{1}{1+\exp\left[(E_r - E_F^r)/k_B T\right]} = \frac{1}{1+\exp\left[(E + eV_{bias} - E_F^l)/k_B T\right]} = f_l(E + eV_{bias}), \end{gathered} \tag{2.21}$$

are the Fermi-Dirac distribution functions in the left and right reservoir, respectively, with k_B Boltzmann's constant and T the absolute temperature. The factor $(1\text{-}f_r)$ guarantees the unoccupied character of the final state of the electron. Similarly, the current density due to electron flow from right to left is given by

$$J^{\leftarrow} = \frac{2e\hbar}{(2\pi)^3 m^*} \int_0^{\infty} dk_z k_z T_{tot}(k_z) \int d^2\mathbf{k}_{\parallel} f_r(E)[1 - f_l(E)]. \tag{2.22}$$

In Eqs. (2.20) and (2.22) the same T_{tot} have been used because of the time reversal invariance of the one electron Schrödinger equation. Also this is dependent only on longitudinal energy E_z (see Appendix A, section A.1), which is a consequence of the independent-electron model used in our calculations. The expression for the net current density is obtained by subtracting Eq. (2.22) from Eq. (2.20):

$$J = J^{\rightarrow} - J^{\leftarrow} = \frac{2e\hbar}{(2\pi)^3 m^*} \int_0^{\infty} dk_z k_z T_{tot}(k_z) \int d^2\mathbf{k}_{\parallel} [f(E) - f(E + eV_{bias})], \tag{2.23}$$

where the 'l' index of the distribution functions have been dropped. Further on we convert the $d^2\mathbf{k}_{\parallel}$ integration into one over $dE_{\parallel}$,by noting that $d^2\mathbf{k}_{\parallel} = (2\pi m^*/\hbar^2)dE_{\parallel}$, and dk_z integration into one over E_z. Integrating over $dE_{\parallel}$ one obtains:

$$J = \frac{em^* k_B T}{2\pi^2 \hbar^3} \int_0^{\infty} dE_z T_{tot}(E_z) \ln\left\{ \frac{1 + \exp[1 + (E_F - E_z)/k_B T]}{1 + \exp[1 + (E_F - E_z - eV_{bias})/k_B T]} \right\}. \tag{2.24}$$

This is the general accepted expression for the static current density which can not be completely worked out into an analytical form and necessitates numerical integration.

2.4 Band Bending under Bias Conditions

An external bias voltage applied across a double barrier structure modifies the equilibrium band profile in the active region. As we have foreshadowed in the previous section, the analysis of the current density implies the determination of the modified band profile, or band bending, to a satisfactory degree of accuracy. The band bending affects the transmission probability in several ways. In a first approximation, it has a direct effect on the quasi-bound levels, decreasing their energy relative to the conduction band edge of the cathode side layer by an amount $1/2(eV_w)$, where V_w is the component of the total voltage that occurs across the double barrier structure, as shown in Fig. 2.4. Physically, this shift represents the decrease in the potential energy at the centre of the quantum well and is unrelated to any Stark effect. We have seen that it is this effect which gives the qualitative explanation of the I-V characteristic. The second effect of the band profile modification is a change in the peak transmission. This can be understood using Eqs. (2.12) and (2.13). Thus the peak of the first resonance of a symmetric double barrier structure under bias will fall below 1.0 because $\Gamma_n^{B_2}$ is larger than $\Gamma_n^{B_1}$. The increase in $\Gamma_n^{B_2}$ reflects an increase in k_R and a decrease in $\alpha_n^{B_2}$, while the decrease in $\Gamma_n^{B_1}$ reflects a decrease in k_L and an increase in $\alpha_n^{B_1}$.

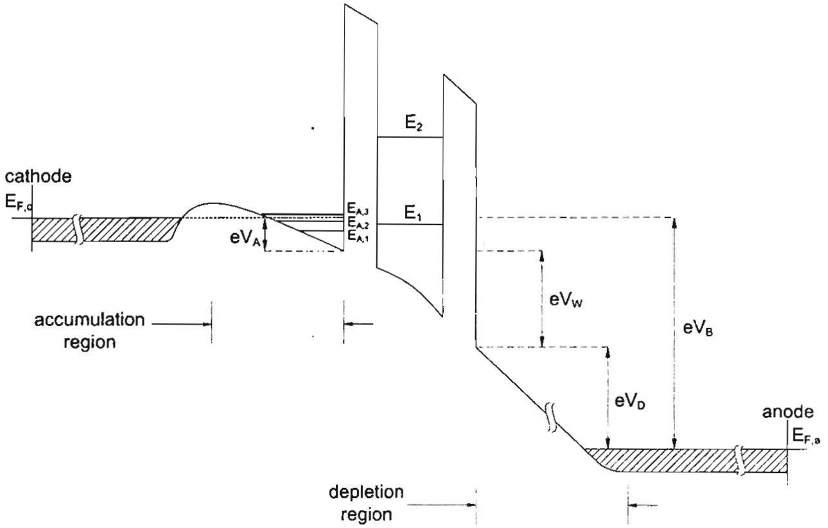


Fig. 2.4 Conduction band bending diagram of a DBRT diode under the bias voltage.

A simple way to estimate the band bending is by treating the double barrier structure as an insulator, modeling the device as a semiconductor-insulator-semiconductor (SIS) diode. Unless the current density is very high, one can then assume that the diode is in a condition of quasi-equilibrium. Under this condition the electron concentration on each side of the structure is characterized by a uniform quasi Fermi level. The band bending is obtained by solving Poisson's equation on each side and connecting the solutions with a uniform electric field across the double barrier structure.

The real DBRT diodes have always interposed between the active region and the high doped contact layers some low doped layers and undoped spacer layers (see Fig. 1.1). Thus, as shown in Fig. 2.4, a positively charged depletion layer forms in the anode side and a negatively charged accumulation layer on the cathode side. The overall band bending is then complicated by the fact that a significant amount of electrons can occupy the quantum well and that both the quantum well charge storage and some fraction of the charge in the accumulation layer resides in quasi-two-dimensional states. To deal properly with these effects one must carry out self consistent calculations of Poisson and Schrödinger equations [23-25, chapter 4, this work]. Such calculations indicate that the charge build up in the quantum well can have a profound effect on the I-V curve. This is addressed in the next section. Also they show that the electron states in the accumulation layer consist of a continuum at energies above the neutral band edge in the cathode side and quasibound states at lower energies [26, 27, chapter 4, this work]. The presence of the spatial quantization in the accumulation layer has been confirmed experimentally in some DBRT devices by the presence of undulations in the I-V curve at voltages below the first peak [26, 28, chapter 4, this work].

2.5 Quantum Well Charge Storage: Intrinsic Bistability

The double barrier structure, like any quantum well system, can store mobile sheet charge during operation. Intuitively, one expects that the sheet charge density σ_W is proportional to the product of the current density and the quasi state life time, i.e., $\sigma_W = J\tau$. Indeed, detailed treatements of this effect in the sequential [29] and coherent [30] approaches have resulted in the following expression for σ_W:

$$\sigma_W^n = \hbar \frac{J}{\Gamma_n^{B_2}}. \tag{2.25}$$

The quantum well sheet charge alters the electrostatic potential across the double barrier structure from its otherwise linear form. The modified potential affects the resonant tunneling current because of the strong dependence of this current upon the position of the quasi bound level relative to the emitter conduction band edge. But, according to the Eq. (2.25), the sheet charge density is proportional to the current density so that the net current density becomes self dependent. This mechanism is known as *electrostatic feedback* [31, 32]. The electrostatic feedback of the space charge dynamically stored in the quantum well determines the phenomenon of *intrinsic bistability* which means that for some range of bias voltage the current is multiple valued [31]. This effect has not been observed unambiguously in any double barrier structure containing symmetric barriers. Instead, such structures typically show *extrinsic bistability* due to the rectification by these devices of their own electrical oscillations in the measurement circuit [33]. Intrinsic bistability has been observed clearly only in a highly asymmetric structure with the more transparent barrier adjacent to the cathode [32, 34]. One possible reason for the absence of experimentally observed intrinsic bistability in symmetric structures is the broadening of the transmission probability by scattering. Another reason may be the quantum size effects in the accumulation layer. Self consistent calculations of I-V characteristic for this kind of structures do not display intrinsic bistability if spatial quantization is accounted for in the accumulation layer [24, 25].

2.6 Effects of the Band Nonparabolicity

In the effective mass model, Eq. (2.1), an implicit assumption is that the electron energy lies close to the corresponding conduction band minimum, so that a parabolic energy dispersion applies to all regions of the double barrier structure. However in many resonant tunneling structures the electron energy is too far away from the local band edge for the parabolic approximation to be valid. For example, electrons that tunnel through an AlAs barrier from GaAs have an energy that is approximately 1 eV below the AlAs Γ-conduction band edge. At this point the tunneling electron wave vector is expected to deviate from the parabolic form $\alpha = \left[2m_B(V_B - E)\right]^{1/2} / \hbar$, where m_B^* is the Γ-valley effective mass in the barrier region. In addition, the ground state level in the GaAs quantum wells of the most used structures is situated at least 100 meV above the conduction band edge, so that the electronic dispersion relation deviates somewhat from $E = (\hbar k)^2 / 2m_W^*$. Both of these effects can be included in the effective mass formalism by invoking a multiband model in which the tunneling electron envelope function is a mixture of conduction and valence band envelope functions [35]. Thus for the barrier regions where the most significant deviation from parabolicity occurs, complex band-structure calculations [36], in a two-band model that connects the conduction band (effective mass m_C^*) with the light-hole band (effective mass m_V^*), can be used. An approximate dispersion relation for the tunneling electrons in these regions is given by [37]:

$$E = E_0 \pm \frac{\hbar^2 \alpha_{\max}}{m_{C,V}^*}(\alpha_{\max}^2 - \alpha^2), \tag{2.26}$$

where $E_0 = E_C - E_G / (1 + m_C^* / m_V^*)$is the energy in the band gap corresponding to the maximum of α,

$$\alpha_{\max} = \sqrt{\frac{E_G}{\hbar^2 (1 / m_C^* + 1 / m_V^*)}} \tag{2.27}$$

is the maximum value of α, and E_G is the energy gap. The + or - sign is chosen depending whether E is greater or less than E_0. They correspond to m_C^* and m_V^*, respectivley. Then to accommodate the nonparabolic dispersion relation with the effective mass envelope functions a so called local parabolic approximation [38] is introduced. In this approximation the nonparabolic wave vectors found with Eq. (2.26) are substituted where appropriate for the parabolic wave vectors and the effective mass in the barriers is approximated by

$$m_B^*(E) = \frac{\hbar^2 \alpha^2}{2|E - E_C|} \qquad \text{for} \quad E > E_0, \tag{2.28}$$

and

$$m_B^*(E) = \frac{\hbar^2 \alpha^2}{2|E_V - E|} \qquad \text{for} \quad E < E_0. \tag{2.29}$$

Nonparabolicity for the quantum well region can be treated using a three band model from **k·p** theory:

$$E = \frac{\hbar^2 k^2}{2m_W^*}(1 - ak^2) \tag{2.30}$$

where the parameter a can be calculated either from Kane's model [39] or by an empirical fit to the band structure.

2.7 Excess Valley Current

Excess valley current denotes the current measured through the DBRT diode, near the valley point and beyond, that is not explained by Eq. (2.24). It does not include the potentially large current components that traverse the double barrier structure through upper quasibound levels or thermionically over the top of the barriers. These components are properly predicted by Eq. (2.24) provided that an accurate form of the transmission probability is obtained at all longitudinal energies.

One of the most important sources of excess current in the double barrier structures is the transport associated with the upper valleys, particularly the X-valley in AlAs barriers. The X-valley related excess current in both single and double barrier structures has been studied theoretically by a number of techniques including a Fowler-Nordheim model [40], the transfer matrix technique [41], a pseudopotential method [42], and an empirical tight binding formalism [43]. Although there is no consensus, a commonly described transport mechanism is one whereby the incident Γ electron transfers to X conduction band at the first heterojunction, maintains an X character troughout the double barrier structure, and subsequently transfers from X conduction band back to Γ conduction band at the last heterojunction. Material systems having a large Γ-X offset, as InGaAs/AlAs, greatly suppress this mechanism by making the initial Γ-X transfer much less probable.

Another important source of excess current, inherent to all materials used for double barrier structures, is LO-phonon scattering. In the sequential picture, this form of scattering allows electrons to transfer from the cathode side into the quantum well at bias voltages past the point of alignment between the quasibound level and occupied states on the emitter side [44]. Once in the quantum well, the electrons can tunnel to the anode side by elastic means. In the coherent picture, as we pointed out, the LO-phonon scattering introduces side bands on the transmission probability curve [17]. In either picture, a large excess current arises in the vicinity of the valley voltage in most DBRT diodes. Materials with lower effective mass tend to have less of this scattering.

Acoustic phonon scattering and alloy scattering have been analyzed by the same sequential mechanism as used for the LO phonons [44]. The conclusion is that acoustic phonons play a much smaller role than the optical phonons in generating excess current. However, the analysis was not carried out for polar-mode acoustic phonon scattering, which is known to be very strong in narrow-band-gap, low effective mass materials. Alloy scattering in ternary compound barriers can also give rise to a significant excess current component that is nearly independent of temperature.

A final excess valley current mechanism inherent to all resonant tunneling diodes arises from the scattering in the cladding layers. This is, perhaps, the most fundamental of all scattering mechanisms, but one of the most difficult to deal with, since it transcends the stationary-state model used in general to model the transport in double barrier strucutures. The stationary-state model separates the quantum mechanical transport in the double barrier structure from the semiclassical transport in the cladding layers. A proper description of this mechanism requires the quantum kinetic theory [25, 45].

References

1. L. D. Landau and E. M. Lifshiz, "*Quantum Mechanics*", 3rd ed., section 134, Pergamon, New York, (1977).
2. P. J. Price, *Phys. Rev.* B **36**, 1314 (1987).
3. M. Büttiker, *IBM J. Res. Develop.* **32**, 63 (1988).
4. C. Mailhiot and T. C. McGill, *J. Vac. Sci. Technol.* B **1**, 439 (1983).
5. D. J. Ben Daniel and C. B. Duke, *Phys. Rev.* **152**, 683 (1966).
6. T. Ando, A. B. Fowler, and F. Stern, *Rev. Mod. Phys.* **54**, 437 (1982).
7. E. Gerjuoy and D. D. Coon, *Superlatt. and Microstruct.* **5**, 305 (1989).
8. D. D. Coon, K. M. S. V. Bandara and H. Zhao, *Appl. Phys. Lett.* **55,** 2453 (1989).
9. H. A. Bethe, "*Elementary Nuclear Theory*", p. 120, Wiley (1947).
10. E. H. Hauge and J. A. Støvneng, *Rev. Mod. Phys.* **61**, 917 (1989).
11. E. H. Hauge, J. P. Falck, and T. A. Fjeldy, *Phys. Rev.* B **36**, 4203 (1987).
12. H. Guo, K. Diff, G. Neofotistos, and J. D. Gunton, *Appl. Phys. Lett.* **53**, 131 (1988).
13. S. C. Kan and A. Yariv, *J. Appl. Phys.* **64**, 1196 (1985).
14. A. D. Stone and P. A. Lee, *Phys. Rev. Lett.* **54**, 1196 (1985).
15. M. Jonson and A.Grincwajg, *Appl. Phys. Lett.* **51**, 1729 (1987).
16. P. J. van Hall and J. H. Wolter, *Superlatt. and Microstruct.* **8**, 305 (1990).
17. N. S. Wingreen, K. W. Jacobsen, and J. W. Wilkins, *Phys. Rev. Lett.* **61**, 1396 (1988).
18. N. S. Wingreen, K. W. Jacobsen, and J. W. Wilkins, *Phys. Rev.* B **40**, 11834 (1989).
19. S. Luryi, *Superlatt. and Microstruct.* **5**, 375 (1989).
20. S. Luryi, *Appl. Phys. Lett.* **47**, 490 (1985).
21. H. J. M. F. Noteborn, "*Quantum Tunneling Transport of Electrons in Double-Barrier Heterostructures*", Ph.D. Thesis, Dept. of Technical Physics, Eindhoven University of Technology, 1993.
22. D. D. Coon and H. C. Liu, *Appl. Phys. Lett.* **47**, 172 (1985).

23. H. Ohnishi, T. Inata, S. Muto, N. Yokoyama, and A. Shibatomi, *Appl. Phys. Lett.* **49**,1248 (1986).
24. M. Cahay, M. McLennan, S. Datta, and M. S. Lundstrom, *Appl. Phys. Lett.* **50**, 612 (1987).
25. W. R. Frensley, *Phys. Rev.* B **36**, 1570 (1989).
26. E. T. Koenig, B. Jogai, M. J. Paulus, C. I. Huang, and C. A. Bozada, *J. Appl. Phys.* **68**, 3425 (1990).
27. T. Fiig and A. P. Jauho, *Appl. Phys. Lett.* **59**, 2245 (1991).
28. J. S. Wu, C. Y. Chang, C. P. Lee, K. H. Chang, D. G. Liu, and D. C. Liu, *Appl. Phys. Lett.* **57**, 2311 (1990).
29. F. W. Sheard and G. A. Toombs, *Appl. Phys. Lett.* **52**, 1228 (1988).
30. K. M. S. V. Bandara, D. D. Coon, and H. Zhao, *J. Appl. Phys.* **66**, 1227 (1989).
31. V. J. Goldman, D. C. Tsui, and J. E. Cunningham, *Phys. Rev. Lett.* **58**, 1256 (1987).
32. E. S. Alves, L. Eaves, M. Henini, O. H. Hughes, M. L. Leadbeater, F. W. Sheard, G. A. Toombs, G. Hill, and M. A. Pate, *Electron. Lett.* **24**, 1190 (1988)
33. T. C. L. G. Sollner, *Phys. Rev. Lett.* **59**, 1622 (1987).
34. A. V. Zaslavsky, V. J. Godman, D. C. Tsui, and J. E. Cunningham, *Appl. Phys. Lett.* **53**, 1408 (1988).
35. E. O. Kane, *J. Appl. Phys.* **32**, 83 (1961).
36. V. Heine, *Proc. Phys. Soc.* **81**, 300 (1963).
37. J. N. Schulman and T. C. McGill, *Phys. Rev.* B **23**, 4129 (1981).
38. E. Wolak, E. Özbay, B. G. Park, S. K. Diamond, D. M. Bloom, J. S. Harris, Jr., *J. Appl. Phys.* **69**, 3345 (1991).
39. T. Hiroshima and R. Lang, *Appl. Phys. Lett.* **49**, 456 (1988).
40. P. J. Price, *Surf. Sci.* **196**, 394 (1988).
41. D. Landheer, H. C. Liu, M. Buchaman, and R. Stoner, *Appl. Phys. Lett.* **54**,1784 (1989).
42. A. C. Marsh, *IEEE J. Quantum Electron.* **23**, 371 (1987).
43. K. V. Rousseau, K. L. Wang, and J. W. Schulman, *Appl. Phys. Lett.* **54**, 1341 (1989).
44. F. Chevoir and B. Vinter, *Appl. Phys. Lett.* **55**, 1859 (1989).
45. W. R. Frensley, "Quantum Transport", in *Heterostructures and Quantum Devices*, eds. N. G. Einspruch and W. R. Frensley, p. 273, Academic Press, London, 1994.

3

Technology

3.1 Introduction

Always, the design is the first step in the fabrication process of an electronic device. But to do this realistically, with maximum chance of success, it is necessary not only to know very well the physical phenomena involved in the functioning of the device but to have knowledge in very great detail about all the potential technological processes used later in the device making. This is more requested in the case of the quantum devices where the small layer thicknesses impose special demands from the very beginning when the structures are grown epitaxially. After the first runs are accomplished, feed back interconnections between the different activities involved in the device realization result in the optimization of this.

In this chapter we describe the technological processes that are necessary for the fabrication of the two and three terminal resonant tunnelling (RT) devices presented in the thesis. For this kind of devices and in general for all the quantum devices, developed from III-V compound semiconductors, one can discuss about two major technological stages in their fabrication. The first one refers to the epitaxial growth of the device structure using MBE or MOCVD systems. The second includes the processes which give "form" to the device and permit the electrical or optical access to it, both at the input and output. Each of these stages has its particular importance in the creation of high performance devices and on no account one can accord priority of one of them. We shall not discuss about the epitaxial growth. A short reference to the materials used for resonant tunnelling devices, whose choice was and is dictated by the progress registered in the growth techniques, has been done in the first chapter. The second stage is presented in detail, grouping the different processes into two classes. The classification is done in a natural way, considering that the "general processes" represent the indispensable steps in processing of any simplest RT device and the "specific processes" refer to the special requirements of a more sophisticated processing. To avoid loading the text with too many data and repetitions, all the information concerning the optimized processes, that we have used or developed during the preparation of the dissertation, have been grouped in the Appendix B, at the end of the thesis. This does not mean that they are less important, on the contrary, they represent in fact our work "fruits" in the field of technology for resonant tunnelling devices.

In the fabrication process of our resonant tunneling devices several processing steps present primary importance. These are photolithography, wet chemical etching, ohmic contact fabrication and selective etching. In this sense they have been treated in a greater extent. Photolithography precedes all the main steps of the device processing. In the presentation of this technological activity we discuss all the factors and conditions that finally lead to high quality definition of resist patterns, specific for the subsequent processes. In general all the mesa definitions necessary in the fabrication of our devices have been carried out using wet chemical etching. Based on the basic knowledge in this field we have investigated different etching systems aiming for solutions that ensure several important necessities in DBRT device processing: relative medium etch rate (~ 0.5 μm/min), good quality of the sample surface and pattern edges after etching, and small underetch along the main crystallographic directions of the GaAs substrate. Good ohmic contacts, low contact resistance and low penetration into the semiconductor material, are essential for an accurate DC and AC characterisation of the DBRT devices. This characterisation offers finally the real device performances necessary in the applications. In the corresponding paragraph we present the contact formation of our successful Ge/Ni/Au metallization scheme and the basic elements used to characterize the ohmic contacts. At last, the selective etching represents the key process in the fabrication of our 3T DBRT devices, this explaining our efforts in finding of a procedure with very high selectivity.

We are convinced that in the process of fabrication of high performance devices each of the processing steps has its importance, so we have endeavoured the optimization of all of them in the specific conditions of our laboratory. In this sense we have to stress that the "recipe" of a certain process works different in different laboratories (slightly different cleanliness conditions, processing substances, etc.), so the best thing is to try to understand as much as possible the mechanism of the process and its direction of evolution. In this way the process can be kept under control when the conditions are slightly modified.

After the presentation of the technological processes we discuss main aspects concerning the design and fabrication of masks, with direct reference to our DBRT devices. This part has been deliberately placed at the end of the chapter to emphasize once more what we have already mentioned in the beginning of this section: before starting the device design, getting knowledge of the technological aspects involved in the device fabrication represents an absolute necessity. Moreover, the technological process description contributes to a better understanding of the "rules" of mask design.

3.2 General Processes

3.2.1 Cleaning

Cleaning operations are performed before all major steps during device processing. It is very important that each of these steps does start after the undesired materials, present on the sample surface, have been removed. The formulation "undesired materials" has to be understood specifically, in function of the subsequent process. For example, the surface native oxide formed because of the sample handling and storage in the (clean room) atmosphere has to be removed before the metal or dielectric deposition, but does not matter in the case of wet or dry chemical etching. Also all major steps necessitate a photolithographic process, so removing of the photoresist between them also represents a cleaning operation. Another aspect is that one cannot obtain completely pure surfaces or chemicals in the real conditions so the term "clean" is referred to being clean enough to result in a reliable, reproducible procedure for the particular process step in question. For cleaning one may employ solvents, acids, bases, or plasma. There are also a number of commercially available proprietary compositions useful for cleaning, photoresist stripping, or oxide removal. Chemical purity is obviously important, so the cleaner solutions have to be in *semiconductor grade*, which is one step above *reagent grade*.

Water plays an important role in the processing chain being used, among others, at cleaning operations, for rinse. The quality of the water has to be high, so deionized (DI) water, with a resistivity higher than 6 MΩ, is the most suited. Stagnant water tends to decrease in resistivity and therefore the rinsing operation will be done in water flow, after the attainment of the corresponding purity conditions.

Organic solvents are effective in removing oils, greases, waxes, and organic materials such as photoresist or spin coating polymers. On the other hand, they are innocuous to almost all materials intended to be kept intact on the device surface such as the semiconductor layer materials themselves, metals, and dielectrics. Solvent cleaning is usually done at elevated temperature, often at the boiling point of the solvent. The most common practice is to immerse the wafer in the heated solvent using agitation. Also a useful procedure is to rinse the wafer several times with heated solvent. The solvents must be discarded and replaced after a number of uses, considering that the contaminants accumulated in the solution affect the degree of cleaning of the subsequent wafers. After the solvents are used, they must in turn be removed from the wafer. Some of them (see Table B1, Appendix B) are not soluble in water. These solvents, however, are soluble to some extent in alcohols or acetone, which are completely miscible with water.

Acids are used in cleaning steps to remove metal contaminants and especially very thin films of oxide formed on the sample surface after wet chemical etching, oxygen plasma etching or ordinary exposure to air. Depending of the previous step, the resulting oxide is between 1 nm and 5 nm thick and can be removed in highly diluted solutions of acid. The acid acts on the oxide and usually will not etch the semiconductor (GaAs). The choice of the acid is in function of the materials present on the sample but a common choice is diluted hydrochloric acid (HCl).

Bases and alkaline solutions can act as cleaners for some types of materials and as well as acids dissolve the GaAs oxides which have an amphoteric character. Again, these are usually used in diluted concentrations. Because the basic solutions tend to cause photoresist adhesion problems their use in the presence of resist patterns will be considered with caution. The most common choice from this category is diluted ammonium hydroxide (NH_4OH).

Plasma etching can also be used for cleaning. Thus, oxygen plasmas are suitable to remove organics such as photoresist from the sample surface (ashering). The plasma causes O_2 to dissociate into highly reactive oxygen atoms that reacts with carbon and hydrogen in resist to produce volatile waste products (CO, CO_2, H_2O). The residual photoresist that may remain in exposed and developed portions of the resist pattern after photolithography, often because of the scumming phenomenon (redeposition of the dissolved unexposed resist in the developed spaces), is removed using a brief oxygen plasma etching step. Oxygen plasmas do not etch dielectric films or GaAs, but produce a shallow oxidation of the GaAs.

During the processing of the resonant tunneling devices the solvent cleaning has been used prior to each application of resist and dielectric films. Also, the removal (stripping) of the resist after the major steps was done with a solvent (acetone or dimethilformamide (DMF)). After each photolithographic step, before the etch or metallization processes, the samples have been exposed for a short period of time to oxygen plasma. Cleaning in diluted HCl has been performed prior to metal and dielectric coatings.

3.2.2 Photolithography

Generally speaking, all the major steps in the fabrication process of an electronic device are based on two essential types of activity, *pattern definition* followed by *pattern transfer*. The pattern definition is accomplished by specific lithographic processes. After this, using *subtractive* (sputtering, wet or plasma etching, etc.) and/or *additive* (lift-off, selective deposition, etc.) techniques the pattern is transferred into the sample. Lithography thus represents the pivot of any device processing and the evolution towards smaller devices has increased its importance.

Lithography is the process by which a pattern is defined in a thin layer, the so-called *resist* layer, which is sensitive to irradiation by photons or particles. The resist layer becomes soluble under the influence of the irradiation, *positive* resist, or insoluble, *negative* resist, in specific developing solutions. The performances of the lithographic process depend both on the technique and the resist used.

There exists a variety of techniques for the lithographic definition of patterns. The techniques can be divided into two main categories: *direct write* and *projection* lithography. In the first case the pattern is written in the resist by scanning a beam of electrons or ions (in exceptional cases photons) over the resist layer. The process is *sequential*, i.e., the parts of the required pattern (pixels) are exposed or irradiated one after another. In the second case the pattern is defined in the resist layer by projecting a *mask* onto the resist film. Both photons (Middle/Deep UV and soft X-rays) or particles (electron or ion beam projection) can be used. The projection can be simple 1:1 shadow image or reduced image of the mask. In the following we shall refer only to the technique used in our laboratory. This belongs to the second category, using UV radiation in the *contact (proximity) printing* variant. Concerning the first category we note that the direct write technique using electrons or light have been used for the fabrication of our masks.

In the contact printing optical lithography the mask is placed in proximity to the sample for alignment, but is clamped directly against the sample for exposure. Our photolithographic system, a Karl Suss MJB 3 UV 300/400, permits both mechanical clamping, *soft* contact, and vacuum clamping, *hard* contact. Between this two extremes there exists another possibility, the *nitrogen pressure* contact. The system light source is a mercury lamp and using specific filters (front lenses) exposures around the h-line (405 nm) and i-line (365 nm) of the Hg spectrum (Middle UV) are possible.

The contact printing lithography, as well as the other techniques using light, is limited by the diffraction effects at the edges of the mask features. This is reflected both in dimensional differences between the resist image and mask features, and in the resolution of the method or ultimate mask dimension replicated in the resist film. The conditions of the contact printing lithography correspond to those for which the Fresnel or near field diffraction occurs.

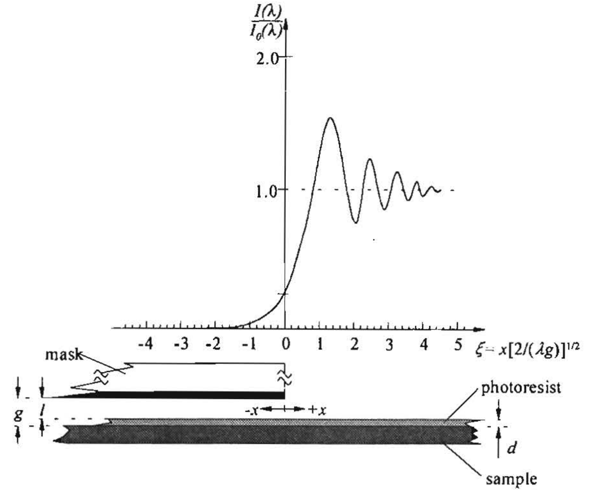


Fig. 3.1 Fresnel diffraction pattern at a straight edge (semi-infinite slit).

The simplest way to account for diffraction phenomena in resist image formation is to approximate the clear mask features by semi-infinite slits. The problem is then reduced to the Fresnel diffraction by a straight edge (see Fig. 3.1) - an example appearing in all optics textbooks [1]. In coherent light, the intensity distribution at the sample level extends under the opaque mask zone and presents undulations (diffraction maxima) in the clear half-plane. In this way, for a certain incident energy on the mask, the *effective* edge of the resist image occurs at a position where the incident energy on the resist E_i equals the resist threshold energy E_T, which defines the minimum energy necessary for a full development of the resist film [2]. Ideally, it would be to control the exposure conditions with such a precision that the resist image effective edge to be situated at the corresponding border between the clear and dark zones of the mask. In practice, however, this is impossible and consequently a dimensional deviation δ from the mask features is registered in the photoresist image. The aspect of the intensity distribution depends on the gap g between the mask and sample, that is the resist thickness d plus the distance mask-resist surface l. Increasing the gap, the light penetration in the dark half-plane increases and the gradient of the intensity decreases [3]. This results in a higher δ and a deterioration of the edge features.

Reducing the clear and dark zones to the finite dimensions of a real mask, the diffraction pattern changes but the general aspect is maintained. The changes concern the intensity gradient and the distribution and amplitude of the diffraction maxima, which now appear too in the dark zone [3]. So, it becomes evident that the edge diffraction effects dictate the minimum line width transferred in the resist film. Fig. 3.2 shows schematically the light intensity distribution of a grid with equal clear and dark narrow lines, at the sample level. The practical minimum period $2b_{\min}$ of the grid, or resolution, which can be replicated by shadow printing using collimated light for exposure can be expressed semi-empirically by [4]

$$2b_{\min} \cong 3\sqrt{\lambda\,(l+\frac{1}{2}d)} \tag{3.1}$$

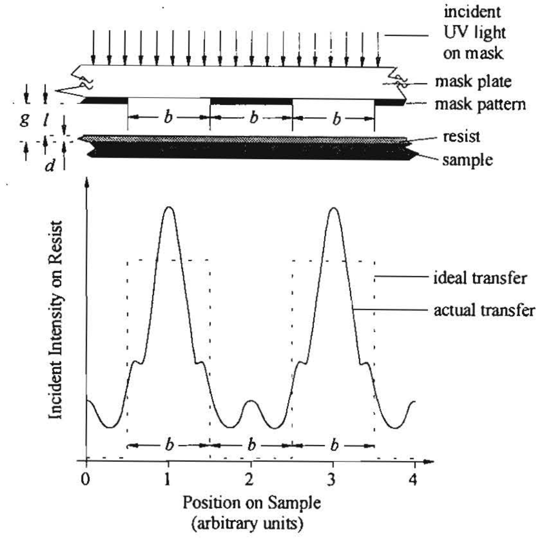


Fig. 3.2. Intensity distribution at the sample level through a grid with equal clear and dark feature [4].

where λ is the wave length of the used light.
Practical resolution refers to the limit dimension which can be obtained in real conditions, considering the (small) tolerances in resist, exposure and development parameters. The optimum resolution could be therefore better than that given by the expression 3.1.

In general the mask features to be transferred into the resist are not simple lines but different geometrical shapes, the most encountered being the rectangles or circles. Near corners diffraction effects result in a much more complicated intensity distribution than the one sketched above for edges [3]. This has a two-dimensional character and finally leads to the roundness of the corners in the resist pattern. The effect is accentuated with the increase of the gap between mask and sample and constitutes a serious limit in the replication of the small features, when, for example, mask squares appear as rhomboids after the transfer in the resist film.

a) Direct Process Positive Resist Photolithography

As we noted at the beginning of this section, the resists used in different lithographic techniques can be positive or negative depending on whether the exposed area is removed or remains during the development process. As a general rule, the positive resists are the preferred choice for the III-V compound device processing, offering higher resolution and more appropriate edge profiles for metallization processes. All current positive photoresists operate on the mechanism involving destruction of a dissolution inhibitor. These resists are composed from three ingredients: an acidic polymer (resin) that dissolves or disperses in aqueous alkali (developer), a photosensitive dissolution inhibitor which forms a solid solution with the acidic polymer and prevents or inhibits the dissolution of the latter in an aqueous alkali, and a suitable organic solvent

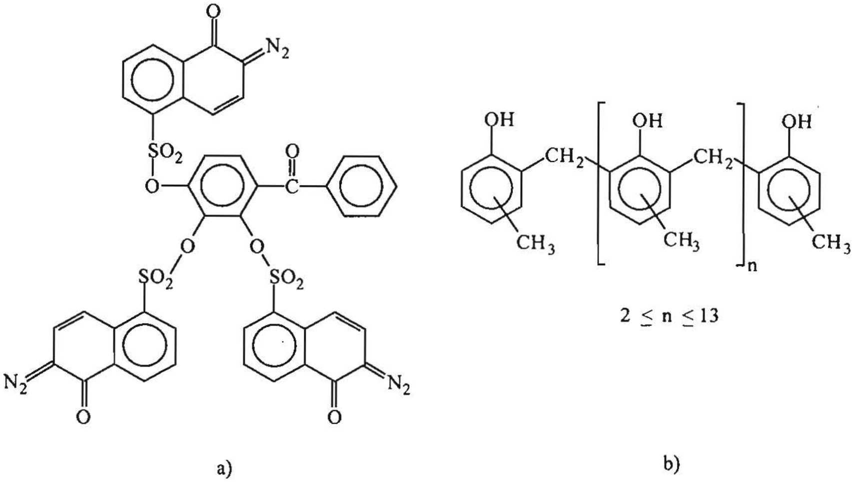


Fig. 3.3 Chemical structure of the two solid components of AZ positive resists: substituted o-naphtoquinone diazide (a) and novolac resin (b).

system. The acidic polymer can be either weakly acidic such as phenol- or cresol-formaldehyde condensation polymers of low molecular weight known as novolac resins or strongly acidic such as acrylic acid containing polymers. The dissolution inhibitor is in all cases based on the ortho-quinone diazide group. In the following we shall refer to the AZ class (Hoechst) positive photoresists which we have used in our photolithographic processes. These resists are based on the novolac resins and the substituted o-naphtoquinone diazides. Fig. 3.3 provides representative chemical structures for these components. The two solid components are dissolved either in a combination of cellosolvate acetate (EGMEA) and xylene, which is toxic, or, in a safer solvent, propylene glycol monomethyl ether acetate (PGMEA), as is the case with the new versions.

On their own the novolac resins are readily soluble in dilute aqueous alkali with a dissolution rate of approximately 15 nm/s. When the diazoquinone derivative is added, typically in the amount of 20-30 weight percent, their rate of dissolution is greatly reduced, to the range of 0.1-0.2 nm/s. A proposed mechanism for this low solubility is based on the alkali-induced azo coupling between the resin and diazonaphtoquinone resulting in a cross-linked structure [5]. On exposure with UV radiation in the range 300-450 nm, the diazonaphtoquinone turns into an acid that is soluble in aqueous alkali and which is no longer a dissolution inhibitor but acts as a dissolution promoter and results in an increased dissolution rate of the novolac resins, up to 100-200 nm/s, depending upon the fraction of inhibitor remaining. In this way the image discrimination for this kind of resist is based on a kinetic effect - the presence of a certain content of inhibitor in the novolac resins determines a much lower dissolution rate of the unexposed resist (although never equal to zero!) and a higher dissolution rate of the irradiated one in comparison with the pure resins [6].

The overall photoprocess of diazonaphthoquinone (DQ) is based on a mechanism proposed early on by Süs [7]. This is illustrated in Fig. 3.4. In the first reaction step, nitrogen is eliminated

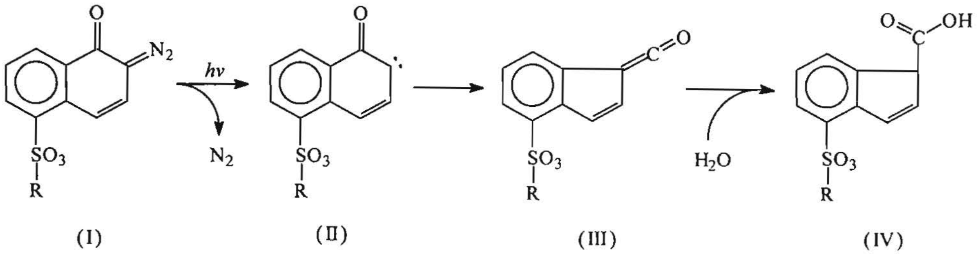


Fig. 3.4 Photoreaction steps of the diazonaphtoquinone (I) [9]. The intermediate products are the carbene (II) and the ketene (III), and the final one is indene carboxylic acid (IV).

and a carbene (II) is formed. This undergoes a Wolff rearrangement with ring contraction, leading to the ketene (III). In the presence of small amounts of water the ketene is finally hydrated to the indene carboxylic acid (IV). The crucial feature of the process is the change from a hydrophobic and nonionizable compound to an ionizable hydrophilic species, indene carboxylic acid. Although not general accepted [8], the nitrogen resulted from the photoreaction seems to be, at least in part, the cause of the acceleration effect of the exposed resist dissolution rate [9,10]. Thus, the escaping gas creates additional free volume and facilitates diffusion of the developer into the polymer matrix. The reaction of Fig. 3.4 describes a somewhat idealized process which is part of the operation principle of the novolac-diazonaphthoquinone based positive resists. During the resist processing some other less helpful reactions do also occur [11, 12]. Between them we mention two which, in fact, are the consequence of the compulsory thermal treatment of the spin coated resist layer, before exposure. First, the thermal decomposition of diazonaphthoquinone leads, in the presence of water, to the formation of the soluble indene carboxylic acid, as after its photodegradation. This has two undesirable effects - on the one hand the solubility of the unexposed resist increases and on the other hand a loss of the photosensitivity is registered - resulting in much lower contrast. Second, if the water content of the resist is insufficient, the ketene, resulted in the photo- or thermo-degradation of the inhibitor, may react directly with the resin and produce crosslinks. A relatively small amount of crosslinking can significantly decrease the resist solubility compromising the lithographic process.

The pattern definition with the AZ positive (dual image) photoresists is a multi step process involving resist spin coating, bake operations, exposure and development. The process flow is shown in Fig. 3.5.

The liquid state of the photoresist makes possible its application to the wafer surface by spinning. The spin coating conditions have to insure the obtaining of resist films which are free of contaminants and pin holes, and have high uniformity and good adhesion. Before resist dispense, the sample surface has to be perfectly clean and dry [13]. Water traces are the major cause of poor adhesion. Therefore a bake step is always included to ensure complete dryness. Also the environmental humidity, which can also affect the later resist photoprocess, must be kept strictly under control (generally below 50% at a temperature of 21 °C). If the resist is applied over an oxide , or any dielectric, usually the poor adhesion can be improved by using an adhesion promoter (e.g. hexamethyl disilazane (HMDS)) prior to resist application. During the resist

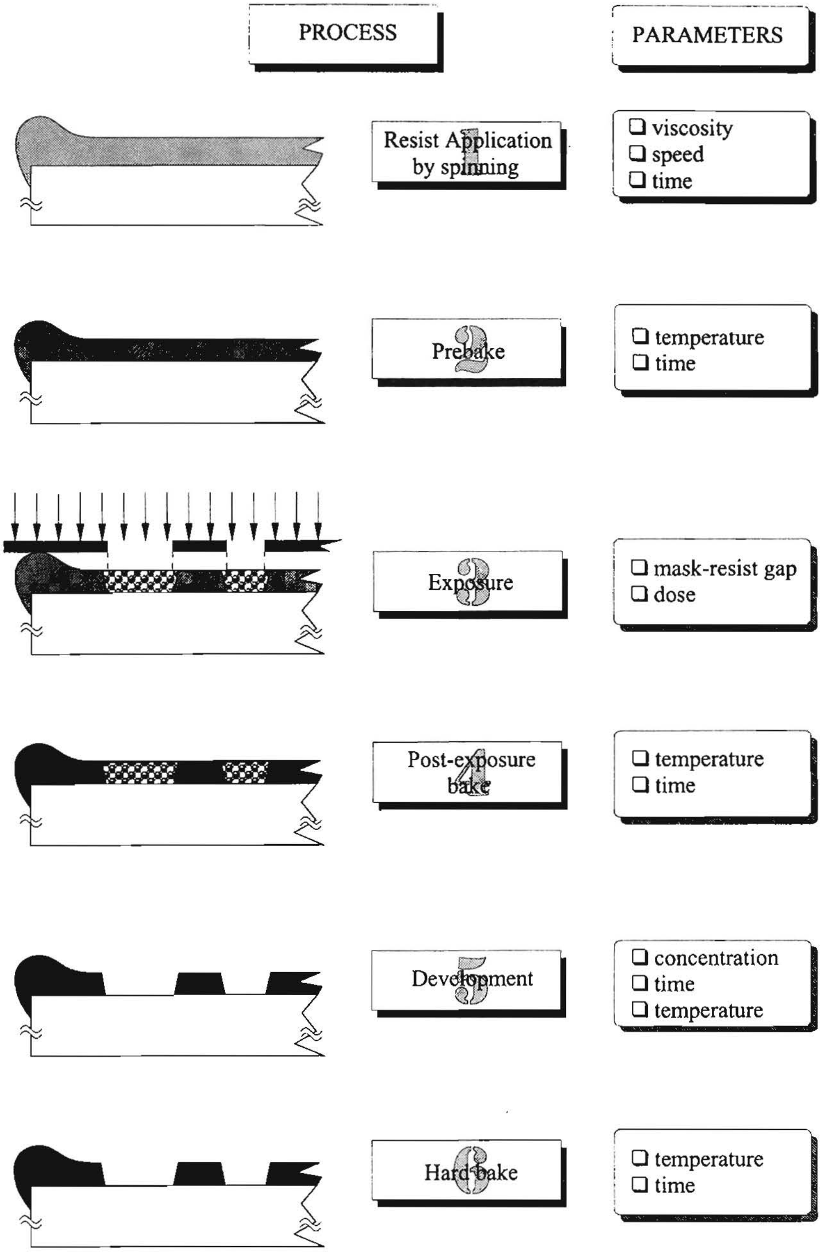


Fig. 3.5 Photolithographic process flow for AZ positive resists.

application the best results are obtained when this slides onto the sample surface without splashing. So, the dispense tool (usually a pipette) will be situated as near as possible to the sample surface and the resist spreading will be done at low spinning speeds or for small surfaces directly with the tool. As soon as the entire surface is wet with resist the spinning speed is increased until a resist specific optimum value (for AZ resists the range is between 3000-6000 rpm) and kept for a specified interval of time (20-40s). During this stage most of the dispensed resist is spun off and together with a dynamic dry process a uniform film is left behind. Acceleration from low to high spin speeds affects the uniformity of the film, higher values giving better results. The thickness of resist films coated on high-speed spinners can be expressed by the following relationship [14]:

$$d = \frac{\kappa S^2}{(rpm)^{\frac{1}{2}}} \tag{3.2}$$

where S represents the percentage of solids in the formulation, *rpm* is the spinning speed in rotations per minute and κ is a constant of proportionality which varies with the intrinsic viscosity of the resin, hence the coated thickness is sensitive to the molecular weight of the polymer. In practice the resist film must be sufficiently thick to cover the underlying topography, so for high step topography more viscous resists, coated at lower spin speeds, are the most suitable. A last problem we mention concerning this step is the edge bead formation after the resist spin coating. This refers to the fact that in the proximity of the sample edges (approx. 10-20 μm, depending on the resist thickness) the film is at least twice as thick as in the rest (at the corners this effect is even more pronounced) and that the sides and sometimes partly the back side of the sample are also covered with resist. When the dimensions of the pattern that is transferred in the resist is at the limit of the resolution of the photolithographic process used then the edge bread represents a serious impediment and must be removed.

During the resist processing many heating or bake operations can be performed. The first one, known as prebake, softbake or pre-expose bake, follows immediately after the resist application and is compulsory. Its role is to convert a liquid-cast resist film into a solid one, with reproducible thickness and sufficient adhesion to the sample surface. The physical removal of the casting solvent till a stable (low) level must be done without degradation, as much as possible, of the resist components. Ideal is to remove all the solvent content from the resist film to avoid in this way its interference with the radiation chemistry which influence the exposure-development conditions. Several physical and chemical factors determine the efficiency of solvent removal and the upper prebake temperature. Thus, the glass transition temperature (T_g) and the decomposition temperature (T_d) of any resist components are the guide for a practical upper limit. For efficient solvent removal, the prebake temperature should be above the T_g of the resin , but should not be high enough to decompose the photo-sensitive component (inhibitor). Also, to remove all of the solvent one requires heating above the T_g of the polymer and the solvent boiling point. In AZ-type resists the diazonaphtoquinone has $T_d = 100$ °C, the novolac resin has T_g in the range 70-120 °C and T_d in the range 150-300 °C, depending of isomeric composition, and as solvents, high-boiling (>125 °C) liquids are used. In these conditions the prebake temperature is limited by the diazonaphtoquinone T_d and up to 15% solvent by weight can remain after prebaking in the range 70-100 °C [15]. Above 70 °C the diazonaphtoquinone becomes already thermal unstable and starts to decompose [16, 17]. Every 10 °C change up in temperature has a significant effect on the extent of its decomposition and, as we have seen, on the subsequent resist dissolution rate. Due to solvent loss the resist films undergo physical changes, as shrinkage and stress. The shrinkage is commonly of the order of 5-10% . Fortunately above 70 °C the resist thickness is almost insensitive to the increase of temperature [18] and thus its reproducibility is not anymore a problem. Since the AZ resist films cannot be baked above their T_g, to induce some flow, the solvent loss produces internal stress in the coating plane. This internal stress is, in part, reduced by the above mentioned solvent retention, due to a plasticizing effect. The residual stress can cause sometimes adhesion failure during development.

A second bake can be introduced after the exposure. This, so called post-exposure bake, is performed at a temperature higher than the inhibitor T_d (e.g. 110 °C), in conjunction with a low prebake temperature (e.g. 75 °C) [19]. The low prebake temperature avoids diazonaphtoquinone thermal degradation and leaves more water in the film for exposure and post-exposure bake, thereby eliminating the uncontrollable crosslinking in the exposed regions. Also, the water and especially the solvent, remained after the prebake step, combined with the high post exposure temperature help to the vertical homogenizing of the photoactive compound (unbleached diazonaphtoquinone) and the photoreaction product (indene carboxylic acid), which after exposure have a variable concentration because of the standing wave effects. Using post-exposure bake the exposed resist shows a very uniform vertical development (constant solubility) across the sample that is very important for the critical dimension control. The "crust" formation on the top of the unexposed resist, caused especially by a higher loss of water at the surface and the subsequent thermal induced crosslinking, is increased in the conditions of the high post-exposure bake temperature, minimizing the dissolution rate of this region during development. All these contribute to increased contrast, improved linewidth control (even in the situation of a high step topography), elimination of standing wave effects, and reduction of scumming. Finally, the post-exposure bake helps the removal of N_2, resulted during the inhibitor photoreaction, from the interface photoresist- sample, promoting adhesion.

The third bake, usually known under the name of postbake or hardbake, follows the development step, finishing the photolithographic process. After development and rinsing the interfacial adhesion is weakened by developer and water penetration along the resist-sample interface. The postbake may restore the lost adhesion and for this reason appears as a very necessary operation before a wet etching process. For AZ type resists improved adhesion has been noted [17, 20] at bake temperatures of 120-150 °C , when the solvent and water are removed completely and the slight melting of the resin (T_g of the novolac is situated in this range) acts as a hot melt adhesive to seal the interface between the resist and sample. However, if the bake temperature is too high the resist starts to flow. This phenomena destroys the lift-off profile and a considerable linewidth distortion (image rounding) may occur. Fortunately, during the postbake another process competes with the resist flow, increasing, at least at the surface, the resin T_g. This is the already mentioned "crust" formation, which, unlike the case of the previous bake operations, is much more accentuated. Surface crosslinking is produced now not only because of the thermal decomposition of the photoactive component but also, equally, by the resin oxidation [21]. The superficial hardening of the resist improves the resist durability under (plasma) etching conditions and maintains, when need be, the lift-off profile.

Resist exposure follows immediately after the prebake step. It plays the central role in the photolithographic process, producing, as we have seen, the photochemical changes in the resist layer that, later, during development, result in a differential solubility. Regarding this step, in certain conditions, i.e. exposure equipment, masks and photoresist, the expected results are obtained by a proper (adequate) choice of the exposure energy or dose. The optical data offered by the resist producer, as absorption spectrum and characteristic (contrast) curve, may constitute the starting point in establishing the dose. It must be remembered, however, that for a positive resist the image discrimination is based on a kinetic effect and consequently its characteristic curve applies for a particular set of processing conditions, that is, for specific prebake time and temperature, developer and development time, etc., which in many situation are not or incomplete specified. So, a step more is to determine in the own conditions the so called "latitude"plot for exposure. This plot figures measured widths of developed lines and spaces (islands and windows) or deviations of these from the nominal feature size (usually the critical dimension is used) on the mask as a function of exposure dose. The intersection point of the curves for lines and spaces, which corresponds to the exact reproduction of the mask feature size, gives the working exposure dose or the *work point* in the corresponding conditions. We can then establish, using the plot, the *exposure latitude* defined as the variation in exposure which keeps the difference between the mask dimensions and resist image dimensions within a specified limit, dictated by the tolerances of the related processes. Similar "latitude" plots can be traced for prebake (temperature variation) and development (time variation) on which the corresponding *work points* are determined. The tolerances expressed by the corresponding *latitudes* (process windows) offer the possibility to play with the step parameter values with a view to the attainment of the proposed goal. For a certain resist the working exposure dose depends of the reflectivity of the support, for metals the dose being less than for dielectrics or semiconductors. As we have seen, the diffraction effects produce deviations from the mask dimensions and the corner rounding, even in the situation of zero space between mask and resist surface. The edge bead resulted after the spincoating operation determines the maintaining of a space, at least of the resist thickness, resulting in the amplification of the above mentioned drawbacks.

The last step in the photolithographic process is development. Resist producers offer the most suitable developers for the different kinds of resist they fabricate. As mentioned at the beginning of this section, the developers for the AZ-type resists are aqueous alkaline solutions. The main parameters that, in given conditions of the previous steps, influence the development process are the concentration and temperature of the developer and the development time. Developer concentration is a quite sensitive variable. At any concentration, small changes in developer normality (± 0.001) can alter the resist dissolution rate by 10% [19]. Upon continued exposure to air, the alkaline developer strength is degraded, by CO_2 pickup and OH^- depletion, to a significant degree (0.003N depletion in 4 hrs.) [22]. This imposes that developer tanks should be purged with N_2 or otherwise covered and kept free of CO_2. Concentrated developer (0.6N), generally increases sensitivity but lowers the contrast. So, the commercial developers, usually delivered in concentrated solutions, will be diluted to convenient concentrations. The dissolution of the resist is a diffusion limited process. Consequently, because the diffusion coefficient of the developer intermixing within the resist polymeric constituent has an exponential increasing dependence with temperature, the rate of development increases in the same way. This circumstance imposes the strict control of the developer temperature during the same process and from one to another. Development time in combination with the exposure dose are very important in the control of the image sizes and resist edge profile. At high doses the line size is essentially dominated by the absorbed energy and independent of the short development time, but at low doses, when the developer dominates the absorbed profile, overdevelopment leads to large image size biases. Also the high doses determine undercut (vertical) edge profiles while the low doses and overdevelopment produce overcut profiles. Static development favours the scumming phenomena and nonuniform development. To avoid these problems moderate agitation is useful, bearing in mind the corresponding increase of the development rate.

In our different lithographic processes we have used several types of Hoechst AZ positive photoresists, that is AZ 1505, AZ 1350J, AZ 4330A and AZ 4620A. Their characteristics and processing "recipes" are presented in detail in the Appendix B. It is worth, however, to describe shortly the principal advantages and drawbacks of these resists, in conjunction with their most suitable uses.

AZ 1505 is a small-thickness resist destined for submicron applications (see Appendix B, Table B2). It demonstrates high sensitivity, high contrast and good process latitude. For these reasons AZ 1505 is the ideal resist for mask making with optical means. In our lab conditions critical dimensions of 1 μm have been obtained with this resist [23] without special precautions, such as the edge bead removing, hard contact exposure, post-exposure bake, etc. AZ 1505 can be used with very good results for low temperature wet chemical etching, especially in acid based solutions [24], but it does not work in the similar process at room temperature, except for the case of shallow etching (gate recess in MESFET or HFET processing [23]). Also, with normal processing, this resist fails in dry etching processes, where practically there is no etch rate selectivity between the resist and the semiconductor or dielectric to be removed.

AZ 1350J is the most used resist in mesa etching processes. Designed for "middle" thicknesses (see Appendix B, Table B2), it demonstrates a very good adhesion in solution chemical etching and, with careful hardening, has a satisfactory resistance in dry etching operations.

The last two mentioned AZ positive resists are thick resists, destined especially for contact-up plating. In combination, they are used for the fabrication of the air bridge interconnections (see section 3.3.3 of this chapter). AZ 4330 and 4620 have high transparency and sensitivity, which permit adequate exposure times. Their high thickness and destination demand special requirements in the bake steps to avoid stress effects. The two resists manifest also good plasma resistance.

Besides the above mentioned positive photoresists, since ~ 1991 dual image resists are available in our laboratory. Their specific processing and utility is described in the next paragraph.

b) Image Reversal Technique and Dual Image Photoresists: an easy way to pattern the resist for lift-off processes

Image reversal refers to the procedure by which a negative image is produced with a conventional positive resist. The image reversal can be accomplished in several ways. A first possibility is to use the already mentioned reaction of the ketene intermediate with novolac resin, occurred in the absence of water, the resulted crosslinks (neutral esters) rendering insoluble the exposed region. Thus, if the exposure is performed under vacuum (virtually anhydrous conditions) and then the resist is flood-exposed, the basic development yields the negative image [11]. Another modality is based on the decarboxylation of indene carboxylic acid that leads to neutral indene derivatives, insoluble in alkaline developers. The decarboxylation process is based catalysed and can be induced at comparatively low temperatures by the addition of a strong base to the resist film. Some of the catalysts, such as imidazole, monazoline or triethanolamine, can be added to the resist before coating [25] while others, such as ammonia, are applied as post-exposure treatments [26] (diffusion process). Subsequent flood exposure and development result again in a negative image of the original mask. At last, in a third and most important idea, the acid generated in the photolysis of diazonaphtoquinone has been utilized in yet another way [27]. So, the adding of a melamine derivative, as a crosslinker, to the novolac resin makes the mixture into an "acid hardening" composition. *Acid hardening resins* are usually crosslinked by strong protonic or Lewis acids. Carboxylic acids are sluggish catalysts at room temperature but become effective at higher temperatures, above 80 °C, when the reaction shown in Fig. 3.6 occurs. The comparative inactivity of the carboxylic acid at room temperature makes possible to use the system in two different ways. The acidic regions produced during exposure can either be removed by a basic developer and the system behaves as a normal positive resist; or they can be crosslinked before development, by heating, and, if subsequently the neutral regions (unexposed) are removed by flood-exposure and development in alkali, the resulted image is a negative one. The crosslinked negative image manifests thermal stability up to high temperatures (> 300 °C) unlike the normal positive resist images and the negative ones based on the previous two procedures, which are thermoplastic and flow when heated near the T_g of the resin. Also, even

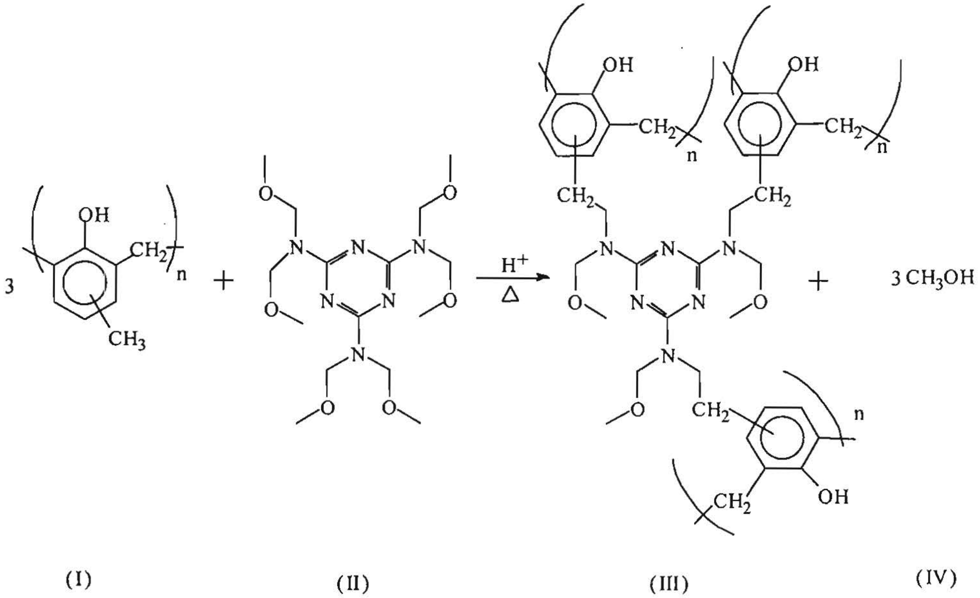


Fig. 3.6 Crosslinking process of the *acid hardening* composition, the novolac resin (I) plus the melamine derivative (II), induced thermally and by the catalytic action of the carboxylic acid. With the crosslinked resin (III) results more the methyl alcohol (IV), as a volatile byproduct.

the positive images prepared with this *dual image* photoresist can be thermo-stabilized producing crosslinking, finally, by flood-exposure and bake. The dual and thermosetting behaviour of this modified positive resist together with some other advantages, such as higher contrast and higher resolution made the procedure to be used for producing commercial image reversal photoresists. The AZ resist class also includes such kind of photoresists, in the AZ 5200 E series. Between them we have used the resists AZ 5206 E and especially AZ 5214E. The process flow for image reversal technique with these resists is shown in Fig. 3.7. Designed for i- and h-lines of the UV spectrum the two resists proved suitable for half-micron lithography, using image reversal process in optimized conditions [28, 29]. For our needs, however, two other possibilities offered by these resists have been exploited.

First, the dual character of the resists permits to be realized , in the positive image regime, the patterning of resist windows with dimensions equal or less than the mask nominal ones. Thus, by introducing a proper flood exposure and a subsequent corresponding bake before the exposure step (see process B.14), the known diffraction effects, materialized in the enlargement of the transferred mask windows, can be limited by the formation at the resist surface of a shallow crosslinked layer insensible to any further irradiation.

Second and much more important, using the dual image photoresists we have been able to develop, in both positive and image reversal regimes, resist patterning processes for *metal (dielectric) lift-off* operations. The lift-off process represents almost exclusively the procedure of metal patterning in III-V compound semiconductor processing. As seen in the process illustration from Fig. 3.8, the resist acts as a vertical separator (standoff) evaporation mask and blanket metal deposition leaves uncovered a narrow rim of the substrate, which is shielded by the resist overhang, as well as the side walls of the polymer pattern. If the sample is then dipped in a specific solvent, this penetrates through the uncovered areas, dissolves the resist and the undesired metal is lifted-off, leaving a clean metal pattern behind. Generally, an aspect ratio of twice the resist thickness to metal thickness is required. From the above short description we observe that the key element in the success of the lift-off process is the undercut edge profile of the resist pattern. With an overcut profile, the deposited metal ensures a complete protection of the photoresist against the solvent action and the process is compromised. Using the conventional positive resist the pattern edge profile can be modified by changing the exposure, bake and development conditions, but in almost all situations an overcut (vertical) profile is obtained (diffraction and interference effects), unless rather exacting additional processes are performed during the normal lithography, such as the post-exposure soaking of the resist film with chlorobenzene. Of course, there exist in addition for this purpose the multilevel techniques [19, 30], but they need also other kinds of resist and, again, complicate the processing in a very high degree. That is why the AZ dual image photoresist offer the simplest way to produce resist patterns with undercut edge profile. The same exposure factors which determine an overcut edge profile for the conventional positive resists will favour in an image reversal processing of the dual resists an undercut profile. So, optimizing the image reversal process with respect to the quality of the resist transferred image we have got at the same time the desired undercut profile for lift-off procedure (see processes B.9 and B.10). And, because we have not always available "negative" masks to use image reversal process, we have developed with the same dual resist a procedure for obtaining the undercut profile in the positive image regime. The procedure is based on the same idea used for limiting the enlargement of the transferred mask windows, described

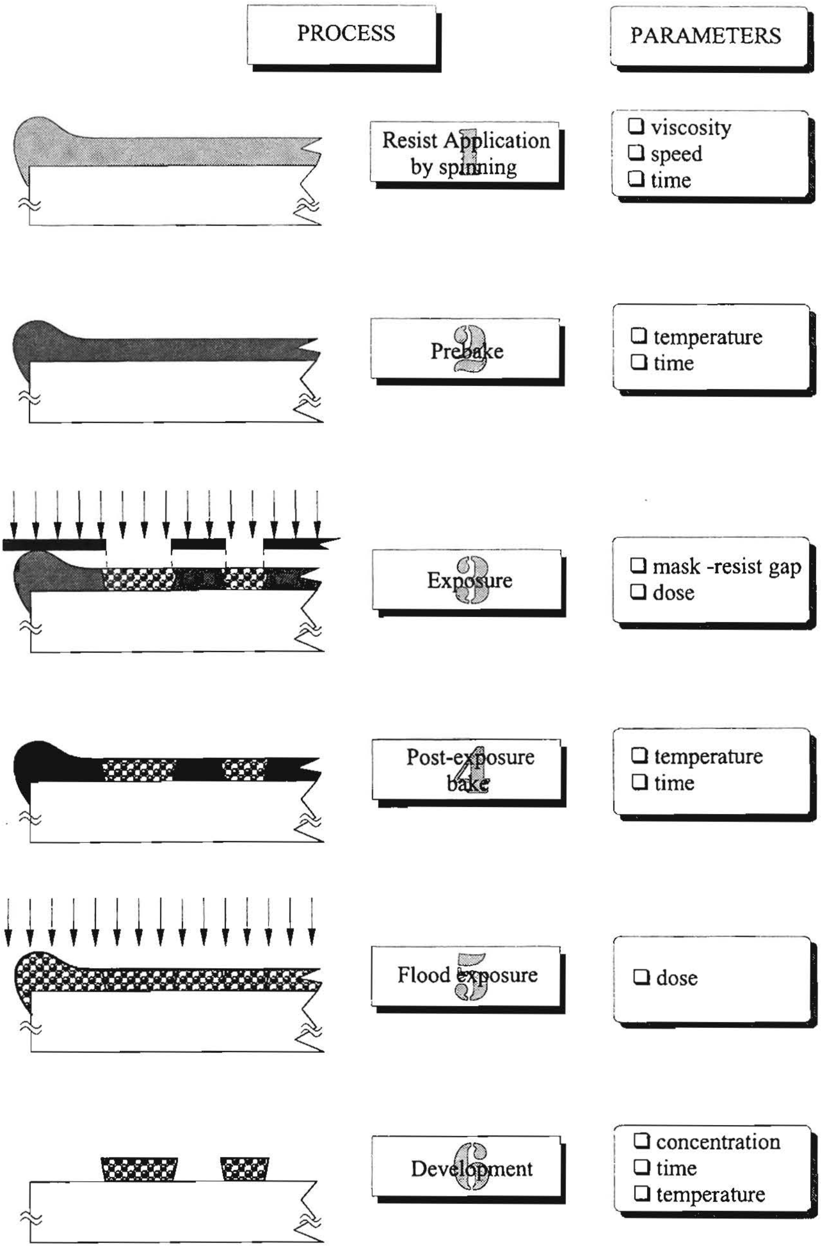


Fig. 3.7 Image reversal photolithographic process flow for dual image resists.

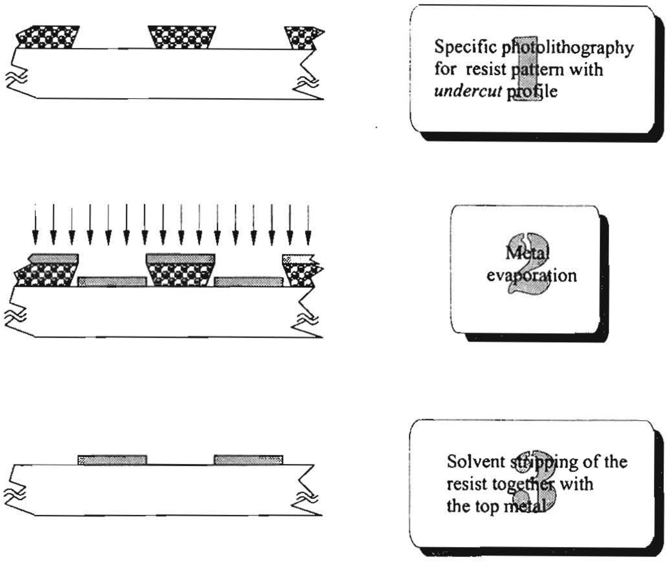


Fig. 3.8 Illustration of the lift-off process.

above, and then has the same sequence of steps, but with other parameters (see processes B.12 and B.13). In this case formation of a crosslinked top layer in the resist film, before exposure, determines in the unexposed areas a vertical gradient of the dissolution rate, orientated to the substrate. This results finally in the creation of an undercut edge profile in the resist pattern.

Concluding this subsection, we observe that the results obtained in the lithographic processing clearly depend on many aspects (variables) and agree with the quip found in [31], that "a good understanding of the principles involved in each aspect makes resist process development merely *difficult*, without such understanding, it is *impossible*". So, besides knowledge in process mechanisms (theory) other things, as careful experimental development, scrupulous attention to detail, and stringent process control (practice), are very important to guarantee the success in this field.

3.2.3 Wet Chemical Etching

Etching or the removing of solid material is an indispensable "tool" in the field of material science and technology, particularly that of semiconductors. Generally, one can distinguish two main types of etching, the dry etching and the wet etching, both of them including a large variety of techniques. Dry etching refers to the techniques that use plasma-driven chemical reactions or energetic ion beams to remove material. Wet etching procedures aim the same thing using liquid etchants. Despite the spectacular progress that dry etching has registered and its expansion to semiconductor device fabrication, especially that of submicron devices, wet chemical etching, with its simplicity (easy access) and multitude of (countless) possibilities, still plays an important role.

Wet chemical etching of semiconductors is used for producing high purity shiny flat surfaces (polishing), for shaping (mesa structures, recessed gates), for cleaning purposes, as well as for characterizing structural and compositional features (diagnostics). Other materials, such as metals and dielectrics, may also be etched in specific solutions, as part of device fabrication procedures. The high specificity of some etchant solutions against the semiconductor nature makes them suitable in selective etching processes. In this section we discuss aspects of the wet chemical etching of GaAs in connection with the definition of mesa structures for device isolation. Selective wet and dry etching processes for GaAs/AlAs system are presented in section 3.3.2.

The fundamental reactions underlying GaAs and in general semiconductor wet etching occur on the material surface and are electrochemical in nature. They involve oxidation-reduction reactions, followed by dissolution of the oxidation products, by complex formation. Etching solutions, therefore, contain an oxidizing agent, such as hydrogen peroxide (H_2O_2) or nitric acid (HNO_3), and a complexing agent, which can be ammonium hydroxide (NH_4OH), sulfuric acid (H_2SO_4), phosphoric acid (H_3PO_4) or citric acid ($C_5H_7O_5COOH$). The solutions also contain a third constituent for dilution. This is very often deionized water but methanol (CH_3OH) may be used also. Here, it must be remembered that almost all the oxidizing and complexing chemicals involved in etching operations contain a high proportion of water in the form in which they are generally used for the preparation of the solutions (see Table B.3).

The explanation of the oxidation-reduction processes is based on the supposition that at the semiconductor surface there exist both anodic and cathodic microscopic (localized) sites [32, 33, 34]. Oxidation of Ga or As atoms occurs at anodic sites, while the oxidant is reduced at the cathodic sites. Both electrons and holes from the semiconductor surface are involved in the oxidation-reduction reactions. Therefore the processes are sensitive to any mechanism that can supply (or remove) electrons or holes at the surface, such as illumination and electrical currents. The Ga and As oxides do not dissolve in the dilution constituent. This is the reason that one introduces in solution the complexing agent, to react with the oxides and transform them in a soluble complex. During the etching, the reacting species are continually used up, so a whole picture of the process must consider also the means whereby these reach the surface and the subsequent reaction products are removed. The transport mechanism can only be the diffusion through the liquid.

One of the two basic mechanisms, mentioned above as being involved in the etching process, limits the dissolution rate. First, the etch rate may be controlled by the rate of the chemical reactions occuring on the surface. This situation is referred to as *reaction-rate*, *surface-rate*, or *kinetically-limited* etching. Second, the etch rate may depend on the transport of chemical species by diffusion to and from the surface. In this case one speaks about the *diffusion-limited* or *mass-transport-limited* etching. Diffusion-controlled processes have lower activation energies (of the order of a few kcal/mol) than reaction-rate limited processes, and therefore are relatively insensitive to temperature variations. They are, however, affected by agitation to a much greater extent, leading to an increase in the etch rate because of the intensification of the mass transport. In contrast, stirring does not change the dissolution rate of the kinetically limited processes. The etch depth is linearly proportional with etching time for the reaction-rate limited case but depends of the square root of the etching time for the diffusion-limited situation. Changes in etching conditions, such as temperature, as well as variations in relative proportions of the etchant constituents can change the limiting regime. Additional factors that influence the rate of etching include type and concentration of doping atoms, lattice defects, and surface structure (patterning).

Mass-transport limited etchants tend to action more isotropic with respect to crystal orientation, so they give a more polishing effect [36]. This occurs because surface protrusions are more available to incoming species, and tend to be etched more rapidly than a smoother surface. Reaction rate-limited etches, on the contrary, tend to preserve surface morphology, although this manifestation can be completely dominated by another tendency, observed almost in all cases: to be highly anisotropic in etching of the crystalline structure through mask patterns. The anisotropic nature of etching behaviour follows from the lack of symmetry in the GaAs lattice and the etch rate dependence on crystal orientation. Surface orientation effects are illustrated in general considering the behaviour of the {111} planes in the etching process [32]. These planes consist either of all Ga atoms or all As atoms, and are referred to as {111}A or {111}B planes, respectively. The planes A and B are bonded alternately by sets of three bonds and sets of one bond. The arrangement is shown diagrammatically in Fig. 3.9. When the GaAs material is sliced to be prepared with {111}-type faces, it is the layers joined by the single bonding that are separated, rather than layers joined by triple bonding. This gives rise to one of the two surfaces being composed completely of Ga atoms ({111}A) and the other one of As atoms ({111}B). There is an important difference between the two sorts of {111} surface. Assuming that the surface atoms are electrical neutral, each Ga surface atom has three bonding electrons and each As atom has five. Three bonds are used up in holding on to the crystal. The surface Ga atoms, therefore, have no free electrons. The surface As atoms, however, have two electrons that can easily take part in any reaction. Accordingly, a (111)B surface should be more chemically active than a (111)A. Surfaces of other orientation have different structure. The {100} surfaces, for instance, consist, in principle, either of Ga or As atoms, both types being double bonded to the lattice. Since neither is preferred, a real {100} surface may be regarded as containing a mixture of Ga and As atoms. Using the above argument, one can predict that, for reaction-rate limited etching processes, a (100) surface etches faster than a (111)A surface but slower than a (111)B surface. This situation has been confirmed for a number of etching systems: NH_4OH:H_2O_2:H_2O

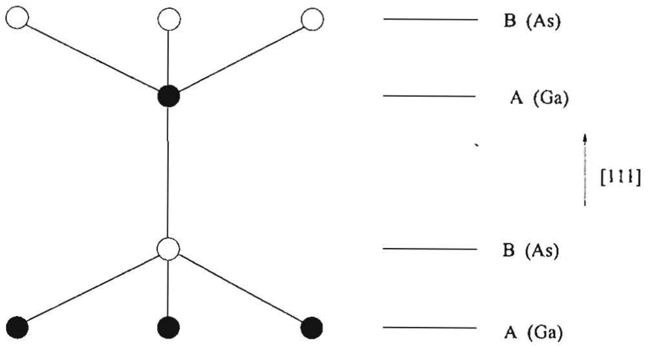
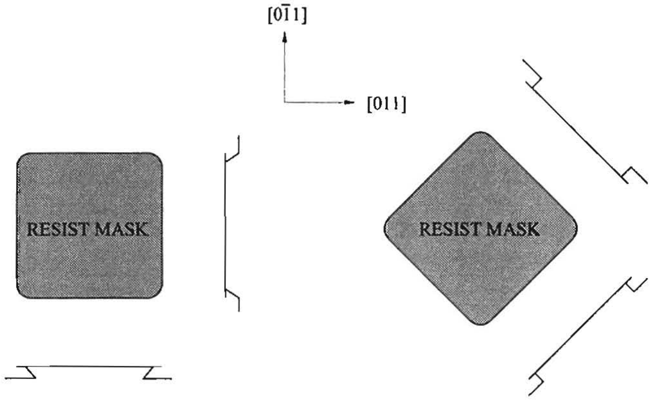


Fig. 3.9 Planes of atoms alternately joined by single and triple bonds in the GaAs lattice.

[35], H_2SO_4:H_2O_2:H_2O [37], H_3PO_4:H_2O_2:H_2O [38], etc.

The above description is somewhat simplistic, but it can be used to explain the general nature of the edge profiles obtained in a (100) oriented sample, by etching through mask patterns defined on its surface. (100) wafer orientation is by far the most commonly used one for device fabrication because it has two natural cleavage directions perpendicular to each other, [011] and $[0\bar{1}1]$. Devices are generally oriented parallel to these directions and the sample cutting into separate chips is not a problem anymore. Fig. 3.10 illustrates schematically the anisotropic etching behaviour generally obtained on such material. The squares in the figure represent mesas formed by etching away the material not protected by the mask pattern. If they are oriented so that their edges are parallel to cleavage directions the profiles have outward slopes (overcut) for the edges along the [011] direction and inward slopes (undercut) for the edges along the $[0\bar{1}1]$ direction (see Fig. 3.10a). Both types of edge profiles may find an explanation in the very slow etch rate for the {111}A crystallographic planes. The fact that the direction of incline changes for the two edge pattern orientations merely reflects the change in polarity of the chemical bonds associated with the {111}A planes in zincblend structures upon a 90° crystallographic rotation about the <100> axis [35]. Changing the mesa orientation by a ~45° rotation of the mask pattern in comparison to the previous positioning, the resulting etch profiles will be midway between the two situations encountered above, that is with essentially vertical walls (see Fig. 3.10b). In all cases, one notes that there is considerable undercutting of the masked material which means that the lateral etch rate is a significant fraction of the vertical etch rate. Unfortunately this fact can not be avoided in wet etching in general, representing an important limitation of the method. Specific etches have slightly different profiles than those from the general representation shown in Fig. 3.10. Both the edge profiles and the grade of underetching are important and must be anticipated in device design and fabrication.

In establishing reproducible and controllable etching processes some aspects must be considered. Thus, freshly mixed etchants may be hot or inhomogeneous, in both cases making necessary a specific interval of time between the preparation and the use. Also, the etching that



a) b)

Fig. 3.10 Anisotropic etching behaviour on a (100) oriented GaAs slice: (a) mask edges are parallel to the <011> cleavage directions and (b) mask edges are 45° to the cleavage planes.

occurs in a small volume of solution can cause its temperature to increase during the process. Compound etchants may change composition over time, depending on temperature and storage technique. Finally, some etching conditions may produce bubbles, usually hydrogen, which can adhere to the surface being etched and cause nonuniform etching. This can be alleviated by agitation of the solution during the etching process.

Although there is clearly a wide choice of GaAs etchants available [33, 39], all the specific device processing requirements can be satisfied by using only a few of these. In this way the process uniformity desideratum is much easier accomplished. Accordingly, we have looked for a few etching solutions suitable for mesa definition as part of our resonant tunneling device processing. The choice has been made after preliminary tests over four well known etching systems: $NH_4OH:H_2O_2:H_2O$, $H_2SO_4:H_2O_2:H_2O$, $H_3PO_4:H_2O_2:H_2O$ and $H_3PO_4:H_2O_2:CH_3OH$. It has to be mentioned here that in all etching processes, accomplished for experiments or device fabrication, the masking material has been photoresist. In our investigations we had in view several necessities (criteria) that etchants must satisfy: relatively medium etch rate (in the order of 0.5 µm/min), good quality of the sample surface and pattern edges after etching, and small underetch action. The tests have been made on GaAs samples covered with resist patterns representing squares, rhomboids and circles of different dimensions. The results are summerized in Table 3.1. As can be observed, for the first two systems the etch rate decreases, normally, with the dissolution, but for the second this is more drastically. The etch rate remains almost the same with the increasing of the NH_4OH concentration but increases evidently (approx. two times) with the increasing of the H_2SO_4 concentration. Both systems give ragged pattern edges, caused probably by the local attack of the resist by the etching solutions. This characteristic is illustrated in Fig. 3.11 for $NH_4OH:H_2O_2:H_2O$ (1:1:30) solution. In the picture, the strong anisotropic character of the etching can be observed as well. Also, the NH_4OH containing solutions present a relatively high underetch comparatively with all the other ones. Considering all the characteristics, between the tried solutions we made the choice for two of them, which have satisfied in a large measure our etching necessities. These have been $H_3PO_4:H_2O_2:CH_3OH$ (1:1:3)

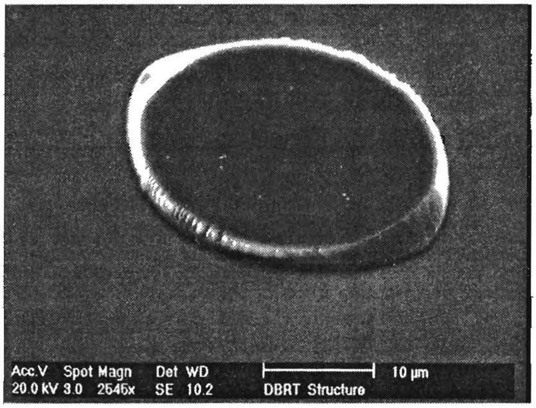


Fig. 3.11 Ragged edges of a mesa etched with $NH_4OH:H_2O_2:H_2O$ (1:1:30) solution. The strong anisotropy of the etching process is also evident.

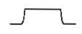
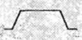
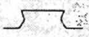
Etching system	*Volume ratios of the constituents*	*Conditions*	*Etch rate [µm/min]*	*Etch profile*		*Aspect*	
				(011)	$(0\bar{1}1)$	*Surface*	*Edge*
NH_4OH:H_2O_2:H_2O	1 : 1 : 10	- RT - ⊛	1.300			a	r
	1 : 1 : 30	- RT - ⊛	0.530			a	r
	1 : 1 : 50	- RT - ⊛	0.300			a	r
	2 : 1 : 10	- RT - ⊛	1.300			a	r
	2 : 1 : 30	- RT - ⊛	0.500			a	r
	2 : 1 : 50	- RT - ⊛	0.290			a	r
H_2SO_4:H_2O_2:H_2O	1 : 1 : 10	- RT - ⊛	1.000			a	r
	1 : 1 : 30	- RT - ⊛	0.140			b	r
	1 : 1 : 50	- RT - ⊛	0.060			b	r
	2 : 1 : 10	- RT	1.700			a	r
	2 : 1 : 30	- RT - ⊛	0.220			b	r
	2 : 1 : 50	- RT - ⊛	0.090			b	s
H_3PO_4:H_2O_2:H_2O	3 : 1 : 50	- RT - ⊛	0.085			a	s
	6 : 3 : 100	- RT	0.110			a	s
H_3PO_4:H_2O_2:CH_3OH	1 : 1 : 3	- RT - ⊛	1.700			a	s
	1 : 1 : 3	- 0°C - ⊗	0.34			a	s

Table 3.1 Wet etching test results on NH_4OH:H_2O_2:H_2O, H_2SO_4:H_2O_2:H_2O, H_3PO_4:H_2O_2:H_2O and H_3PO_4:H_2O_2:CH_3OH systems. The introduced symbols have the following meaning: RT - room temperature,⊛ - slow magnetic stirring, ⊗ - no agitation, a - acceptable, b - bad, s - straight, r - ragged.

at 0°C and H_3PO_4:H_2O_2:H_2O (6:3:100) (shadowed in Table 3.1). Their preparation and the etching conditions are described in Appendix B.

3.2.4 Contacts

Metal-semiconductor interfaces or contacts represent the "gate ways" through which we have electrical access to and control the versatile properties of the semiconductor devices. In spite of their obvious importance, however, contacts have been viewed in the past as an adjunct field, progressing primarily by semi-empirical means. In the field of compound semiconductors especially, up until about a decade ago, contacts evolved largely by trial and error. The impressive advancement registered in this field, marked primarily by the advent of the high quality epitaxial grown heterostructures and the progressive shrinking of the device dimensions, imposed very stringent requirements concerning the contacts. The exceptional performance characteristics of devices such as HBT, HEMT or DBRT diode could not be fully realized without improved contacts, with shallow penetration into the semiconductor and having very low specific contact resistance, high uniformity and stability. During these realizations the technologist's empirical approach, often involving as many as five metals deposited on top of the compound semiconductor was supplanted by material scientist's methodical one in the quest of contacts with the above mentioned properties. In this section the metallization schemes used for contacting our DBRT devices are presented. Mainly, we discuss the n-type contact developed on the emitter and collector GaAs contact layers of both two and three terminal devices. Also, we refer to the base metallization of the three terminal devices.

From the point of view of the electrical behaviour of the metal-semiconductor interface there are two principal types of contacts, the low-resistance contact and the rectifying contact. The low-resistance contact allows the current flow with negligible resistance. The commonly accepted term "ohmic contact" attributed to this makes reference to the desired linear I-V characteristics in both directions. At the other end of the spectrum is the rectifying contact, also known as the "Schottky barrier". Here a charge dipole, with its associated depletion region, is a necessary prerequisite and the resulting contact exhibits rectification properties, the current which draws under reverse bias being negligible.

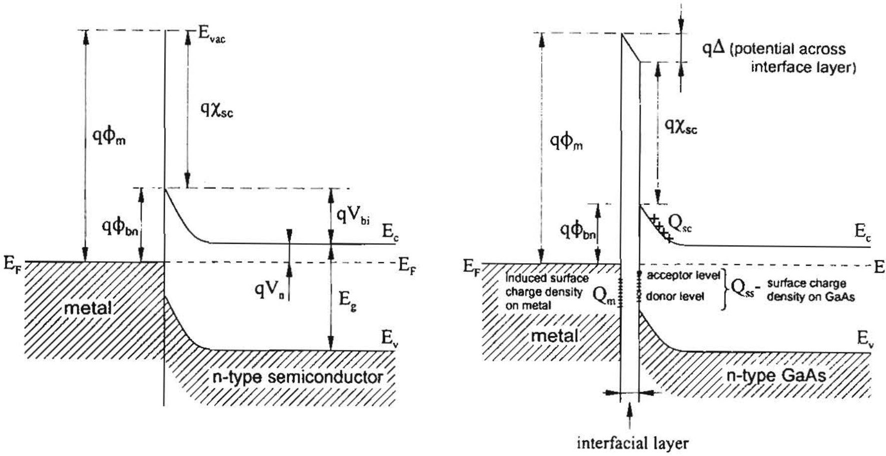
When a metal makes intimate contact with a semiconductor, their individual Fermi levels adjust to be identical on either side of the interface. As a result of the stabilization of this "local " Fermi level position, a potential barrier (Schottky) for carriers moving across the interface is formed. Ideally, the magnitude of this barrier is dependent on the metal work function, ϕ_m, and the semiconductor affinity, χ_{sc}. Thus, the electron barrier height, ϕ_{bn}, is given by

$$\phi_{bn} = \phi_m - \chi_{sc}, \tag{3.3}$$

and hole barrier height, ϕ_{bp}, by

$$\phi_{bp} = \frac{E_g}{q} - (\phi_m - \chi_{sc}). \tag{3.4}$$

where E_g is the band gap of the semiconductor and q the electronic charge (see Fig. 3.12a). Thus, for a n-type semiconductor, an ideal ohmic contact is obtained when $\phi_m \le \chi_{sc}$ and $\phi_{bn} \le 0$. Conversely, an ideal rectifying contact results when $\phi_m \ge \chi_{sc}$ and $\phi_{bn} \ge 0$. In this situation, the



(a) (b)

Fig. 3.12 Energy band diagrams for metal-semiconductor (GaAs) contacts: (a) ideal Schottky contact; (b) considering the surface Fermi level pinning because of states associated to surface defects [45].

choose of the suitable metal could result in the obtaining of the desired electrical properties of the contact. In practice, in the case of GaAs, however, this ideal behaviour is not observed. The barrier height for a metal/GaAs system is found to be virtually independent of the metal work function. This is thought to be due to the "Fermi level pinning" at the semiconductor surface, which occurs when the surface state density is 10^{13} cm^{-2} or more [40]. Initially, it was considered that these surface states could be intrinsic to the semiconductor surface, arising from the discontinuity of the lattice [41, 42]. Subsequently, however, investigation on "well cleaved" (110) GaAs surfaces showed that there are no detectable intrinsic surface states in the band gap of GaAs [43, 44]. Therefore the pinning of the Fermi level has an extrinsic determination. It could result from submonolayer coverages of different metals or oxygen, being independent of the nature of the contamination and roughly at about 0.8 eV below the conduction band minimum [44]. In connection with this an "unified defect model" for Schottky barrier formation has been elaborated [45], in which the surface states are associated with native acceptor or donor defects, for example, arsenic and gallium vacancies, produced by the interaction of GaAs with the contaminant (see Fig. 3.12b). Also, another idea argues that the Fermi level energy position at the interface is not governed by the surface state density but rather it is related to the work functions of microclusters of one or more interface phases, resulting from surface reactions occurred before or during metallization [46]. This means that an effective work function can be defined as the weighted average of the work functions of different interface phases. For GaAs (and in general the III-V compounds) the effective work function is primarily due to the work function of the arsenic (group V element) resulting from the fact that, at the interface native oxide-semiconductor, excess of this element is observed. Whatever the cause of the surface Fermi level pinning on GaAs would be, this reality makes impossible to obtain contacts with $\phi_{bn} = 0$ (ohmic contacts) by

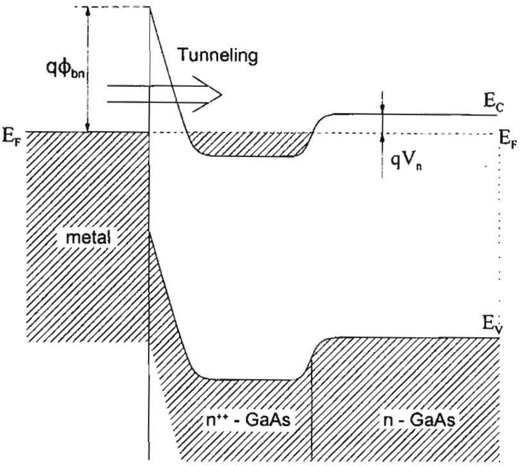


Fig. 3.13 Energy band diagram for a tunneling metal-GaAs contact.

simple choosing a metal with the appropriate work function. Consequently, two general procedures have been conceived to overcome this problem.

A first method is to dope a thin layer of GaAs beneath the contact to as high a level as possible. This results in a narrow depletion region and a corresponding thin electron barrier (see Fig. 3.13). In these conditions the barrier becomes thin enough for most of the conduction to occur by tunneling. For good ohmic behaviour this n^{++} layer must be doped to a level of $5x10^{19}$ cm^{-3} or more. Achieving these high doping levels in GaAs is not an easy task by epitaxial growth, only MBE being capable of this. However, appropriate metallization schemes including also a dopant, can, after specific thermal treatments, result in the obtaining of the highly doped layer at the metal-GaAs interface. This procedure is remotely the most commonly approach in the fabrication of ohmic contacts on GaAs and we also made use of it.

Another general procedure has in view the reducing of the effective Schottky barrier height to GaAs by the use of heterojunctions. In practice the method is to grow another semiconductor layer on top of GaAs and deposit the metal onto it. The chosen semiconductor must have an associated metal/semiconductor barrier height $\phi_{bn} \leq 0$ and be lattice matched to GaAs. This is illustrated in Fig. 3.14 in the case of the InAs as intermediate semiconductor layer between metal and GaAs. For an abrupt heterojunction the large discontinuity in the conduction band of InAs/GaAs implies an associated high barrier (see Fig. 3.14b). In order to overcome this and to improve the lattice matching, a graded heterojunction must be used (see Fig. 3.14c).

Principally there are three mechanisms that govern the current flow through a metal-semiconductor (GaAs) junction, suggested also by the above presentation. The first one, thermionic emission (TE), is dominant in the moderately doped semiconductors (~ $10^{17}cm^{-3}$). The barrier height is small, the carrier can easily surmount the top of this barrier by thermionic emission. For high barrier, as is the case of the GaAs, the vast majority of electrons are not able to overcome it and rectifying (nonohmic) behaviour is observed. The second mechanism,

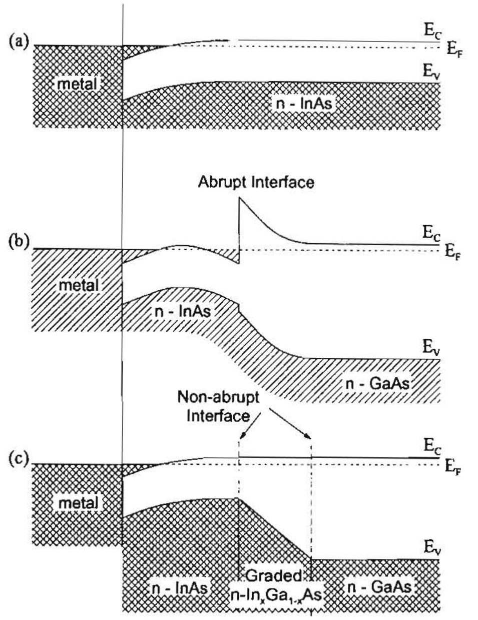


Fig. 3.14 Energy band diagrams for low barrier ohmic contacts : (a) metal/n-InAs; (b) metal/n-InAs/n-GaAs; (c) metal/n-InAs/graded n-Ga$_{1-x}$ In$_x$As/n-GaAs [46].

thermionic-field emission (TFE), is applicable for intermediate doped semiconductors (between ~ 10^{17} and ~ 10^{18}cm^{-3}). In this situation both the thermionic emission and tunneling through the upper part of the barrier are important for the carrier conduction. The third mechanism is the field-emission (FE), effective in heavily doped semiconductors (> 10^{18} cm^{-3}). Because the depletion region is narrow, the electrons tunnel through the whole barrier. The dominant mechanism of the current flow depends primarily on temperature, barrier height, doping concentration profile, charge carriers effective mass and dielectric constant. Besides, several other factors, such as the semiconductor interfacial layers or the stoichiometry of the semiconductor surface, influence the transport mechanism.

The electrical properties of the ohmic contacts are characterized by their specific resistance r_c defined as

$$r_c = \left(\frac{\partial J}{\partial V}\right)^{-1}_{V=0} \tag{3.5}$$

or

$$r_c = \lim_{\Delta S \to 0} R_c \Delta S, \qquad [\Omega \text{cm}^2], \tag{3.6}$$

where R_c is the total contact resistance and S is the contact area. Theoretical expressions of the specific contact resistance can be obtained for each dominant mechanism, starting from the corresponding theoretical I-V characteristics [47,48]. Thus, one can show that r_c is determined predominantly, in each case, by the following factors:

$$r_c \approx \exp\left[\frac{\phi_{bn}}{kT}\right] \quad \text{for TE;} \tag{3.7}$$

$$r_c \approx \exp\left[\frac{\phi_{bn}}{E_{00}\coth\left(\frac{E_{00}}{kT}\right)}\right] \quad \text{for TFE,} \tag{3.8}$$

and

$$r_c \approx \exp\left[\frac{\phi_{bn}}{E_{00}}\right] \quad \text{for FE,} \tag{3.9}$$

where k is the Boltzman's constant, T temperature and E_{00} is the tunneling parameter defined by

$$E_{00} = \frac{qh}{4\pi}\sqrt{\frac{N_{D,A}}{\varepsilon m^*}}\ . \tag{3.10}$$

In eq. (3.10) q is the electronic charge, h is the Plank's constant, $N_{D,A}$ is the dopant concentration, ε is the dielectric constant of the semiconductor, and m^* is the electron effective mass. E_{00} is a very useful parameter in predicting the blocking or ohmic characteristics of a metal-semiconductor contact. For $kT/E_{00} >> 1$ the TE mechanism dominates and, in general, the contact behaves as a Schottky barrier (nonohmic). For $kT/E_{00} << 1$ the FE is effective and the contact exhibits ohmic characteristics. In the range where $kT/E_{00} \approx 1$ a mixed mode of transport occurs. The current transport mechanisms and the expressions for the specific contact resistance suggest the same possible ways for achieving ohmic contacts on GaAs as those ones sketched formerly when we have discussed the metal-semiconductor junction.

Our metallization scheme for n-type contact consists of three metal layers - 20nm germanium (Ge), 15nm nickel (Ni) and 200nm gold (Au) - deposited in this sequence on n^+-GaAs with doping around $2x10^{18}$ cm^{-3}. The deposition has been accomplished in high vacuum (base pressure ~ $3x10^{-8}$ mbar) using an e-beam process (Leybold L560 UV equipment). Then, the metallization has been annealed using a rapid thermal annealer (RTA) in hydrogen and nitrogen atmosphere, at 400°C for 60 s (see Appendix B). This metallization process originates from Duisburg University [49] and [50], but it was optimized in our laboratory conditions, concerning the annealing temperature and time [23]. The metallization data presented above situate, apparently, our scheme among the numerous variants of the most commonly used contact material for n-GaAs, eutectic AuGe (12% wt. or 27% at. Ge)-Ni. We say "apparently" because, for the explanation of its metallurgical behaviour, some elements as noneutectic composition between Au and Ge, the deposition sequence, the atom ratio Ni-Ge, and the RTA procedure, make us able to look beyond the actual understanding of the eutectic AuGe-Ni contact formation and to consider also the image offered by the "new" generation of contacts, based on the solid-phase regrowth of compound semiconductors. So, before explaining the contact formation in the case of our metallization scheme we shall consider the closest and most representative example from each of the two main groups of n-type contacts, alloyed and based on solid-phase regrowth.

The alloyed AuGe based contacts evolved basically from the metallization scheme using single eutectic AuGe to that with eutectic AuGe with some content of Ni [51] or an overlay of Ni [52,53] and, finally, to eutectic AuGe sandwiched between two Ni layers [54]. In the last two cases also a Au capping layer has been used to improve more the electrical properties of the contact [55] and for measuring and wiring reasons. The mechanism of contact formation using single eutectic AuGe was believed to be relative simple: during heating AuGe eutectic is molten and some of GaAs substrate is dissolved in the melt; on cooling Ge segregates from the melt together with the solidifying semiconductor. In this way a high doped layer appears at the interface contact-GaAs, forming a "tunneling" type contact. In reality the "alloying" process is more complex, the contact tending to "ball up" and become non-uniform both laterally and in depth. The Ni deposition was included originally to act as a wetting agent and prevent the AuGe metal from "balling up" during alloy [51, 52]. The investigation of the contact structure with cross sectional transmission electron microscopy (TEM) and modern methods of surface analysis, as microprobe Auger electron spectroscopy (AES), secondary ion mass spectrometry (SIMS), x-ray energy dispersive spectrometry (EDS) microanalysis and Rutherford backscattering spectroscopy (RBS) revealed, however, that Ni plays an active role in the metallurgical mechanism of contact formation [53, 56]. We illustrate this with the optimized metallization scheme described in [57]. The metallization consists of a Ni(5nm)/AuGe(27% at. Ge)(100nm)/Ni(35nm)/Au(50nm) sandwich. The evolution of the contact microstructure under heat treatment is shown in Fig. 3.15. At the initial stage, the Ni layer adjacent to semiconductor reacts with GaAs to form Ni_xGaAs ternary phase whereas Ge diffuses from the AuGe layer to the upper Ni layer, giving Ni_xGe compounds. Later, the Ni and Ge migrate to the bottom Ni_xGaAs phase. Finally, Ga reacts with Au to form β-AuGa phase and Ge replaces Ga in the Ni_xGaAs phase, resulting in the formation of the NiAs(Ge) grains at the interface. The entire evolution of the metallurgical process described above is sustained by many facts. Thus, Ni has a high reactivity with GaAs starting to react with it above 200 °C [58]. Also, it is known that Ni, adjacent to a AuGe layer, acts as a sink

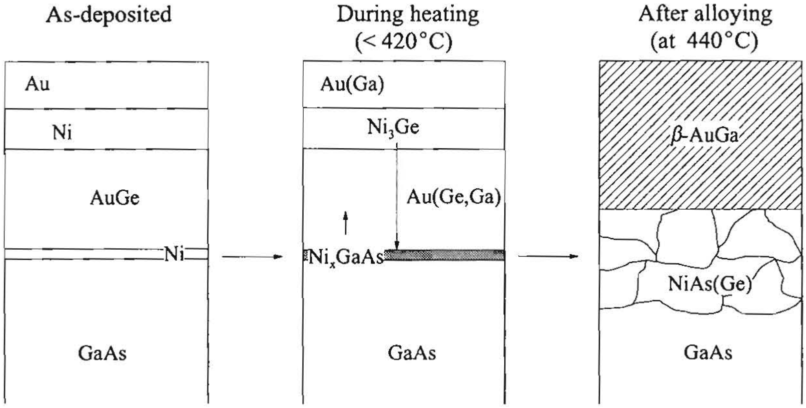


Fig. 3.15 Schematic illustration of the Ni(5nm)/AuGe(100nm)/Ni(35nm)/Au(50nm) metallization on n^+-GaAs at various stages of the alloying process [57].

for Ge and extracts it at temperature higher than 250 °C to form different compounds [59]. The solid state diffusion of Ge into the Ni layer results in the destruction of the eutectic composition throughout most of the original AuGe layer before to reach its melting point [60]. In this sense one can assert that, at least in the stage of contact formation, the metallurgical reactions proceed basically between solid-phases and the melting is avoided when the Ni is present. This fact and the presumed initial smooth interface between Ni_xGaAs layer and GaAs explain the improved uniformity of this contact in comparison with that based on single eutectic AuGe. The NiAs(Ge)/GaAs interface is considered to play the dominant role in the ohmic behaviour of the contact by the formation of a Ge highly doped layer (n^{++}) underneath the NiAs(Ge) grains.

The second metallization scheme we discuss is based on solid state interactions. It consists of a Ni(40nm)/(Au(5nm))/Ge(100nm) layer sequence [61]. The middle Au layer is very thin and is not absolutely necessary for the ohmic behaviour of the contact [62] but it decreases appreciably the specific contact resistance. The mechanism of contact formation involves the solid-phase regrowth of a heavy doped GaAs layer by reaction-driven decomposition of an intermediate ternary phase, Ni_xGaAs, developed at the contact interface. Instead of Ni/Ge system other combinations can be used, as Ni/Si, Pd/Ge , Pd/Si, Pd/In or Pt/In [63]. During annealing the contact microstructure transforms as seen in Fig. 3.16. At the initial stages of heating, ~ 200 °C, the Ni reacts with GaAs resulting ternary Ni_xGaAs compounds:

$$x\text{Ni} + \text{GaAs} \rightarrow \text{Ni}_x\text{GaAs}. \tag{3.11}$$

It is believed that a small fraction of Ge would be incorporated in the Ni_xGaAs layer, because of the strong binding energy between Ni and Ge. After annealing at 300 °C most of the Ni was included in the Ni_xGaAs phase. Above 300 °C, Ge starts to react with Ni_xGaAs, the reaction

$$x\text{Ge} + \text{Ni}_x\text{GaAs} \rightarrow x\text{NiGe} + \text{GaAs} \quad \text{(epitaxial regrowth)} \tag{3.12}$$

being driven to the right by free energy considerations [63]. During this reaction some of the Ge atoms should remain in the regrown GaAs. The reaction continues and at about 400 °C, Ga in the

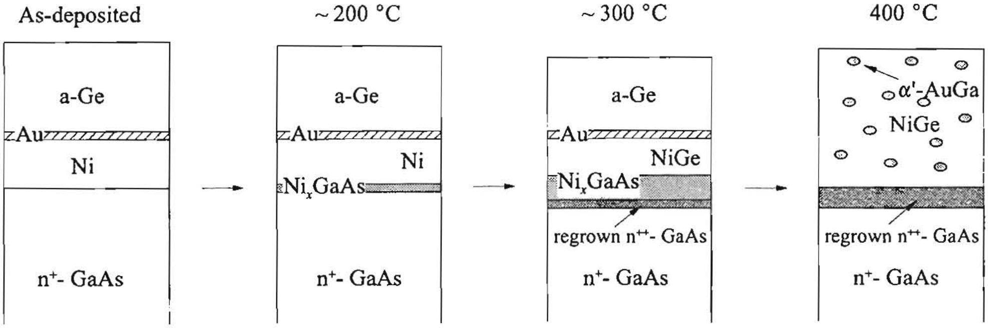


Fig. 3.16 Schematic illustration of contact formation by solid-phase regrowth mechanism for Ni(40nm)/(Au(5nm))/Ge(100nm) metallization on n^+-GaAs [61].

Ni_xGaAs starts to react with Au to form α'-AuGa (orβ'-AuGa). The Ga vacancies in the regrown GaAs layer are then occupied by Ge, resulting in the attainment of a very high doping level. An excess of Ni atoms versus Ge atoms in the initial layers determines the formation of NiAs grains in the final NiGe layer. The As vacancies left behind, in the GaAs regrown layer, can be occupied by Ge and finally, the n-doping in this layer will decrease (compensation) degrading the electrical properties of the contact.

The two examples presented above now helps us to discuss our metallization scheme. We still make some specifications. First, the annealing plateau at 400°C for 60s is preceded by a temperature ramp, starting from 100°C, with a total duration of 10 s. Then, the Ni and Ge layer thicknesses indicate a higher number of Ni atoms than Ge atoms (the ratio is 1.37). At last, it is known that Ge is stable on GaAs, with very little interdiffusion between them, up to 450°C [64] and it reacts with Ni to form Ni_2Ge between 150-300°C and NiGe between 250-600°C [65]. In this conditions, at the initial stages of annealing, above 200°C, in our metallization Ge reacts with Ni to form Ni_2Ge (see Fig. 3.17). In the same time, the Ni, with its high affinity for GaAs, starts to diffuse through the Ge layer to reach the GaAs substrate. Consequently, at the interface metallization-semiconductor, the Ni_xGaAs ternary phase develops. These reactions continue until above 300°C when Ni_2Ge transforms in NiGe, the remaining Ni combining with unreacted Ge and of course with the substrate. At this temperature Ge starts to react with Ni_xGaAs phase as well, in concordance with the reaction expressed by eq. 3.12. When the annealing temperature reaches the 400°C plateau, we suppose that the contact microstructure looks like in Fig. 3.17d. We expect that before 400°C the Au layer is relatively stable [61], the diffusion of the Au atoms towards the GaAs substrate and, conversely, of the Ga atoms into the Au layer being relatively small. During annealing at 400°C, as seen in Fig. 3.17e, the whole Ni_xGaAs ternary phase is "consumed", leading to the increase of the regrown GaAs layer and the formation of NiAs. The Ga resulted at the NiAs phase formation reacts with the Au to give the β-AuGa phases. Also, some Ga from the regrown layer undergoes the same reaction, the Ga vacancies being occupied by Ge atoms which enhances in this way the doping of this layer. The increase of the annealing temperature or time will result in the continual development of the β-AuGa and NiAs phases,

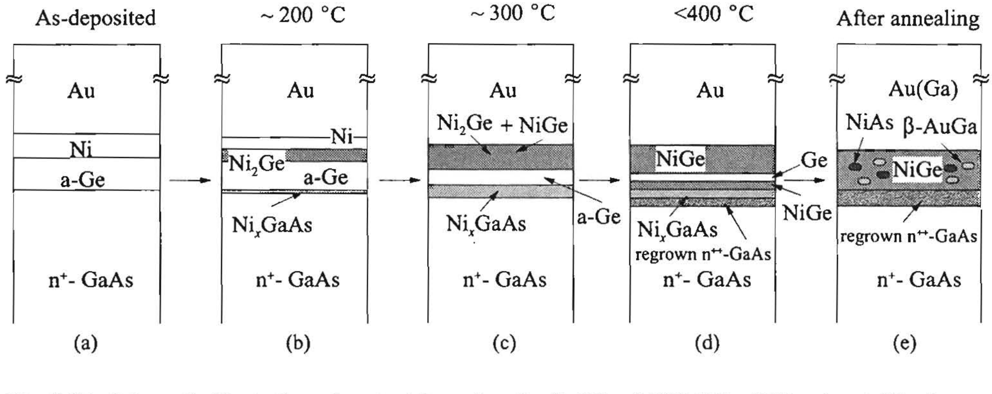


Fig. 3.17 Schematic illustration of contact formation for Ge(20nm)/Ni(15)/Au(200nm) metallization on n^+-GaAs, used for the n-type contacts of the DBRT devices.

which, first, degrade the contact morphology (β-AuGa melts at 370 °C) [23] and then affects its uniformity and electrical characteristics. The NiGe and NiAs phases in our description correspond in one way to the NiAs(Ge) phases noticed in [56] and [57].

The use of the RTA process has a benefic influence on the contact performances. The rapid and controlled heating process determines especially the proceeding of the solid-phase reactions at the interface. Also, this annealing procedure can reduce the phase segregation, the Au diffusion to the GaAs surface and the growth of the undesirable phases, as β-AuGa and NiAs, by grain boundary diffusion. As a result the RTA process leads not only to low contact resistance but also to better uniformity, improved surface morphology, better edge definition and less penetration of the metallization into the semiconductor [53, 66, 67].

In connection with the degree of uniformity of the contact is also the existence of the GaAs native oxide. Except when an in-situ cleaning of the GaAs surface is performed prior the metal deposition, a 1-5 nm thick oxide has to be expected on the semiconductor surface, depending upon its preparation [68]. This oxide layer is probably sufficient to retard or even inhibit the nucleation of the Ni_xGaAs phase, the reaction starting first in the "weak" oxide areas [68,69].

In conclusion, concerning our metallization scheme, we believe that with optimum deposition and annealing conditions the uniformity and the penetration of the contact into the GaAs can be acceptable, even for the micron devices, because of the uniform Ni_xGaAs phase formation and the regrown of the very thin n^+-GaAs layer. The excess of Au, however, makes the latitude of the contact processing very narrow and, more, its thermal stability very fragile. For our proposals, however, the low contact resistance demonstrated (see chapter 4) made from this metallization scheme an alternative choice.

The second metallization, used for the base contacting of the three terminal DBRT devices, has been Ti/Pt/Au. This type of metallization is known as giving a good p-type contact on p^+-GaAs [70] or a performance Schottky barrier for the MESFETs or HEMTs [23, 46]. We have chosen it for a relative high thermal stability and low penetration into the semiconductor. Because the metallization is deposited on the undoped GaAs material of the quantum well and it is not annealed, the contact exhibits a rectifying electrical behaviour (see chapter 5).

3.3 Specific Processes

3.3.1 Thin Film Dielectric Deposition and Etching using Plasma Systems

Dielectric layers, as silicon oxide (SiO_2) and silicon nitride (Si_3N_4) are used in the GaAs based device processing particularly for interlevel dielectrics, protection from chemical and mechanical harm or as mask during different process steps. Silicon nitride is a better diffusion barrier than silicon oxide and therefore is superior for encapsulation (environmental sealing or glassivation). Also, Si_3N_4 is preferred for capacitor dielectrics because it has a higher dielectric constant than SiO_2. The lower dielectric constant of SiO_2 makes it suitable as a general interlevel dielectric, to separate metal crossovers and pads with minimum capacitance. High quality dielectric layers can be grown by *chemical vapour deposition* (CVD). In this technique the reactant gaseous species flow in the vicinity of the heated substrate and appropriate reactions take place to grow the material on the substrate surface. The temperatures required to drive these reactions for the obtaining of resonable deposition rates and good film properties are very high, often 700 °C to 1000 °C. GaAs and in general compound semiconductor based devices cannot stand such high temperatures. Not only does arsenic evaporate in this temperature range, but metals (contacts), commonly present on the sample from the previous processing steps cannot be exposed to these extremes. Therefore CVD is not a useful technique in the processing of GaAs based devices. Plasmas are used instead, to generate chemically reactive species at relatively low temperature. Plasma-assisted deposition processes are denoted as *plasma-enhanced chemical vapor deposition* (PECVD). For semiconductor device processing low pressure plasmas or glow discharges are used. These plasmas are characterized by gas pressures in the range of 0.1 torr to several torr, electron densities of 10^9 to 10^{12} cm^{-3}, and electron energies of 1 to 10 eV. The mechanism of PECVD process is based on the nonequilibrium nature of the plasma state, in which the chemical species (atoms, molecules or radicals) with temperatures near the ambient coexist with high energetic electrons that can exhibit temperatures of thousands of degree Kelvin. The electron energy is sufficient to break molecular bonds and create chemically active species in the plasma. Moreover, any of these species can be excited to higher electronic energy states by further interactions with electrons. Consequently, chemical reactions that usually occur at high temperature can be made to occur at low temperature in the presence of an activating plasma state. Even in PECVD it is preferable to heat the substrate in order to increase the deposition rate, to reduce the incorporation of contaminants and to improve the quality of the deposited layers. The plasma conditions result in creation of virtually all possible radicals from the incoming gases, which may participate in the deposition of the dielectric materials. Electrically most of the species in glow discharges remain neutral. With the exception of microwave electron cyclotron resonance (ECR) plasma's (see below), the degree of ionization (the ratio of the ion and neutral species) is of the order of 10^{-4} -10^{-5} [31]. Although the gas ionization is small, it is adequate to provide sufficient numbers of reactive species. The multitude of chemical species and the nature of the plasma state make processing results extremely sensitive to all possible parameters: gas type, gas flow, gas delivery position, pressure, electrode geometry, power, radio (microwave) frequency, wafer temperature, and wafer material. Unlike CVD, in the case of the PECVD the compositional control of the thin film material is difficult. CVD growth results generally in near stoichiometric composition of the deposited material. PECVD yields films that are amorphous in nature, with very little short-range structural ordering and in which the chemical bonding may vary. Moreover, in the plasma-assisted deposition process, chemical species other than the desired ones are often included in the film.

In many situations, after the deposition of the thin dielectric films, the device processing involves the opening of windows in these materials. As in the case of semiconductors, the dielectrics can be etched using solutions (wet etching) or plasmas (dry etching). Plasma etching of dielectrics has several advantages over wet etching. The most important are the greater etch control and the etch anisotropy, manifested by a more rapidly etching in the vertical direction (normal to the material surface) than in the horizontal one. The last feature permits the reduction of the undercutting to nearly zero, which in some processing situations is an indispensable condition. Unlike the plasma deposition, the etching process in general uses both the neutral reactive species (radicals) and positive ions created in the plasma. The process is denoted as *reactive ion etching* (RIE), being performed at even lower pressures than the deposition process, typically between 0.01 and 0.1 torr. The neutral radicals form the chemical component of the etch process. These are normally adsorbed on the material, reactions take place resulting volatile products and, finally, these desorb being evacuated. The ions are extracted from the plasma and bombard the substrate. The function of the ionic bombardment is to increase desorption of the reaction products and/or to enhance the reaction of the neutral radicals with the substrate material. Also, the energetic ions can remove material by mechanical action. For these reasons they are denoted as physical or sputter component of the process. Using special reactor configuration the ions impinge more or less vertically on the substrate and in this way determine the anisotropic character of the etching.

Both the deposition and etching of SiO_2 have been accomplished in an Oxford Plasma Technology (OPT) cluster tool [23]. The cluster consists of three different process chambers (stations) connected to a general transfer chamber in which the samples are introduced using an additional load lock. In all chambers the plasma is created and sustained remote and independent from the substrate location, in a so called downstream configuration. The advantage of such a configuration is that the damage produced by the ion bombardment is minimized. Two of the three stations, number 3 and 1, are dedicated for the deposition and etching, respectively of SiO_2 and Si_3N_4.

Station 3 is a microwave downstream PECVD reactor. For the deposition of the SiO_2 films, nitrous oxide (N_2O) is used to provide the oxygen and silane (SiH_4) is the supplier of the silicon. Both gases are diluted with nitrogen (N_2). As seen in Fig. 3.18, part of the gases, N_2O diluted with N_2, are introduced through a gas ring in the top of the microwave cavity and flow along a quartz tube. These gases are microwave excited and transformed into the plasma state. The other part of gases, SiH_4 diluted also with N_2, are supplied in the vicinity of the substrate table via a second gas distribution ring. The CF_4/O_2 mixture, which can be introduced in the top of the reactor as

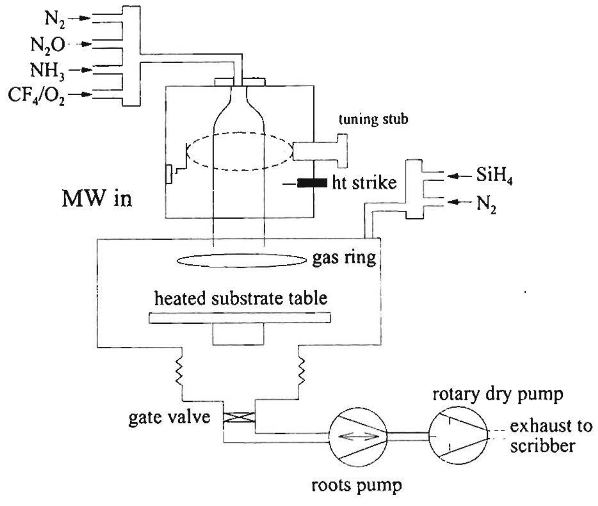


Fig. 3.18 Schematic layout of the microwave downstream PECVD reactor used for the silicon oxide deposition [23].

well, is destined for the chamber cleaning purposes. The flow of all gases is regulated by mass-flow controllers. The process pressure is adjusted by the bottom gate valve and the substrate table can be heated until 400 °C. The microwave power is supplied at a frequency of 2.45 GHz, using an antenna located outside the quartz tube. Opposite to the antenna in the cavity there is a manual tuning stub to reduce the reflected power.

In this reactor configuration the SiO_2 deposition is a four-step process [71]. First the top supplied gases, N_2O and N_2, are microwave excited and reactive species are created in the plasma. The main plasma constituents of the N_2O are N_2 and O_2 molecules and O and NO radicals [72]. Then, in a second step, the different species, in majority neutral, are transported by diffusion and pumping outside the plasma region. There, during the third step, the activated species mix with neutral silane or disilane and gas phase reactions result in the formation of precursor molecules or radicals. NO radicals are relative unreactive [73], so the oxygen atoms play the main role in the generation of the gas phase precursors. Finally these precursors and, if need be, excess atomic oxygen are involved in CVD reactions on the heated substrate resulting in deposited SiO_2 material.

In analogy with the plasma deposition of the organic polymeric materials, it is believed that the gas phase precursors contain the bonding group which is the structural building block in the deposited SiO_2 film [71]. If this is true, the local bonding groups observed in the SiO_2 film can be traced back to precursor species with the same particular structural building block. Thus, the low temperature deposited films contain beside the local bonding groups Si-O-Si, present at all temperatures, the Si-OH structural block. This suggests that one possible precursor species is silanol ($(SiH_3)OH$). Increasing the temperature deposition the film content of Si-OH groups decreases and the general structural block is Si-O-Si linkage. The precursor species responsible for the incorporation of this bonding group is disiloxane ($(SiH_3)_2O$) [74]. Starting from these precursors the surface reaction of the CVD process involves the elimination of hydrogen (H) atom from the SiH_3 group of disiloxane via a reaction in which the H terminal atoms are replaced by bridging oxigen (O) atoms. Such a deposition process reflects two different aspects observed in the plasma polymerization as well: the above-mentioned retention of the local bonding groups found in the precursor molecules and a crosslinking reaction that requires atomic species. In this context the CVD surface reaction necessitates O atoms too and can be written as

$$(SiH_3)_2O + O \rightarrow SiO_2 + H_2, H_2O, \text{etc.} \tag{3.13}$$

The etching of SiO_2 has been performed in the station 1 of the OPT cluster system. This is an *electron cyclotron resonance* (ECR) plasma etch reactor. The configuration of the reactor is shown schematically in Fig. 3.19. The ECR discharge is generated in the upper cavity of the reactor, remote and independent from the substrate stage, situated downstream, in the process chamber. The key elements in producing the ECR discharge [75] are the nonuniform static magnetic field produced by the pair of magnets (magnetic coils) which surrounds the discharge cavity and the microwave power, supplied on top of the reactor using a wave guide. The magnetic field is mirror like. In the absence of the accelerating electric field (microwave power), the motion of an electron in such a magnetic field is well known: it will spiral with ever decreasing transverse orbits into the converging field until it is reflected (see Fig. 3.20a). Then, the electron reverses direction and spirals out of the mirror with increasing orbits. The radii of these orbiting

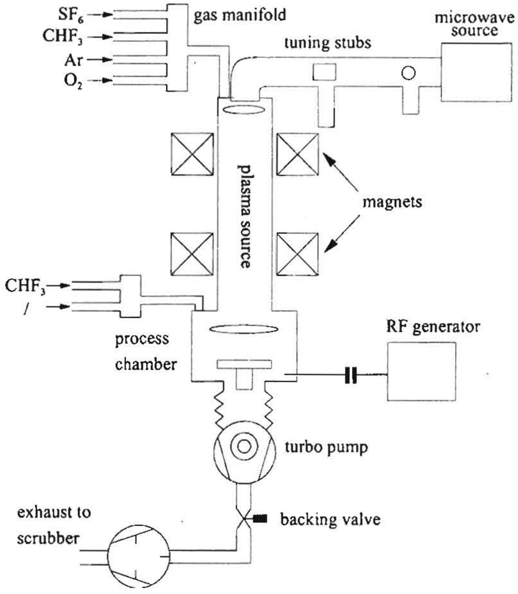


Fig. 3.19 Schematic layout of the ECR plasma etch reactor used for silicon oxide etching [23].

trajectories are small. For example, an electron with 4 eV energy will have a radius of 0.05 mm in a field of 0.0875 T (see below). Now, if the transverse, time varying microwave electric field is present in the mirror a strong acceleration of the electron takes place when it passes through

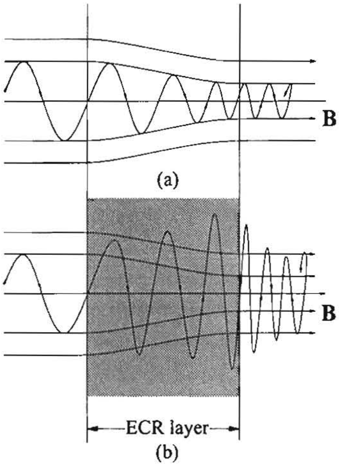


Fig. 3.20 Electron motion in a mirrorlike static magnetic field (a) without an accelerating electric field (b) with a perpendicular electric field $\mathbf{E}_0 e^{j\omega_{ce} t}$ where the shaded region represents the ECR layer [75].

a region where the microwave excitation frequency ω equals the electron cyclotron frequency ω_{ce} (see Fig. 3.20b). This is in fact the electron cyclotron resonance condition which reflects the optimum coupling of the electron to the microwave energy. The condition can be written quantitatively as

$$\omega_{ce} = \frac{eB}{m_e} = \omega. \tag{3.14}$$

where e is the electron charge, m_e is the mass of electron and B is the strength (induction) of the static magnetic field. Since the used microwave excitation frequency is 2.45 GHz eq. 3.14 gives a value for B of 0.0875 T. The accelerating region is very thin (<1mm) and referred to as an ECR layer. If the electron is outside of this region, little microwave energy is coupled to it. However, the electron can experience in principle many oscillations during the time it spends in the ECR layer. Consequently, if there are less (no) collisions, considerable energy can be imparted to electrons in this layer with relatively small electric fields (10-100V/cm) [75]. Upon leaving the ECR zone the accelerated electrons possess energies usually greater than 10 eV. They move throughout the discharge volume and excite, dissociate and ionize neutrals and ions. From the above, it is clear that ECR is a coupling technique for low-pressure discharges where the electrons can orbit many times between elastic and inelastic collisions. Depending on used gases (variable gas mixtures) and flow rates, the ECR discharges can operate over a relatively wide range of low pressure, 10^{-5} Torr to over 100 mTorr [75]. At these low pressures collisional energy coupling through the discharge volume can be ignored. One can say that the ECR layer serves the function of a hot cathode filament in dc discharges, providing the high-energy electrons required to sustain the discharge. The plasma generated using ECR phenomenon can be varied from weakly to highly ionized state by changing discharge pressure, gas flow rates, and input microwave power.

As in the case of the PECVD reactor, it is possible to introduce gas mixtures at two different levels. Thus, there is a top gas inlet ring for the access of sulfur hexafluoride (SF_6), trifluoromethan (Freon® 23)(CHF_3), argon (Ar) and O_2, and a bottom one, in the process chamber, which make possible the use of CHF_3 in the vicinity of the substrate. The gas flow is regulated by mass flow controllers. The bottom valve and an automatic pressure controller ensure the adjusting of the process pressure. The substrate table position can be changed with respect to the plasma region. Also, a heating/chilling system establishes and maintains the desired table temperature. On the substrate table additional RF power, at a frequency of 13.56 MHZ, can be supplied to obtain a certain DC bias.

For SiO_2 etching we have used a mixture of SF_6, CHF_3 and Ar introduced through the top gas inlet ring. All the gases traverse the ECR discharge zone and dissociate. The plasma chemistry is extremely complex and not entirely established, especially for SF_6 [76, 77]. Both SF_6 and CHF_3 supply species which are involved in the SiO_2 etching [76,78]. Thus, SF_6 dissociates in fluorine (F) atoms and fluorosulfur ($S_x F_y$) radicals. Atomic F is an active etchant for SiO_2, the reaction which takes place resulting in the formation of volatile silicon fluoride (SiF_4) and O_2:

$$SiO_2 + 4F \rightarrow SiF_4 + O_2. \tag{3.15}$$

Since the etching process in this case is of chemical nature it has an isotropic character. Concerning CHF_3, the low working pressures promote the formation of fluorocarbon radicals

(CF_x, $x \leq 3$) in the plasma. The CF_x radicals does not etch spontaneously SiO_2, the etching process being more complex and involving the ionic component of the plasma as well. Thus, it is believed that the unsaturated CF_x species which impinge on the oxide surface form and maintain continuously a thin fluorocarbon layer ($\leq$ 10 Å). From this layer halogen atoms are transferred to the substrate and by ion-induced reactions gaseous species are produced, resulting finally in the SiO_2 etching [59]. The entire process can be expressed by the following general reaction:

$$CF_x + SiO_2 \rightarrow SiF_4 + (CO, CO_2, COF_2). \tag{3.16}$$

The involving of ions determines this etching process to be anisotropic.

The etching mechanism depends on the balance between the main reactive plasma species, F atoms and CF_x radicals, which can be controlled by changing the flows of the generating feed gases, SF_6 and CHF_3. Besides the role which it has in the second etching channel mentioned above, Ar is important to stabilize plasma and to reduce the etch rate by dilution [76]. Finally, we point out that the etching process stops when the GaAs surface is reached. This selectivity is caused by the involatility of the gallium fluoride (GaF_3) formed on the semiconductor surface and/or the formation of a fairly thick fluorocarbon layer (as in the case of Si [76] there is no reaction with GaAs that can gasify the polymer film) which prevents attack by sputtering or chemical etching. The parameters of the SiO_2 dry etching process used for our DBRT device fabrication are presented in the Appendix B.

3.3.2 Selective Etching

In the processing of the new generation of electronic and opto-electronic devices, from III-V compound semiconductors the selective etching represents a very helpful tool. The procedure is based on the specificity manifested by some etchants against the material nature.

For some quantum devices, e.g. the three terminal DBRT devices described in chapter 5, the etch depth control during their fabrication must be less than the thickness of the 'active' layers, which very often are in the range of several nanometres. Also, the etch uniformity is very important for the performances of such devices and this requirement imposes even higher control [79]. The use of etchants with very low etch rates could be, in principle, a solution in these cases, but there is another aspect. All the quantum devices have the 'active ' layers embedded in specific cladding layers for contacting. These are, usually, much thicker than the functional structure. Consequently, it is unrealistic to consider the use of low rate etchants because of the too long etch time and consequences which devolve from that, such as etching mask destruction, strong underetch, etc.. In this way, a highly selective etch that removes one material much faster than another can resolve the problem in a satisfactory way. Therefore, in the last period much interest has been paid to develop highly selective etch processes, using both wet and dry etchants. The most studied combination in this sense has been GaAs/$Al_xGa_{1-x}As$, which constitutes also the system used in the fabrication of our resonant tunneling devices. Accordingly we refer in the following only to this system. Here, we specify that the problem of the GaAs/$Al_xGa_{1-x}As$ selective etching, as of any other combination of materials, presents two aspects in conjunction with the symmetry of the situation. One can speak about the selective etching of GaAs over $Al_xGa_{1-x}As$ and conversely of $Al_xGa_{1-x}As$ over GaAs. In general, the device processing necessitates the use of the GaAs over $Al_xGa_{1-x}As$ selective etching but there are also situations when the opposite selective process could help. This circumstance explains the much higher interest in finding etchants which concerns the first aspect. As we shall see, in the processing of our three terminal resonant tunneling devices both aspects have been involved, so the entire problem of the GaAs/$Al_xGa_{1-x}As$ selective etching is presented.

Selective etchants are classified after their selectivity. The quantitative expression of this property is given by the ratio of the etching rates of the faster removed material (GaAs or $Al_xGa_{1-x}As$) to the slower removed one ($Al_xGa_{1-x}As$ or GaAs).

Initially, the selective wet etching of GaAs over $Al_xGa_{1-x}As$ has predominantly utilized $NH_4OH{:}H_2O_2$ solution. The selectivity of this solution has been improved by solution agitating [80], by providing a steady jet stream of the etching solution [81], or by adjusting its pH values [82]. However, the selectivity values reported with $NH_4OH{:}H_2O_2$, less than 30, are too low to permit the etch control at the scale of the new quantum devices. Moreover, this solution attacks the photoresist mask, requiring dielectrics for masking during etching. A solution that has demonstrated much higher selectivity and no resist attack is citric acid ($C_6H_8O_7$):H_2O_2. Many investigations using this solution have shown that, in specific concentration of the two constituents, the selectivity is dependent on Al mole fraction and varies from over 100 for GaAs/$Al_{0.3}Ga_{0.7}As$ [79, 83, 84, 85] to around 1500 for GaAs/AlAs [79]. These very high values permit a good control of the etch depth and uniformity.

Dry etching offers also several possibilities for GaAs over $Al_xGa_{1-x}As$ selective etch processes. Selective reactive ion etching in Freon® 12 (CCl_2F_2)-based plasmas [86,87,88] and silicon tetrachloride ($SiCl_4$)/silicon tetrafluoride (SiF_4) or SF_6-based plasmas [79, 89, (90, 91)] have become alternatives of choice for the processing of many quantum devices, due to the high selectivities (> 200) achievable, also dependent on the Al mole fraction.

Concerning the selective etching of $Al_xGa_{1-x}As$ over GaAs only wet etching technique offers a solution. Thus, for $x > 0.5$ hydrogen fluoride (HF) manifests a very high selectivity ($\gtrsim 10^7$, for $x = 1$) [92], being used in the epitaxial lift-off experiments.

In the following we discuss the selective etching processes that have been used for the fabrication of our three terminal resonant tunneling devices. First, the selective characteristics of the $C_6H_8O_7/H_2O_2$ solution are described as they result from our experiments. In the same part, a short analysis of the selective removing of very thin AlAs layers, as components of DBRT structures, is inserted. Then, the main aspects of the selective dry etching using CCl_2F_2 are presented. This procedure has been investigated and developed at Duisburg University, Semiconductor Devices Department, but we have used it in our device processing [98].

a) Selective Wet Etching

As mentioned above, $C_6H_8O_7{:}H_2O_2$ solution at specific volume ratios manifests very high selectivity for GaAs over $Al_xGa_{1-x}As$. The etch characterization of this solution differs, to a certain extent, between sources from literature [79, 83, 84, 85]. For this reason and, of course, for the establishing of the selective properties in the conditions of our laboratory, we have done detailed etch tests. We focused on the determination of the etch characteristics and selectivity for the GaAs/AlAs system, involved in the structure of our resonant tunneling devices. All the etch tests have been realized at room temperature, with low stirring. We used photoresist masks (AZ 1350J) and measured the etch depth with a Tencor Alphastep. Details concerning the $C_6H_8O_7/H_2O_2$ solution preparation are presented in Appendix B.

In the first stage of our experiments we have established the etch rates of GaAs for different volume ratios (r_v) of the $C_6H_8O_7$:H_2O_2 solution. Because the resonant tunneling structures have the largest part of the top cladding layers made from n^+-GaAs we have used in this tests n^+-GaAs substrate wafers. Before each etch operation in $C_6H_8O_7$:H_2O_2 solution, the native oxide has been removed from the surface of the GaAs sample, using HCl:H_2O (1:4). Without this step there is an uncontrollable delay of the normal etch start in the $C_6H_8O_7$:H_2O_2 solution. Fig. 3.21 shows the results. These are very similar to those reported in [79]. The etch mechanism is reaction rate limited for $r_v > 2$ and diffusion rate limited for $r_v < 2$ [93]. In the inset of Fig. 3.21 the linear time dependence of the etch depth demonstrates the reaction-rate limited behaviour of the $C_6H_8O_7$:H_2O_2 solution, for $r_v = 4$.

For AlAs, the etch rate measurement of any solution is complicated by the fact that the surface of this material is not stable in the atmosphere. Also, it is known that for $3 < r_v < 10$ of the $C_6H_8O_7$:H_2O_2 solution, there is a significant decrease of the etch rates of $Al_xGa_{1-x}As$ in comparison with GaAs when x increases from 0.3 to 0.45 [79]. So a very low etch rate of the AlAs in the same volume ratio range was expected. For these reasons the etch rate of AlAs and than the selectivity for GaAs over AlAs have been established indirectly, only for one composition of the $C_6H_8O_7$:H_2O_2 solution . In this sense, we have used a particular structure with an AlAs stop layer of 3 nm (see Fig. 3.22). The layer sequence on top of the AlAs layer reflects the structure used for our resonant tunneling devices, in front of the first barrier, but with a much thicker undoped GaAs layer. Also, we have chosen a $C_6H_8O_7$:H_2O_2 solution with $r_v = 4$. The reason has been dictated by the highest selectivity demonstrated by this solution composition for

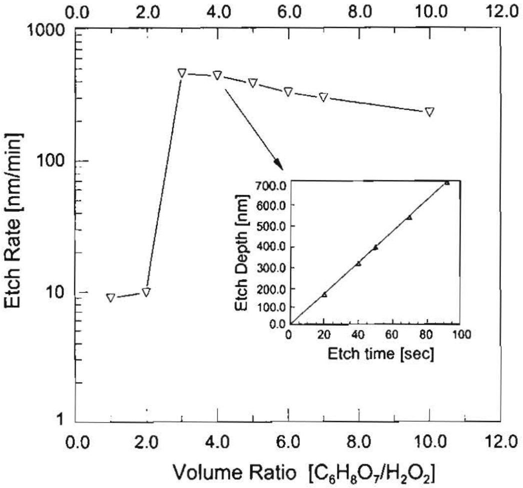


Fig. 3.21 Etch rate of GaAs as a function of $C_6H_8O_7/H_2O_2$ volume ratio at room temperature (21 °C). The inset shows the etch depth versus etching time for GaAs in 4:1 $C_6H_8O_7/H_2O_2$ solution.

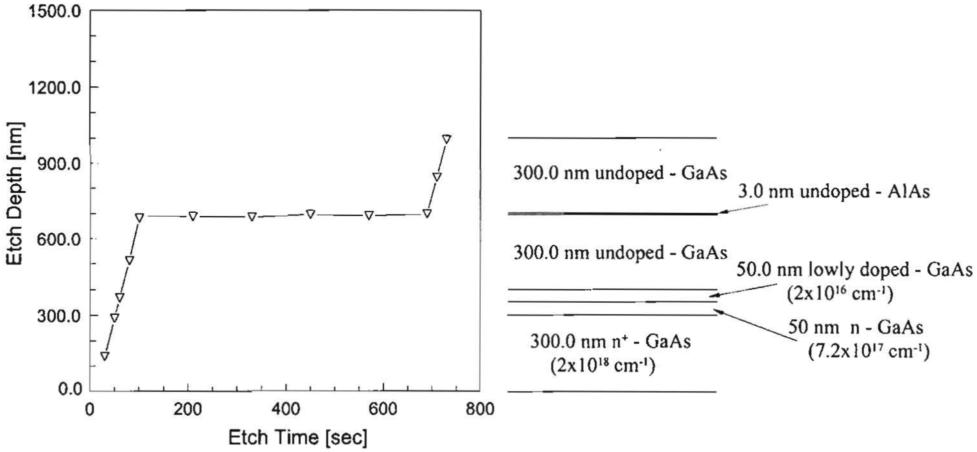


Fig. 3.22 Etch depth versus time for the test structure sketched in the right side of the graph using a solution of 4:1 $C_6H_8O_7/H_2O_2$.

GaAs over $Al_{0.45}Ga_{0.65}As$ [79]. Fig. 3.22 shows the etch depth as a function of the etch time for the test structure mentioned above. As observed, the GaAs layers were etched rapidly. After this, the etch rate decreased precipitously, when the AlAs layer was encountered. Immediately the AlAs was etched through to expose the lower GaAs layer, the etch rate increased again. It should be noted that it took over 10 min to etch through only 3 nm-thick AlAs layer. This translates to an etch rate of less than 0.3 nm/min and a selectivity for GaAs over AlAs of around 1500. Also this experiment has shown that there is no visible difference in etching of doped and undoped GaAs, with the $C_6H_8O_7$:H_2O_2 (4:1) solution.

In our three terminal DBRT device fabrication we also had to remove selectively the first AlAs barrier of the structure. As mentioned above HF can be used for selective etching of AlAs over GaAs. In epitaxial lift-off experiments [92] it was shown that the selectivity in concentrated HF (48%) can be higher than 10^7. In these experiments the etch rate of AlAs has been reported as being as high as 1 mm/hour, depending on HF concentration and temperature. Our processing conditions are different than those of the epitaxial lift-off process. Thus, if in our case the etch is accomplished in the usual way, vertically through resist opened patterns, in the epitaxial lift-off the etching proceeds in the horizontal plan, the solution using an as narrow access space as the thickness of the AlAs layer. So we expect that the AlAs etch rate is considerably higher in our case. Because the thickness of the AlAs barrier is very small (2.5 nm) and the underlying GaAs layer (quantum well) is also thin (5.0 nm) we have used a very diluted HF solution and the etching process takes place at low temperature and for a very short time. In this way we have avoided any risc of the GaAs layer penetration. Details concerning this selective process are found in the Appendix B.

b) Selective Dry Etching

Selective RIE of GaAs to AlAs, using pure CCl_2F_2 etching gas, has been successfully tested and then used for three terminal DBRT device fabrication at Duisburg University, Semiconductor Devices Group. The work was performed in a conventional planar RIE system, developed by Oxford Plasma Technology, using RF power at 13.56 MHz to ignite and maintain the plasma.

In the etching process of GaAs the main role is played by chlorine (Cl) atoms which forms by ion-induced reactions volatile products with both Ga and As, gallium chloride ($GaCl_3$) and arsenic chloride ($AsCl_3$). Removal of $GaCl_3$ is slower than $AsCl_3$ (high vapour pressure) and in pure chlorine plasmas will be the limiting step of the etching rate [76]. Besides Cl atoms, other species contained in the plasma, as CF_x radicals, are involved in the GaAs etching [94, 95]. As mentioned above, when we discussed about SiO_2 dry etching, halocarbon radicals can deposit on the substrate surface to form a polymer thin layer [86, 94]. In the present case this layer is continually removed by ion bombardment but it slows down significantly the etch rate in comparison with the etching process using pure chlorine plasmas.

Concerning AlAs, the very low etching rate in CCl_2F_2 discharges is caused by the formation of nonvolatile AlF_3, with a high sublimation temperature of 1291 °C [86]. Although we did not measure exactly the AlAs etch rate, establishing only the evident role of etch stopping played by a very thin AlAs layer (2.5nm), we expect that the selectivity in our processing conditions is ~ 450 [87]. The exact description of the process parameters is presented in Appendix B.

3.3.3 Gold Plating and Airbridge Fabrication

Plating in microelectronics serves to many purposes. Thus, it may be used to create a second level of metallization, which increases metal thickness and improves conductivity, to provide propitious (favourable) surface for wire bonding, to form backside plated heat sinks, or to fabricate bridge interconnections. On GaAs and in general III-V compound based devices the plating metal is almost always gold. Comparative with other metals, suitable for plating and used in the silicon processing, gold has many advantageous properties. Of course, the first of them is that gold plates very well. Then, it has high electrical and thermal conductivity, is easily soldered or welded, is resistant to oxidation , and is ductile. Many of these parameters exhibit degradation as other metals are alloyed with gold. For this reason, the gold used for electronic devices is highly pure, usually 99.99% or better. Obtaining of such a purity and also some other requirements, specific for gold plating in microelectronics, can be accomplished easiest using commercial plating compositions intended for semiconductor use. In this way we have proceeded for our plating necessities in the processing of the DBRT devices. As a main goal, we have used electroplating of gold for air bridge fabrication. In the same time, however, we have obtained thick metal on the device pads which has been benefit for 'on wafer' ac and dc measurements and of course for wire bonding when it was necessary. In this section the gold plating process is described, first, as it was developed in our laboratory. Then, we show how this process has been used for the fabrication of airbridges.

The plating equipment has been assembled in our laboratory. As seen in Fig. 3.23, in

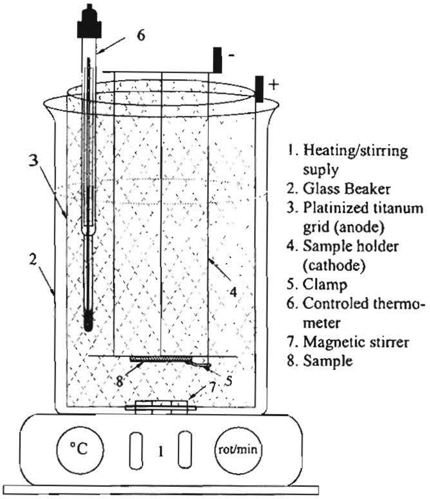


Fig. 3.23 Gold plating setup.

it looks as a classical deposition equipment used in electrochemistry. There are, however, some special requirements, imposed by the nature of plated material (gold) and the characteristics of the plating area. To avoid any contamination the anode material is from platinized titanium. Also, its cylindrical shape ensures a radial symmetry of the electric field and therefore of the current lines. The *plating factor* (exposed area divided by total wafer area) of our samples is, in general, very small. This makes the achievement of the correct current density difficult by simple use of the sample as a cathode. To resolve this problem the sample is mounted on a plate whose exposed area is much larger than the sample plating area and is also plated during the electrodeposition process. In this way the larger exposed area of the cathode determines the current density, almost independent of the variations in the pattern on the sample. The bath is maintained at a constant temperature using a hot plate controlled by the electrical response of a thermometer. The hot plate ensures also a tunable magnetic stirring of the solution. At last, the current is driven by a constant dc current source-monitor. Not figured in Fig. 3.23, a pH-meter and a gravimeter are used to check periodically the corresponding bath characteristics.

The plating bath has been prepared using a PUR-A-GOLD 202 B proprietary composition from Enthone-Omi. This bath, as the most dedicated to semiconductor use, is a mild acidic solution which utilizes for plating the gold cyanide complex ($Au(CN)_2^-$). The source of the cyanide complex is the potassium gold cyanide ($KAu(CN)_2$), and the buffered acidic character is ensured by the orthophosphoric acid (H_3PO_4). We suppose that except these basic components other ingredients are included, such as carrying current species (conductive salts) and levelling agents. The role of the levelling agents is to improve the *throwing power*, which expresses the ability of the plating bath to produce uniform deposits of uniform thickness on cathodes having irregular surfaces.

The $Au(CN)_2^-$ complex is very stable in aqueous solution, with the equilibrium constant of

5×10^{-39} [96]. This means that the Au is tightly bound in the cyanide complex and few free Au^+ ions exist in the solution. Therefore, gold plating does not proceed in the simplified manner of the classical free metal ion electrodeposition. The electrochemical reaction involves the entire complex and can be expressed by

$$Au(CN)_2^- + e^- = Au + 2CN^- . \tag{3.17}$$

Although gold plating has an old and vast utilization, in very different fields, fundamental, well documented studies of the kinetics and electrochemistry of the process are surprisingly sparse.

The thickness and quality of the deposited gold layer is determined by many parameters. First of all, there are of course, the two parameters contained in the Faraday's electrolysis law, the current density and time. Other factors are temperature, agitation, solution gold content, pH, specific gravity, levelling agents, and cathode geometry, position and orientation. The producer of the bath components indicates the values of almost all the above parameters, for an optimum process (see Appendix B). In practice, however, for specific lab conditions and especially plating patterns, starting from the producer specifications the value of some parameters has to be altered.

Our plating experiments have been done on thin gold layers, deposited on a GaAs substrate. The plating pattern, defined in photoresist, was similar with that one used later in the fabrication of the air bridges. As a rule, the photoresist thickness has to be larger than the thickness of the layer to be plated, so we used AZ 4330 resist, also involved in the air bridge processing. Initially, the plating conditions have been generally in the limits indicated in the process specifications: density current $J = 0.15$ A/dm^2, temperature $\theta = 40°C$, pH = 5.7, solution gold content c_{Au} = 10g/l, and agitation $r_a = 80$ rot/min. We made an option for a lower temperature value to avoid too much solution evaporation during the process, our plating bath being open, in an extractive cabinet. In these conditions we obtained an Au plated layer which had a good aspect (surface morphology) but was quite nonuniform in thickness. The lack of uniformity has been observed both at the individual pattern level and along the whole sample. On the individual patterns the plating rate was higher near the resist wall than in the middle, giving rise to the so called "rabbit ears". This phenomena of the "rabbit ears" appearance seemed to be the most important drawback. To explain it and how a uniform plating on our patterns can be obtained, some considerations concerning the transport of the ions and levelling agents at the plating surface have to be done [97]. The local rate of electrodeposition varies with position depending on how many gold complex ions and levelling agents reach the open cathode surface. In general, the different ion species are transported in the solution both by the action of the established electric field (migration) and also by diffusion. When stirring is involved, we speak too about the transport by convection. In the case of the gold cyanide complex anions, because the deposition reaction (eq. 3.17) takes place to the cathode, the diffusion (convection) and migration fluxes are opposite. To obtain deposition the diffusion (convection) must dominate the migration of the gold containing ions. So we consider that in normal plating conditions the prevalent mechanism of transport of the $Au(CN)_2^-$ ions to the cathode electrode is by diffusion (convection). Concerning the levelling agents, it is assumed that they are transported to the plating surface only by (convective) diffusion [97]. There, they are adsorbed and direct or indirect, by their reaction products, act to inhibit the metal deposition. In these conditions we can suppose that both ions and inhibitors move in the solution mainly by diffusion (convection). Now, since a higher concentration of inhibitors is expected at points on the electrode that are exposed to a high diffusive flux, metal deposition is selectively suppressed in this points. Hence, as the surface plating advances, protrusions tend to attenuate and cavities tend to fill. The rather complicated geometry of the resist pattern would result in highly nonuniform diffusive fluxes which make difficult to estimate the evolution of the process. In a simplified picture, however, considering the overcut profile of the resist edge pattern, we can consider that at low agitation levels the diffusion fluxes are lower in the middle of the patterns and higher nearby walls, as predicted in [97]. This determines the "rabbit ears" appearance of the plating profile, indicating a low level of the levelling agents. Explaining the nonuniform plating causes we could discern the way to obtain a better profile layer. Thus, in the quasi-absence of the levelling agents, the modality of equalization of the diffusion fluxes along the pattern surface is the increase of the agitation and/or temperature. And, indeed, experiments performed using the same starting conditions, except the agitation and temperature, which have been settled to higher values, gave much better results. Other improvements, as the use of a disc shape like cathode instead of an rectangular one and a stirrer which gives a more uniform agitation, contributed also to the establishing of the optimum plating process. The conditions of the optimized gold plating process are specified in the Appendix B.

After finishing the gold plating process we could start the fabrication of the airbridge interconnections. In connection with this, we have benefited by collaboration with the Duisburg University [49, 98].

The airbridges have several advantages over the dielectric crossovers extensively used in microwave integrated circuits (MICs). These include low parasitic capacitance, the ability to carry substantial current and immunity to edge profile problems [31]. Our choice for this type of interconnection was determined especially by the first two qualities. Low parasitic capacitance, between the bridge and any metallization beneath, follows from the large spacing and low dielectric constant of the intervening isolator material. Air has a much lower dielectric constant than any other dielectric, and the space under the airbridge is usually considerable higher than the thickness of typical dielectrics (several hundreds of nanometres). In this way the airbridge crossovers are less capacitive than the dielectric type by a factor (typically) of five to twenty [31]. The substantial metal thickness of plated bridges permit device operation at high current density. Also, the risk of shorts between metallizations, at high current densities, is almost inexistent in the case of airbridges.

The fabrication process of airbridges involves many steps (see Fig. 3.24). We have discussed about the photolithographic steps in subsection 3.2.2a. Here, we stress only some processing aspects specific to the airbridge fabrication, contained otherwise in a technical form, in the process recipes presented in Appendix B (processes B.8 and B.9). As seen in Fig. 3.24-1 the process starts with a lithographic step by which windows are opened in the resist mask over the metallization places that necessitate interconnection. Then a thin Au layer (50-100nm) is applied over the entire sample. Further on, the second lithographic step follows. This time the resist patterning involves the windowing of both the areas opened in the first resist layer and the connection areas between these (see Fig. 3.24-3). The sample is then plated, the thin metal layer conducting the plating current to all parts of the wafer. After the plating operation, the two resist layers and the thin Au coating between them are removed, leaving the plated bridges. For the first lithographic operation we have used AZ 4330A, a 3 μm thick resist. After development and rinse a prolonged bake have been performed (see process B.8). This postbake is very important to

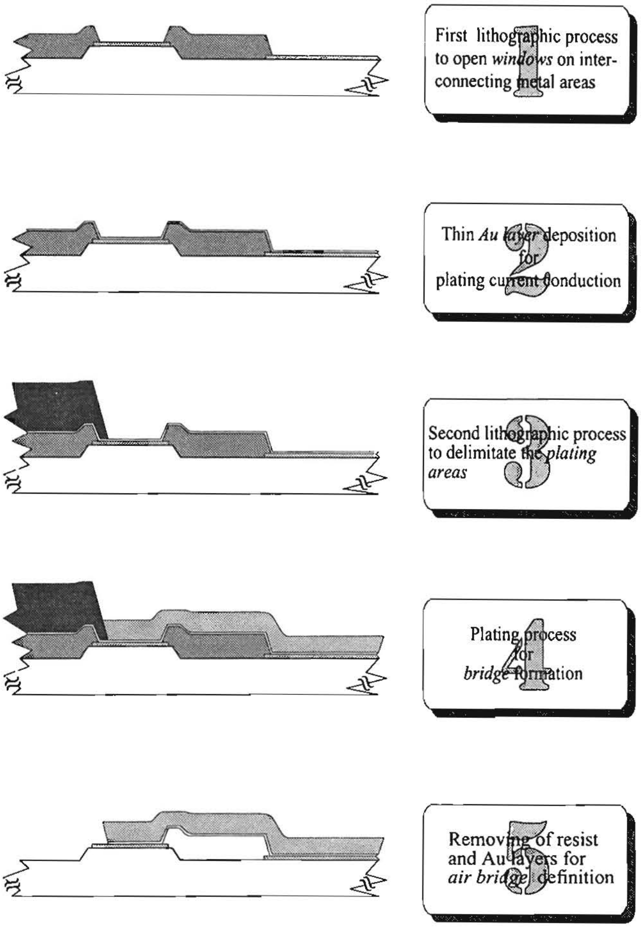


Fig. 3.24 Airbridge fabrication flow using gold plating process.

ensure a perfect dryness of the resist layer, otherwise, during the softbake of the second resist layer, the metal coating swells up in some places and the whole process would be compromised. The thin Au intermediate layer can be evaporated or sputtered. Sputtering, however, produces better coverage of all the features than does evaporation. The second lithographic process uses the 6 μm thick resist AZ 4620A. To avoid the stress on the metal layer and its deterioration the soft bake of this resist has been done after a special procedure: first drying at room temperature, then baking on the hot plate and finally keeping the sample again at room temperature a certain interval of time before resist exposure (see process B.9). As in the first lithographic step, a postbake operation has been performed too, but using different conditions. In this case the resist layer must be sufficiently baked to withstand the plating environment. Finally, the resist and metal

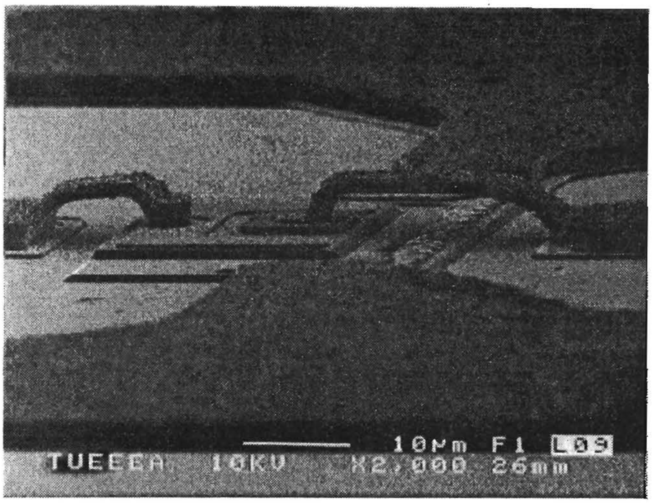


Fig. 3.25 Gold bridges on a three terminal DBRT device.

layers have been removed individually, one after the other. Proceeding in this way any damage of the airbridge structure is avoided. The entire procedure is presented in the Appendix B. Fig. 3.25 shows air bridges realised during three terminal DBRT device processing.

3.4 Design and Fabrication of Masks

In this section we present some guiding lines concerning the design and fabrication of the masks, as they have been "crystallized" during our involving in this field. Here we specify that usually the proper design of the device "physical" aspect and that one of the corresponding masks is accomplished in parallel. For this reason the term "mask design" designates very often both types of activities and we adopted this acception.

The mask design is very important in the process of device fabrication and, of course, for its later functioning. A "well thought" set of masks will ensure a "smooth" processing afterwards, without insurmountable difficulties. The device geometry together with the shape, dimensions and disposing of the contacts and pads have impact on the device performances. They determine too the access to the electrical and/or optical measurements for device characterization and testing. So it is evident that the design of the masks has to be accomplished with discernment, taking into account also that high quality masks (see below) have considerable costs.

In the previous sections about technology, we have seen that the device processing is structurized into two main categories of processes, subtractive and additive. Also it was pointed out that each process is preceded by a lithographic step, which uses a specific mask. A mask consists of a high quality glass plate (slide) which surface is coated by a patterned thin film. For the patterned layer hard-surface (chromium, iron oxide, silicon) or soft-surface materials (photoemulsion, photoresist) can be used. The patterns contained in a mask destined to a certain process form a *reticle* or *layer*. The term "reticle" is used particularly in the mask fabrication activity and "layer" in the mask design. As a rule, if the lithography is accomplished with a positive resist, the mask consists in a *light field* reticle (layer) for a subtractive process and a *dark field* reticle for an additive one. Of course, the above rule reverses when a negative resist is used. Light and dark field terms refer to the fact that the "drawn" patterns on the mask are opaque or transparent, respectively, for the specific energetic source of resist sensitizing, the rest of the mask being in opposite condition.

During the mask design we distinguish many stages. In the first one, the different "physical elements" of the devices (mesas, contacts, pads, etc.) and additional test configurations together with their form and dimensions are established. This is done in concordance with their functionality and working conditions, with the goal to obtain an optimum access to the properties of the device structure. All the options outlined during this step have to take into account the technological limits existing in the laboratory.

The data from the first stage are then used in the next step to establish the shape and dimensions of the mask patterns corresponding to the different elements of the devices. This is perhaps the most important and difficult part of the mask design because now it has to be taken into account all the deviations from "ideality" introduced by the different technological processes, later, during device fabrication. The correct estimation of these deviations permits the design of mask patterns which will reproduce, after processing, the initial data of the device elements without technological difficulties. We discuss some aspects concerning this delicate problem. When photolithography is used for pattern definition as in our case, this constitutes the most important source of deviations. In section 3.2.2 about photolithography we have pointed out that the edge diffraction effects determine in the resist image *dimensional deviations* from the mask features. If we define the dimensional deviation δ as the difference between the dimension of the resist image and that one of the resist pattern (see Fig. 3.26a) then for the light field reticles (masks) $\delta < 0$ and for the dark field ones $\delta > 0$. An estimation of the δ values for different types of resist is given in the Appendix B. Another type of deviation from ideality which can be introduced during the photolithographic process is the deviation from the designed superposition of two (consecutive) resist images (see Fig. 3.26b). This kind of deviation is caused by the misalignment. There are many sources of misalignment. One of them proceeds from the impossibility to focus in the same time both the reticle and sample surface planes. This leads to the impossibility to use correctly the alignment marks from the mask. The above drawback proceeds from the gap between the mask and sample surfaces, introduced by the resist thickness and/or edge beads (see subsection 3.2.2a). There are means to remove the edge beads but if the resist thickness is the cause of the misalignment, this represents an unavoidable reality which has to be considered. Estimations of the misalignment introduced in this case, for different resists used in our photolithographic processing are presented in the Appendix B. Other sources of misalignment are the design of the alignment marks, the bending of the mask plate, and the alignment mark shifts introduced in the different layers during the mask fabrications (see below).

Between the two main groups of subtractive processes, wet and dry etching, the first one produces dimensional deviations of the patterns transferred in the sample versus mask features. Because of the underetch phenomena registered during the wet chemical etching, the transferred features can be smaller (mesas) or larger (holes) than the designed ones on the mask. The

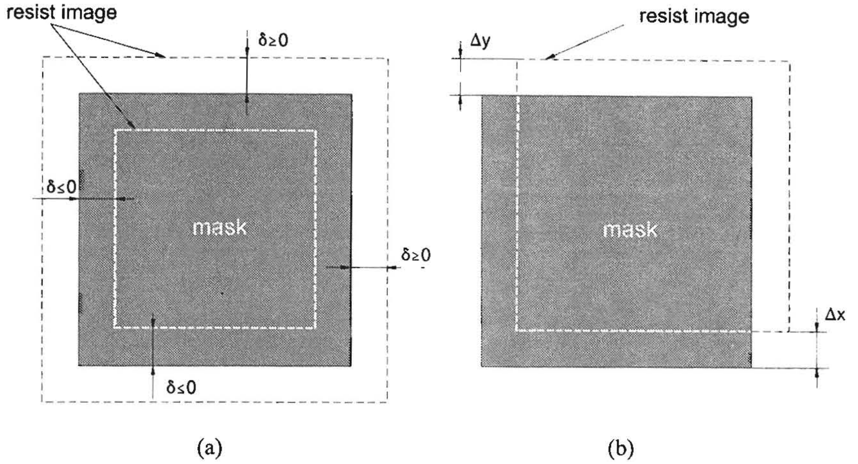


Fig. 3.26 Deviations from "ideality" introduced during the photolithographic process: (a) dimensional deviations due to diffraction edge effects and (b) misalignment.

underetch depends of the resist adhesion: it can be limited by a proper choice of the resist, but never be avoided. In the Appendix B, the underetch produced of the wet chemical etching processes used in our device fabrication is evaluated.

After the complete estimation of the mask feature geometry they are drown using a specific software program. For our masks we have used the XDALI program, developed at the Delft University for IC mask design. The drawing strategy is to realize a "horizontal" grouping of the patterns, on devices and test configurations, and a "vertical"one, on layers corresponding to all lithographic processes during device fabrication. Finally, we get layered patterns for each device, separately. During this stage we have to design the alignments marks too. The alignment marks permit the correct superposition of the resist images corresponding to patterns which belong to different layers. During processing the alignment marks from a layer are transferred onto the sample, via the resist image. Subsequently, these "marks" on the sample are used for the alignment of the next layer. These marks always have to ensure an alignment with an accuracy better than the smallest geometrical difference between the superimposed layer features. Evident in this calculus we have to take into account also the deviations from ideality appearing during the device processing, discussed previously. When the difference is very small the adequate design of the alignment marks becomes crucial. The problem involves in fact, the foresight of sufficient and suitable geometrical elements which, by sequential transfer onto the sample using specific processes, ensure the alignments of the features with the desired accuracy. As an example, in Fig. 3.27 there are shown the alignment marks designed for the processing of our three terminal DBRT devices. Usually the alignment marks on the different layers determine a strict sequence of the processing steps initially established. It is much better, however, if the design of these permits the change of the processing sequence, in this way offering flexibility to the device fabrication. The alignment marks presented in Fig. 3.27 ensure such a flexibility for

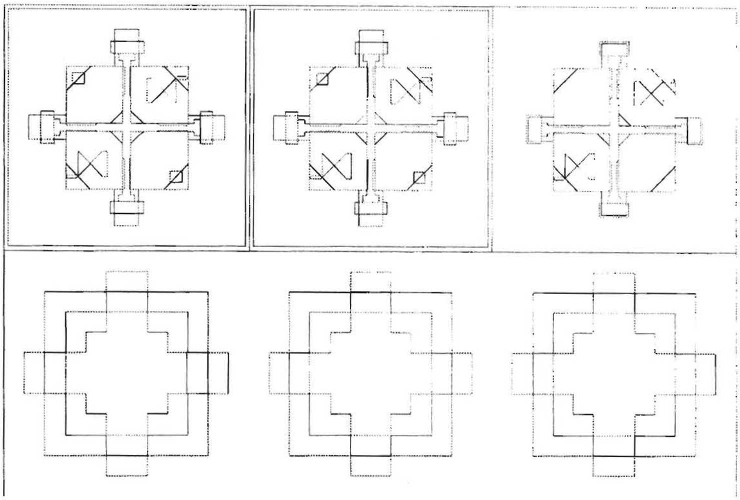


Fig. 3.27 Alignment marks designed on the masks destined to the three terminal DBRT device fabrication.

the DBRT device processing.

The last stage of the mask design consists in the integration of all the patterns drawn previously for devices, test configurations and of course, alignment marks into a *main cell*. The arrangement of the features has to be logically and clear. Different "fields" of patterns are realized, the grouping of these being made after common characteristics of the devices they represent. The fields are labelled in a concise and suggestive manner. Finally, the device patterns have to be arranged in a matrix form with lines and rows spacial separated, which go along the whole cell. In this way, after processing, if necessary, the separation of the devices in chips is easy to be accomplished, without the risk to be damaged.

The mask fabrication comes to materialize the former design activity. We have not been directly involved in the fabrication of our masks for the DBRT devices, but our experience shows that a permanent contact with the mask manufacturer (maker) during the whole period of fabrication leads to a rapid and good quality execution. Some aspects which can appear after the mask fabrication have to be pointed out. Actually, the mask fabrication is accomplished in a number of ways [99]. In principle, all of them are based on a lithographic process, during which a sophisticated electron-beam or optical imaging system is used. The imaging systems realize the conversion of the digitized data representing the designed layers into a "physical" image, developed in the radiation sensitive material lying on the mask plate. Our masks have been realized with an optical imaging system at MESA Research Institute, University of Twente. This system is a pattern generator which takes the digitized data and converts them to a series of optical flash exposures on the photoresist-coated chromium plate [99,100]. Using the pattern generator, reticles have been fabricated after all layers representing the main cell. After this, because our lithographic technique is based on the contact printing optical exposure (see section 3.2.2), from each reticle (layer) obtained previously, arrays of the main cell have been generated on other chromium plates. This operation has been accomplished with a photorepeater, which is a laser interferometer-metered step-and-repeat camera [99,100]. So, each ready mask contains finally arrays of the main cell corresponding to different technological steps of the device processing.

After fabrication the masks must be checked attentively and all the nonconformities with the design or defects due to processing have to be rectified. The glass plate of the mask can have different thicknesses. The choice of the plate thickness must be done in connection with the own lithographic equipment. If there is any possibility of plate bending then thicker plates will be preferred. This will determine a higher energy absorption by the plate, but with a longer exposure time the problem can be resolved. Finally, we point out a drawback which can appear during the second main step of mask fabrication, the photorepeating of the main cell reticles. The photocopying operation is submitted to positioning errors during the horizontal movement of the stage on which the mask glass plate lays. The errors translate finally in misalignments of the corresponding array elements situated on different masks (layers). The alignment deviations can be in this situation as high as (up to) 0.5 μm [100].

3.5 Conclusions

Except for the epitaxial growth, the technological processes used for the fabrication of the two and three terminal RT devices are presented in detail. The basic idea of the presentation is the understanding of the process mechanism and their direction of evolution. In this sense an exhaustive description of the processes is carried out, based on pertinent studies in the literature and our own experience. Special attention has been paid to the dual image resist photolithography for specific applications, wet chemical etching, selective etching of GaAs and AlGaAs, and gold plating for air bridge interconnections. Based on accepted descriptions of the metallurgical behaviour of two of the most used metallization schemes for ohmic contacts, the contact formation of our Ge/Ni/Au contact sequence is illustrated.

Some guiding lines concerning the design and fabrication of masks, as they have been "crystallized" during the development of a mask set for the 3T device processing, are also included. It is concluded that between this activity and the knowledge gained in the technological field there is a strong interdependence.

The level of optimization of the technological processes described in this chapter permits a high quality processing of the 2T and 3T DBRT devices presented in the next sections. Also, they can be used in good conditions for the approach of other kinds of modern compound semiconductor devices, such as HBT's [98], MESFET's [23], and HEMT's [23, 101].

References

1. F. L. Pedrotti and L. S. Pedrotti, "*Introduction to Optics*", 2nd. ed., Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1993.
2. D. A. McGillis and D. L. Fehrs, *IEEE Trans. Electron Devices* **ED-22**, 471 (1975).
3. K. A. Valiev, "*The Physics of Submicron Lithography*", Plenum Press, New York, 1992.
4. D. Windmann and K. U. Stein, "*Semiconductor Technologies with Reduced Dimensions*", in Proc. 2nd Eur. Solid State Circuits Conf., 1976, p. 29.
5. W. S. DeForest, "*Photoresist Materials and Processes*". McGraw-Hill Book Co., Inc., New York, 1975.
6. D. Meyerhofer, *IEEE Trans. Electron Devices* **ED-27**, 921 (1980).
7. O. Süs, *Ann. Chem.* **556**, 65 (1944).
8. F. A. Vollenbroek, W. P. M. Nijssen, C. M. J. Mutsers, M. J. H. J. Geomini, M. E. Reuhman, and R. J. Visser, *Polym. Eng. Sci.* **29**, 928 (1989).
9. A. C. Ouano, *Polym. Eng. Sci.* **18**, 306 (1978).
10. W. D. Hinsberg, C. G. Wilson, and K. K. Kanazawa, *Proc. SPIE* **539**, 6 (1985).
11. J. Pacansky and D. Johnson, *J. Electrochem. Soc.* **124**, 862 (1977).
12. J. Pacansky and J. R. Leyrla, *IBM J. Res. Dev.* **23**, 42 (1979).
13. K. Skydmore, *Semiconductor International*, February 1988, p.56.
14. L. F. Thomson and R. E. Kerwin, *Annu. Rev. Mater.* **6**, 267 (1976).
15. A. Ouano, in *"Macromolecular Solutions"*, eds R. Seymour and G. Stahl, Pergamon Press, Elmsford, N.Y., 208 (1982).
16. F. Dill and J. Shaw, *IBM Res. Dev.* **21** 210 (1977).
17. J. Shaw, M. Frisch, and F. Dill, *IBM Res. Dev.* **21**, 219 (1977).
18. T. Batchelder and J. Piatt, *Solid State Technology* **26**, 211 (1983).
19. W. M. Moreau, "*Semiconductor Lithography - Priciples, Practices and Materials*", Plenum Press, New York, 1988.
20. M. Andrasi, *Thin Solid Films*, **67**, 229 (1980).
21. J. Bieron and R. Conley, *J. Appl. Polym. Sci.* **1**, 171 (1963).
22. A. Johnson, *Kodak Microelectronics Seminar* G-136, 100 (1982).
23. J. G. van Hassel, "*Dry Processing of GaAs-based MESFETs and Pseudomorphic HEMTs*", PhD Thesis, Eindhoven University of Technology, ISBN 90-386-0050-X, 1995.
24. H. H. P. M. van den Heuvel, "*Vervaardiging van GaAs/AlGaAs Ridge Waveguide Lasers met nat-chemisch geëtste spiegels*", Afstudeerverslag, Eindhoven University of Technology, 1993.
25. C. G. Wilson, in "*Introduction to Microlithography*", eds. L. F. Thomson, C. G. Wilson, and M. J. Bowden, *ACA Symp. Ser.* **219**, 117 (1984).
26. M. L. Long, *Proc. SPIE* **469**, 72 (1984).
27. W. E. Feely, J. C. Imhof, and C. M. Stein, *Polym. Eng. Sci.* **26**, 101 (1986).
28. M. Bolsen, *Hoechst Bereich Informationstechniek*, July 1988.
29. M. E. Reuhman-Huisken and F. A. Vollenbroek, *Microelectronic Engineeering* **11**, 575 (1990).

30. A. Reiser, *"Photoreactive Polymers - The Science and Technology of Resists"*, John Willey & Sons, New York, 1989
31. R. Williams, "*Modern GaAs processing Methods*", Artech House, Inc., Boston, 1990.
32. B. Tuck, *Journal of Materials Science* **10**, 321 (1975).
33. W. Kern and C. A. Deckert, in *Thin Film Processes*, eds. J. L. Vossen and W. Kern, Acdemic Press, New York, 1978, p. 401.
34. K. Sangwal, *Etching of Crystals : Theory, Experiment, and Application*, North-Holland, Amsterdam, 1987.
35. J. J. Gannon and C. J. Nuese, *J. Electrochem. Soc.* **121**, 1215 (1974).
36. D. W. Show, *J. Electrochem. Soc.* **128**, 875 (1981).
37. S. Iida and K. Ito, *J. Electrochem. Soc.* **118**, 769 (1971).
38. Y. Mori and N. Watanabe, *J. Electrochem. Soc.* **125**, 1511 (1978).
39. S. Adachi and K. Oe, *J. Electrochem. Soc.* **130**, 2427 (1983).
40. J. Bardeen, *Phys. Rev.* **71**, 717 (1947).
41. J. Tamm, *Phys. Z. Sowjetunion*, **1**, 733 (1932).
42. W. Shockley, *Phys. Rev.* **56**, 317 (1939).
43. J. van Laar, P. W. Huiser, and T. L. van Roy, *J. Vac. Sci. Technol.* **14**, 894 (1977).
44. W. E. Spicer, I. Lindau, P. Skeath, and C. Y. Su, *J. Vac. Sci. Technol.* **17**, 1019 (1980).
45. W. E. Spicer, S. Eglash, I. Lindau, C. Y. Su, and P. R. Skeath, *Thin Solid Films*, **89**, 447 (1982).
46. J. M. Woodall and J. L. Freeouf, *J. Vac. Sci. Technol.* **19**, 794 (1981).
47. A. Y. C. Yu, Solid-State Electronics **13**, 239 (1970).
48. C. Y. Chang, Y. K. Fang, and S. M. Sze, *Solid-State Electronics* **14**, 541 (1971).
49. R. M. Bertenburg, '*Technologie koplanarer monolitisch integrierter Mikrowellenschaltungen auf Galliumarsenid unter Verwendung von Submicron-Feldeffekttransistoren*', PhD thesis, Gerhard Mercator Universität-Gesamthochshule-Duisburg, 1992.
50. R. A. Bruce and G. R. Piercy, *Solid-State Electronics* **30**, 729 (1987).
51. N. Braslau, J. B. Gunn, and J. L. Staples, *Solid-State Electronics* **10**, 381 (1967).
52. W. D. Edwards, W. A. Hartman, and A. B. Torrens, *Solid-State Electronics* **15**, 387 (1971).
53. M. Ogawa, *J. Appl. Phys.* **51**, 406 (1980).
54. K. Heime, U. König, E. Kohn, and A. Wortmann, *Solid-State Electronics* **17**, 835 (1974).
55. S. J. Chua, S. H. Lee, R. Gopalakrishnan, K. L. Tan, and T. C. Chong, *Thin Solid Films*, **200**, 211 (1991).
56. T. S. Kuan, P. E. Batson, T. N. Jackson, H. Rupprecht, and E. L. Wilkie, *J. Appl. Phys.* **54**, 6952 (1983).
57. Y. C. Shih, M. Murakami, E. L. Wilkie, and A. C. Callegari, *J. Appl. Phys.* **62**, 582 (1987).
58. M. Ogawa, *Thin Solid Films*, **70**, 181(1980).
59. M. Wittmer, R. Pretorius, J. W. Mayer, and M-A. Nicolet, *Solid-State Electronics* **20**, 433 (1974).
60. E. Relling and A. P. Bhota, *Appl. Surf. Scie.* **35**, 380 (1988-89).
61. H. R. Kawata, T. Oku, A. Otsuki, and M. Murakami, *J. Appl. Phys.* **75**, 2530 (1994).
62. K. Tanahashi, H. J. Tanaka, A. Otsuki, and M. Murakami, *J. Appl. Phys.* **72**, 4183 (1994).
63. T. Sands, E. D. Marshall, and L. C. Wang, *J. Mater. Res.* **3**, 914 (1988).
64. W. T. Anderson, Jr., A. Christou, and J. E. Davey, *J. Appl. Phys.* **49**, 2998 (1978).

65. M. Wittmer, M-A. Nicolet, and J. W. Mayer, *Thin Solid Films*, **42**, 51 (1977).
66. C. Lin and C.P. Lee, *J. Appl. Phys.* **67**, 260 (1990).
67. S. J. Pearton, F. Ren, A. Katz, T. R. Fullowan, C. R. Abernathy, W. S. Hobson, and R. F. Kopf, *IEEE Trans. Electron Devices* **39**, 15 (1992).
68. C. J. Palmstrøm and D. V. Morgan, "Metallization for GaAs Devices and Circuits", in "*Gallium Arsenide - Materials, Devices, and Circuits*", ed. M. J. Howes and D. V. Morgan, John Willey&Sons, Chicester, 1985.
69. A. Iliadis and K. E. Singer, *Solid-State Electronics* **26**, 7 (1983).
70. A. Katz, S. N. G. Chu, B. E. Weir, C. R. Abernathy, W. S. Hobson, S. J. Pearton, and W. Savin, *IEEE Trans. Electron Devices* **39**, 184 (1992).
71. G. Lucovsky, P. D. Richard, D. V. Tsu, S. Y. Lin, and R. J. Markunas, *J. Vac. Sci. Technol.* **A4**, 681 (1986).
72. T. A. Cleland and D. W. Hess, *J. Electrochem. Soc.* **136**, 3103 (1989).
73. D. L. Smith and A. S. Alimonda, *J. Electrochem. Soc.* **140**, 676 (1993).
74. P. A. Longeway, R. D. Estes, and H. A. Weakliem, *J. Phys. Chem.* **88**, 73 (1984).
75. J. Asmussen, *J. Vac. Sci. Technol.* **B7**, 883 (1989).
76. D. L. Flamm, "Introduction to Plasma Chemistry", in "*Plasma Etching-An Introduction*", eds. D. M. Manos and D. L. Flamm, Academic Press, Inc., London 1989.
77. R. d'Agostino and D. L. Flamm, *J. Appl. Phys.* **52**, 162 (1981).
78. S. K. Ray, C. K. Maiti, and N. B. Chakraborti, *Semicond. Sci. Technol.* **8**, 599 (1993).
79. M. Tong, D. G. Ballegeer, A. Ketterson, E. J. Roan, K. Y. Cheng and I. Adesida, *J. Electron. Mater.* **21**, 9 (1992).
80. R. A. Logan and F. K. Reinhart, *J. Appl. Phys.* **44**, 4172 (1973).
81. K. Kenefick, *J. Electrochem. Soc.* **129**, 2380 (1982).
82. J. J. Lepore, *J. Appl. Phys.* **51**, 6441 (1980).
83. C. Juang, K. J. Kuhn, and R. B. Darling, *J. Vac. Sci. Technol.* **B8**, 1122 (1990).
84. G. S. DeSalvo, W. F. Tseng, and J. Comas, *J. Electrochem. Soc.* **139**, 831 (1992).
85. B.-Y. Mao, J. A. Nielsen, R. A. Friedman, and G. Y. Lee, *J. Electrochem. Soc.* **141**, 1083 (1994).
86. K. Hikosaka, T. Mimura, and K. Joshin, *Jpn. J. of Appl. Phys.* **20**, L847 (1981).
87. A. Seabaugh, *J. Vac. Sci. Technol.* **B6**, 77 (1988).
88. Y. K. Su, Y. Z. Juang, S. C. Shei, and B. C. Fang, *Solid-State Electronics* **36**, 1779 (1993).
89. A. A. Ketterson, E. Andideh, I. Adesida, T. L. Brock, J. Baillargeon, J. Laskar, K. Y. Cheng, and J. Kolodzey, *J. Vac. Sci. Technol.* **B7**, 1493 (1989).
90. S. Salimian, C. B. Cooper III, R. Norton, and J. Bacon, *Appl. Phys. Lett.* **51**, 1083 (1987).
91. S. J. Pearton, U. K. Chakrabarti, W. S. Hobson, and A. P. Kinsella, *J. Vac. Sci. Technol.* **B8**, 607 (1988).
92. E. Yablonovitch, T. Gitter, J. P. Harbison, and R. Bhat, *Appl. Phys. Lett.* **51**, 2222 (1987).
93. M. Otsubo, T. Oda, H. Kumabe, and H. Miki, *J. Electrochem. Soc.* **123**, 676 (1976).
94. R. E. Klinger and J. E. Green, *Appl. Phys. Lett.* **38**, 620 (1981).
95. J. Vatus, J. Chevrier, P. Delescluse, and J. F. Rochette, *IEEE Trans. Electron Devices* **ED-33**, 934 (1986).
96. F. A. Lowenheim, *Electroplating*, McGraw-Hill Book Co., New York, 1978.
97. J. O. Dukovic, *IBM J. Res. Develop.* **37**, 125 (1993).

98. A. Wiersch, Private communications.
99. D. J. Elliott, “Integrated Circuits Mask Technology”, McGraw-Hill, New York, 1985.
100. A. Kooy, Private communications.
101. J. Wellen, "*Modelling, design and fabrication of a GaAs-based integrated photoreceiver for short distance optical communication*", PhD Thesis, Eindhoven University of Technology, ISBN90-386-0260-X, 1997.

4

Two Terminal DBRT Devices

4.1 Introduction

To start the three terminal (3T) DBRT device approach without being sufficiently involved in all problems concerning the DBRT diodes it would undoubtedly be nonrealistic. So, naturally, an important part of the work spent during the elaboration of this thesis has been devoted to the study of the two terminal DBRT devices from all points of view, fabrication, characterisation (electrical measurements) and modeling. This activity not only paved the way for the realisation of the 3T resonant tunneling devices but offered us interesting insights concerning the physics of the DBRT diode.

For our DBRT diodes we made use of structures grown from GaAs/$Al_{0.4}Ga_{0.6}As$ and GaAs/AlAs material systems, which have been grown by MBE at the Solid State Physics Department of Eindhoven University.

The DBRT diode processing has been accomplished in several ways. One can say that there existed a natural evolution in this sense, from simple to complex. The first diodes have been processed in the most direct way, using only three main fabrication steps, device area definition by mesa etching and twice contacting. These normal (vertical) diodes, contacted on the front and back side, proved to be a true confirmation of the quality of the structures which we have used, giving the "green" signal for more complex processings. The necessity of such processings was imposed by measurement reasons. Both the DC and especially AC measurements require devices with an active area as small as possible. In this situation the electrical access to the device is realised with connections via measurement pads placed on the front side of the sample. This idea has been carried out in two ways. In a first variant the interconnections between top contacts and pads lie directly on the sample, the mesa edges being isolated from them with a thin SiO_2 film. In the second variant no isolation was necessary because air bridge interconnections have been used. All the three modalities of processing are presented in the following. The first two processing ways constitute the subject of the next section from this chapter and the third one will be described in chapter 5 because it is in connection with the fabrication of the 3T-DBRT devices (simultaneous processing using the same masks). During processing activity we have realised the optimization of many of the processes involved in the different ways of device fabrication. A detailed presentation of all these processes has been done in the previous chapter. Here we point out only the fact that in connection with the first two procedures of device processing we paid special attention to the wet chemical etching, the lift-off process using dual image photoresist, in both direct and image reversal variants, and the photolithographic process (with the same kind of resist) suitable for the limitation of the dimensional deviation of the resist image versus mask pattern.

After the device fabrication this chapter describes the electrical measurements of the 2T-DBRT devices. Both room temperature and temperature dependent DC measurements are presented and evaluated comparatively concerning the different device structures, with reference also to the performances of similar devices described in the literature. We refer then to the AC measurements, carried out at room temperature.

The last part of the chapter is devoted to the modeling of the DBRT diode. First, the model developed in the group is presented. Then the results of the model are discussed in comparison with our DC measurements previously introduced.

4.2 Device Fabrication

As mentioned above the 2T-DBRT devices have been fabricated from MBE grown structures based on the GaAs/$Al_xGa_{1-x}As$ material system. Totally, three structures have been processed, the layer sequence and their characteristics being shown in Table 4.1. For identification we have maintained the structure name used at the growth place. As can be seen, concerning the thickness and doping profile of the layers forming the cladding regions, the structures are quite similar. The superlattice (layer 3) inserted before the active (double barrier quantum well) region for the structures W309 and W488 is intended to improve the quality of the next grown layers. Both of the structures W115 and W309 have been grown on n^{++}-GaAs substrate while for the structure W488 a semi-insulating (SI) substrate has been used. The active region has AlAs barriers in the structures W309 and W488, and lower $Al_{0.4}Ga_{0.6}As$ barriers in the structure W115.

The first processing way we describe in this section represents the most usual and simplest procedure to fabricate semiconductor (GaAs) diodes. This means that it includes not more than three main steps: one for the device definition and another two for the fabrication of the contacts. Sometimes a sample polishing step is used for device separation reasons. This kind of processing necessitates only one mask containing patterns which finally will be replicated on the sample, leading to the corresponding device fabrication. In our case we have benefited of a mask with square, rhombic and circular patterns with characteristic dimensions of 5, 10, 20, 30, 40 and 50 μm. The mask has been available in both variants, light and dark field. The use of one or the other variant depends of the photolithographic process intended to be used, direct or image reversal process. Also, using one of the mask variant, filled patterns (dots) of the above shapes are replicated in the resist film and then transferred into the sample but making use of both of the two variants, in a suitable sequence, pierced patterns (rings) with the same geometries can be obtained. In the following we refer to both possibilities. Everywhere, an image reversal photolithographic process has been used. As seen in Fig. 4.1 the dot device processing contains three main steps. First, using the mask, the top device contact is deposited by the lift-off procedure. This means that at the end of this step (specific) metal dots lie on the sample surface. These dots are subsequently used as mask for the next main step, the mesa wet chemical etching for device definition. Finally, the second contact is deposited on the whole back side of the

Layer	Material	Thickness	Doping	Wafer / Structure			Observations
				W115	W309	W488	
13	GaAs	500.0	$2.00\text{x}10^{18}$ (n) $1.00\text{x}10^{18}$(n)	▨ 	 ▨	 ▨	contact layer
12	GaAs	50.0	$7.20\text{x}10^{17}$(n)	▨	▨	▨	buffer
11	GaAs	50.0	$2.00\text{x}10^{16}$ (n)	▨	▨	▨	buffer
10	GaAs	2.5	undoped	▨	▨	▨	spacer
9	$Al_{0.4}Ga_{0.6}As$ or AlAs	5.6 2.5	undoped	▨ 	 ▨	 ▨	barrier
8	GaAs	5.0	undoped	▨	▨	▨	quantum well
7	$Al_{0.4}Ga_{0.6}As$ or AlAs	5.6 2.5	undoped	▨ 	 ▨	 ▨	barrier
6	GaAs	2.5	undoped	▨	▨	▨	spacer
5	GaAs	50.0	$2.00\text{x}10^{16}$ (n)	▨	▨	▨	buffer
4	GaAs	50.0	$7.02\text{x}10^{17}$ (n)	▨	▨	▨	buffer
3	GaAs + $Al_{0.33}Ga_{0.67}As$	2.5 2.5	$2.00\text{x}10^{18}$ (n)		▨ ▨	▨ ▨	x100 ➢superlattice
2	GaAs	100	$2.00\text{x}10^{18}$ (n)	▨	▨	▨	contact layer
1	GaAs	100	$2.00\text{x}10^{18}$ (n)	▨	▨	▨	buffer
	GaAs or GaAs	$\sim 4.0\text{x}10^{5}$	(n^{++}) (SI)	▨ 	▨ 	 ▨	substrate

Table 4.1 Structure of the MBE grown wafers W115, W309 and W488.

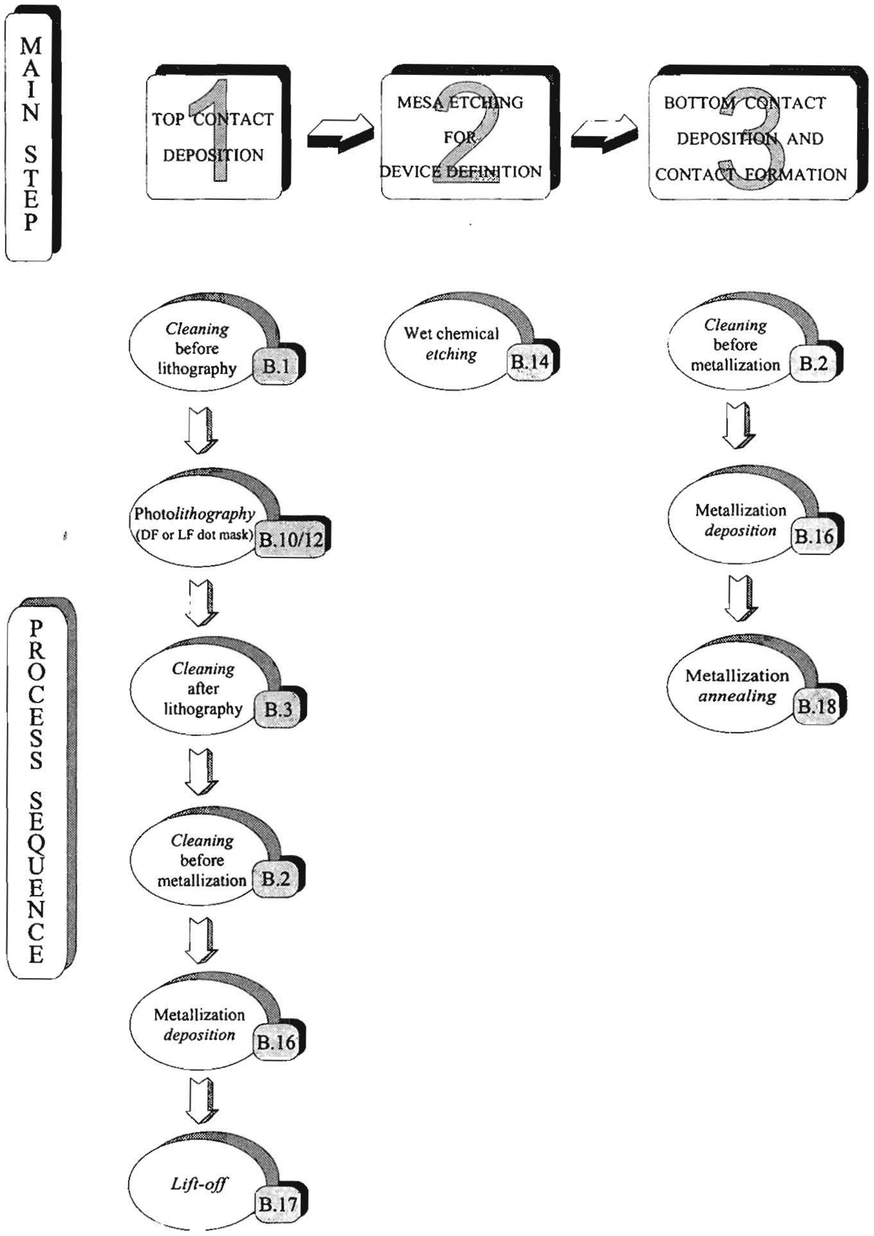


Fig. 4.1 Processing flow for the dot 2T DBRT device fabrication. "B.i" denote the process "recipe" detailed in the Appendix B.
]

sample. The ring processing is very similar with the above description of the dot device fabrication except that in the first main step a supplementary lithographic process with the complementary mask is used.

The second modality of 2T-DBRT device processing has been accomplished using masks from a set designed in collaboration with Duisburg University of Technology (DU), Semiconductor Devices Group, for the fabrication of both HEMTs (MESFETs) [1] and DBRT diodes. The mask patterns used during the processing to obtain the different device components (mesas, contacts, pads) are likewise those introduced in our masks for the fabrication of the same kind of devices (see section 5.3 of the next chapter). A typical device layout is shown in Fig. 4.2. The specific design of the pads permits the access with Cascade probes for microwave measurements and of course better conditions (on wafer) for other kinds of electrical measurements. We specify also that square shape devices with the characteristic dimension of 4, 6, 10 and 25 μm are obtained using these masks. The processing procedure contains four main steps in this case (see Fig. 4.3). Because of the mask design the most difficult step is the deposition of the contacts (step 3). Thus, between the patterns destined for mesa etching and those for top contact deposition there is a relative small difference (~ 1.5 μm). Moreover, the patterns for both top and bottom contact deposition are situated on the same mask. In these conditions, using optical lithography and wet chemical etching, the positioning of the top contact on the mesa surface without reaching its edges have constituted a real problem. We have surmounted this difficulty developing a specific lithographic process for the limitation of the dimensional deviations of the resit image from the mask pattern (see section 3.3.2b and Appendix B, B.2). With the same purpose, the mesa definition has been performed using a wet chemical etching process with low underetch. Because the interconnections between the top contacts and measurement pads lie on the sample surface we have to choose the suitable orientation of the sample before the lithographic process for etching, which can ensure finally the positioning of the middle pads on the part of the mesas with overcut profile (there is no rotational symmetry of the

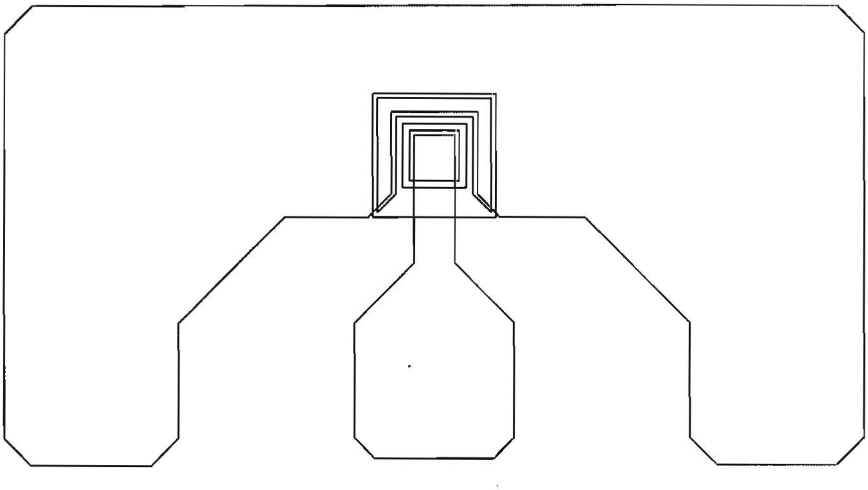


Fig. 4.2 2T DBRT layout on the EUT-DU mask set.

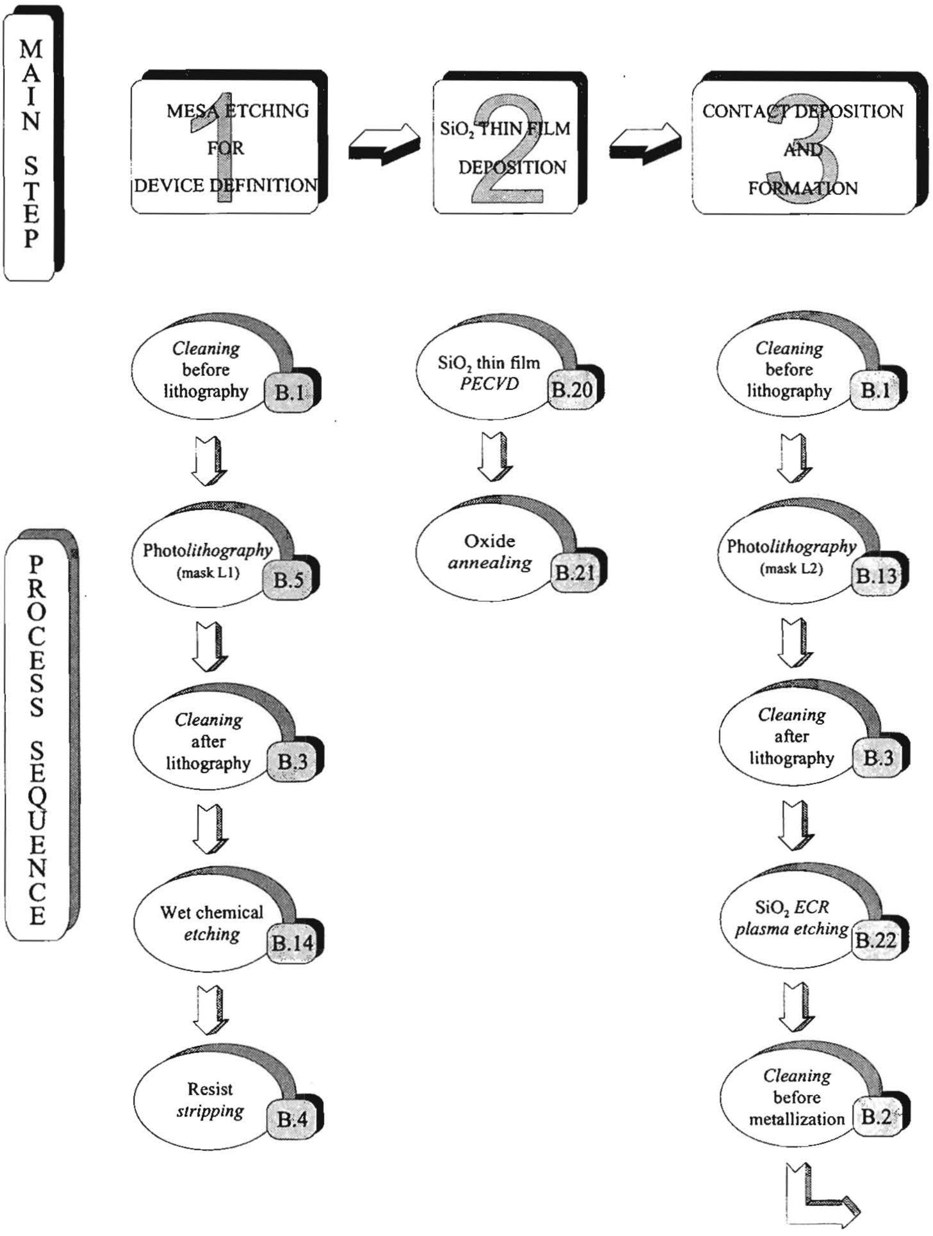


Fig. 4.3 Processing flow of the 2T DBRT device fabrication using EUT-DU mask set. "B.i" denote the process "recipe" detailed in the Appendix B.

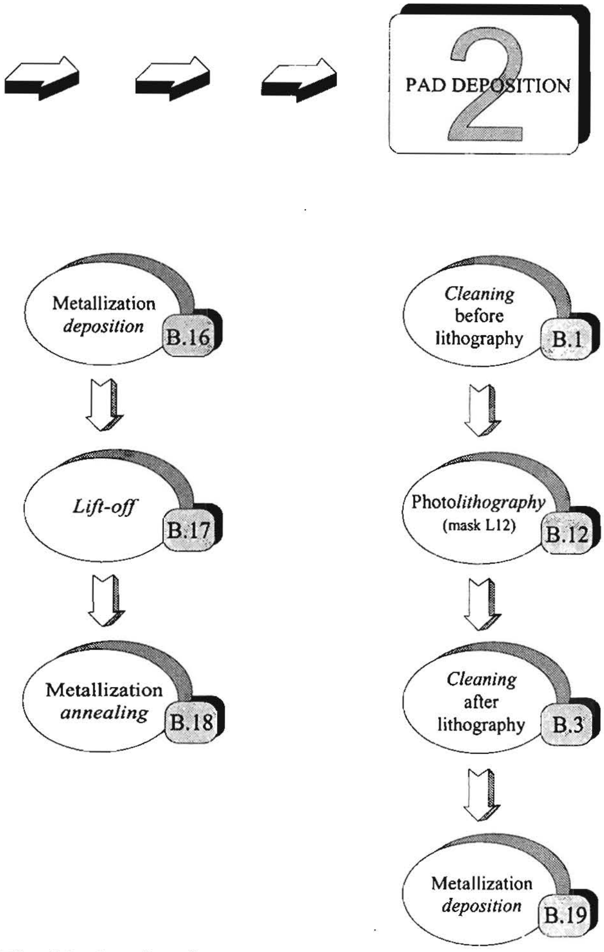


Fig. 4.3 (continued)

masks). At last, the use of the SiO_2 for electrical isolation imposes the use of the primer in the subsequent photolithographic processes and also a specific metallization for the pads, both of them from adhesion reasons. Fig. 4.4 shows a 2T-DBRT device processed by means of the EUT-DU mask set.

The third procedure for 2T DBRT device processing, with air bridge interconnections between the top contact and measurement pads, is part of the 3T DBRT processing and, as mentioned, will be described in the next chapter. We here only specify that the devices fabricated in this way have square active areas with the characteristic dimension of 10, 14, 20, and 28 μm.

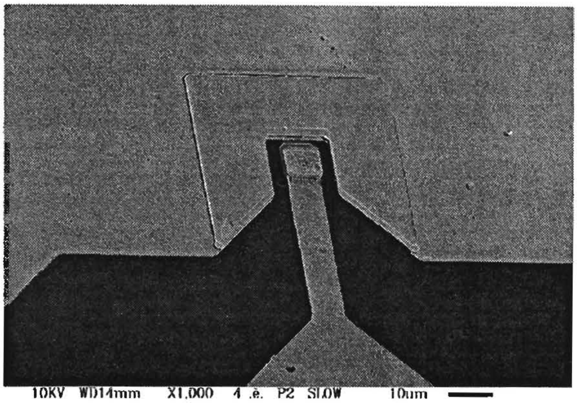


Fig. 4.4 2T DBRT device processed by means of the EUT-DU mask set.

4.3 Measurements

4.3.1 DC Measurements

As known, the I-V characteristic of the DBRT devices represents a powerful characterization tool of the device structure. In the following we make a minute analysis of the DC data for our 2T DBRT devices from which we infer specific features of the resonant tunneling process.

a) Room Temperature DC Measurements

The room temperature (RT) DC measurements have been accomplished on-wafer, using high quality probes, with an HP 4141B DC source/monitor system, computer controlled by the IC-CAP 4.4 software package. The system is equipped with four programmable source/monitor units (SMUs), two programmable voltage sources, two voltage monitors, and an active ground unit. The SMUs can be used in the voltage mode - constant voltage supply with the measure of the corresponding current- or inverse, in the current mode. Each SMU output is connected to a triaxial output connector. The centre conductor is connected to the Sense line, a high input impedance voltage monitor. The inter-conductor, Force line, is the output of the voltage source or current source. This "Force" and "Sense" configuration enables accurate voltage measurements regardless of the voltage drop along the output cable (Kelvin connections). The outer-conductor is connected to the Guard terminal of the SMU, which maintains the same voltage as that of the "Force" terminal. This helps to accurately measure low currents. The ground unit (GNDU) has a "Force" and "Sense" line similar to that of the SMU with the view to keeping the input voltage at 0 V, which helps more the accurate voltage measurements by the SMU. For our 2T DBRT device measurements we have used a SMU and the GNDU.

Before to present the DC results of the 2T DBRT devices we specify that after processing, for each sample, we have determined the contact resistance of the device ohmic contacts. This has been accomplished using the "classical" transmission line model (TLM) [1,2] on specific metallized test pattern defined during processing. The test pattern is composed from similar ohmic contact areas, differently spaced, as illustrated in the inset of Fig. 4.5. The contacts have the width W, the length L, and are separated by a variable distance l_i. The whole pattern is mesa isolated to restrict current to flow only across the distance l_i. In this conditions the total resistance between two consecutive contacts can be expressed as:

$$R_T = 2R_C + R_{Sh}\frac{l_i}{W} = 2R_{ShC}\frac{L_T}{W} + R_{Sh}\frac{l_i}{W}, \text{ for } L_T \ll L \tag{4.1}$$

where R_C is the contact resistance of the ohmic contact area, R_{Sh} is the sheet resistance of the semiconductor contact layer between the contacts, R_{ShC} is the sheet resistance of the material under the contact and L_T the transfer length (the lateral distance required for the current to flow into (or out of) the ohmic contact. The plot of the R_T versus the contact separation (l_i) yields R_C and R_{Sh}. Fig. 4.5 exemplifies the procedure in the case of a processing of the sample W488, using our designed masks where the sequence of l_i's is 3, 6, 12, 24, 48 μm. The EUT-DU masks contain also test patterns for ohmic contact determination with l_i's of 2.5, 5, 10, 20 and 40 μm. In general, we have obtained for our DBRT devices $R_C \leq 0.1$ Ωmm, which indicate a good quality of the ohmic contacts [2]. As observed in Fig. 4.5, if we know R_{ShC}, which in alloyed contacts is different from R_{Sh} (lower), the plot R_T (l) permits also to determine L_T and then the specific contact resistance $r_c = R_C W L_T$ [8]. These two quantities are also used for the characterization of the ohmic contacts. Technically, the R_{ShC} can be found by measuring the contact end resistance R_E [1, 2, 3]. However considering $R_{ShC} = R_{Sh}$ we found values of the r_c around $1\text{x}10^{-6}$ Ωcm^2,

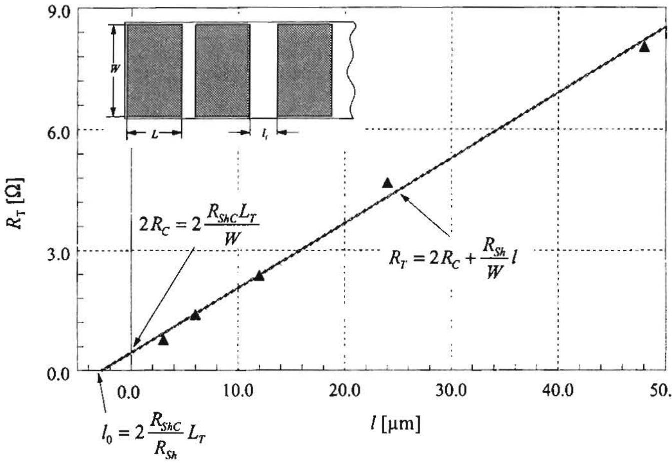


Fig. 4.5 Plot of the measured resistance versus distance between consecutive test ohmic contacts on a pattern as shown in the inset. The sample has been processed from the W488 wafer. The characteristic ohmic contact quantities extracted in this case have the values $R_C = 0.045$ Ωmm, $R_{Sh} = 32.2$ Ω/□, $L_T = 1.38$ μm and $r_c = 1.24\text{x}10^{-6}$ Ωcm^2. L_T and r_c are evaluated considering $R_{ShC} = R_{Sh}$.

which, even taken into account the underestimation of this quantity caused by the approximation, proves again the good quality of our ohmic contacts.

Fig. 4.6 shows representative I-V characteristics of the 2T DBRT devices proceeding from the three different samples which we have processed, namely W115, W309, and W488. Characteristic data of these measurements are given in Table 4.2. The positive polarity corresponds to the situation when the electrons are injected from the substrate side towards the top of the device.

Wafer	Forward bias (+ on top)					Reverse bias (- on top)					Asymmetry	
	J_p x10^3 [A/cm^2]	V_p [V]	J_v x10^3 [A/cm^2]	V_v [V]	$r_{p/v}$	J_p x10^3 [A/cm^2]	V_p [V]	J_v x10^3 [A/cm^2]	V_v [V]	$r_{p/v}$	V_p^-/V_p^+	$r_{p/v}^-/r_{p/v}^+$
W115	2.89	0.66	1.25	0.86	2.32	2.79	0.74	1.26	0.94	2.21	1.12	0.95
W309	7.69	0.96	2.46	1.25	3.1	8.69	1.04	3.02	1.27	2.88	1.08	0.93
W488	4.16	0.96	1.22	1.29	3.42	4.32	0.89	1.41	1.10	3.06	0.92	0.895

Table 4.2 Characteristic DC data of 2T DBRT devices from samples W115, W309, and W488.

As observed, the negative differential resistance (NDR) region of the characteristics is distorted by oscillations of the devices in the measuring circuit. From this reason the values of the valley current density (J_v) and voltage (V_v) can not be determined quite accurately. To have a discussion basis we iterate here that the structure of the devices proceeding from the three samples have the same geometry and doping profile in the cladding regions. The differences regard the substrate used for growth, n^{++}-GaAs for W115 and W309, and SI-GaAs for W488, and, much more important, the barrier material and thickness, $Al_{0.4}Ga_{0.6}As$/5.6 nm for those from W115 and AlAs/2.5 nm for those from the other two samples, W309 and W488.

It is quite difficult to make a real comparison of the DC properties of our devices to the data from literature. On the one hand they refer to structures with a high variety of structural parameters, which do not fit from a study to another and of course with ours. On the other hand, they are sometimes inconsistent with each other, that is understandable if we take into account the high sensitivity of the DBRT device electrical properties to the quality of the active structure (growth conditions) [4]. However systematic experimental studies describing the influence of some structural parameters, such as the barrier thickness and energetic height [5-9] or the quantum well width [9,10], are useful in the DC analysis of our devices. The above remarques suggest us that a comparative analysis of the DC results seems to be the most indicated, some defining characteristics of the tunneling process in DBRT devices being pointed out. In this sense, we first observe that the devices proceeding from all the structures have asymmetric I-V characteristics regarding the polarity. However, concerning this aspect we retain that the I-V characteristic asymmetry of the W488 devices is in opposite sense than that corresponding to the other two kinds of devices. Then, the W115 devices have the lowest peak current density (J_p) corresponding to the lowest V_p. J_p's for the W309 and W488 devices (identical structures on different substrates) are different but correspond to the same V_p for the forward bias. Further on, we remarque that besides the main peak all the I-V characteristics show a small feature

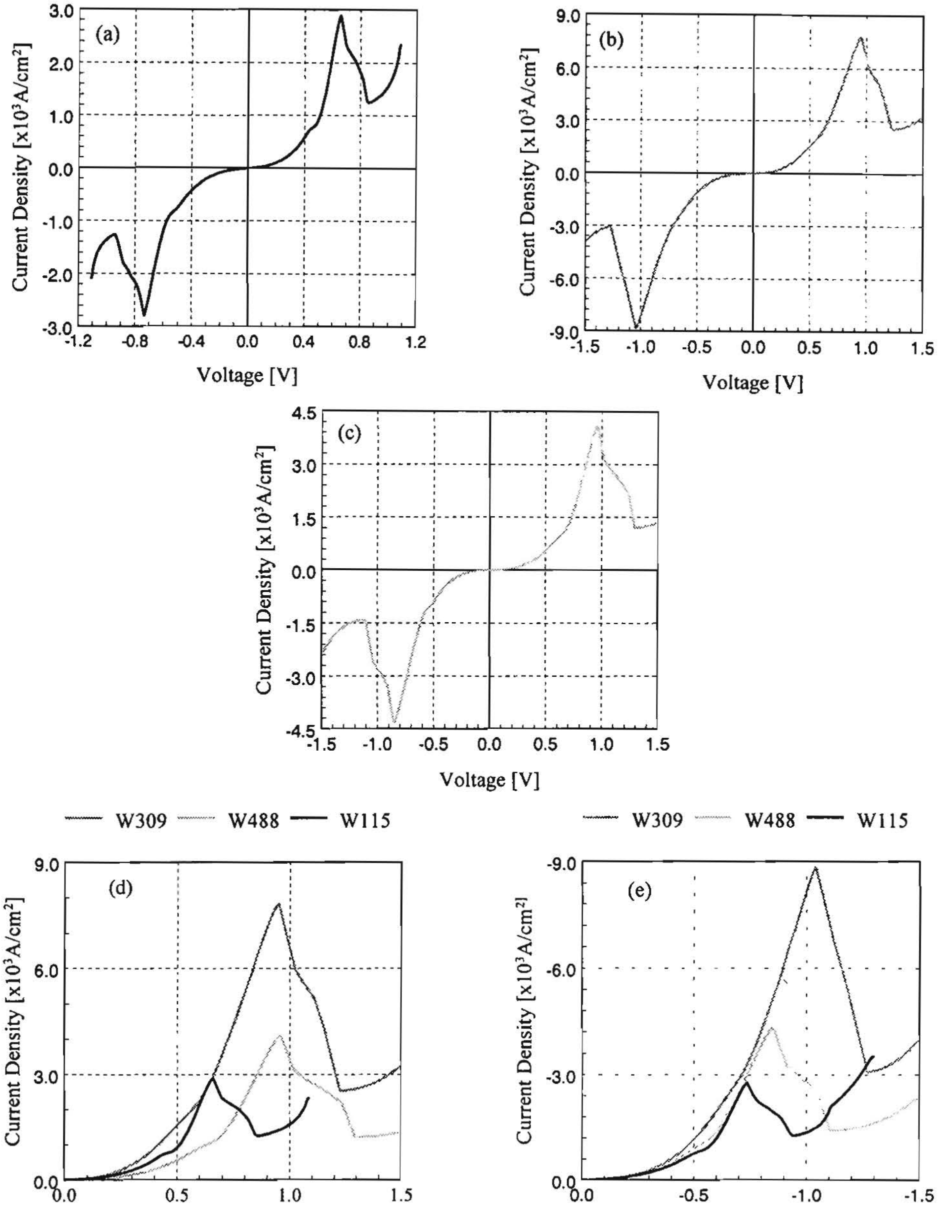


Fig. 4.6 I-V characteristics for 2T DBRT devices proceeding from samples W115 (a), W309 (b), and W488 (c). (d) and (e) show all the characteristics, by comparison, for the positive and negative polarity, respectively.

(undulation) at lower biases. Finally, we mention that the peak to valley ratio ($r_{p/v}$) increases from sample to sample in the order W115, W309 and W488, manifesting a similar asymmetry for all the samples.

To explain some of these characteristics first we make some remarques concerning the DC behaviour of a DBRT diode. These are also useful for the discussion of the I-V characteristic temperature dependence which we present later on, in the next paragraph. Fig. 4.7 shows a detailed picture of the band diagram for the emitter and barrier (active) regions which mainly determines how the current flows through the device. As mentioned already in chapter 2 and is clearly presented in Fig. 4.7, under bias conditions an accumulation layer arises in front of the first barrier where, energetically, the electrons reside in both usual conduction band three-dimensional (3D) states and in quasi-two-dimensional (2D) states. Using Fig. 4.7 and the expressions of the current density corresponding to the two energetic domains from the accumulation layer (see Eqs. 2.24 and 4.5) we can assert that the current flow depends on two main factors: the longitudinal electron supply and the transmission probability through the DBRT structure. The first factor is in connection with the 3 D electron energy distribution in the tunneling direction and 2D electron concentration corresponding to the different emitter quasi levels. Concerning this aspect a common quantity for the two kind of electrons used to calculate the current density is:

$$g(E_z) = \frac{m^* k_B T}{\pi \hbar^2} \ln\left[1 + \exp\left(\frac{E_F - E_z}{k_B T}\right)\right] \tag{4.2}$$

$g(E_z)$ is the integral of the Fermi-Dirac distribution function $f(E)$ times the density of states over the 2D **k** space parallel to the interfaces and has the meaning of electron surface concentration in the plan perpendicular to the tunneling direction corresponding to a certain longitudinal energyE_z. In Eq. 4.2 E_F and E_z are measured from the minimum of the conduction band (E_C) at the interface with the first barrier. Fig. 4.8 shows the dependence of $g(E_z)$ for a particular E_F and

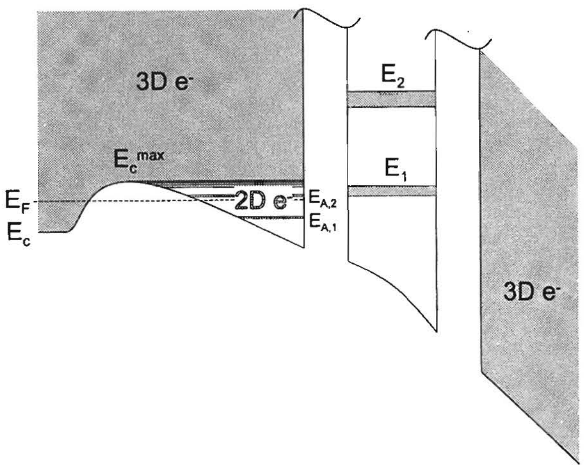


Fig. 4.7 Detail of the accumulation and barrier region band diagram.

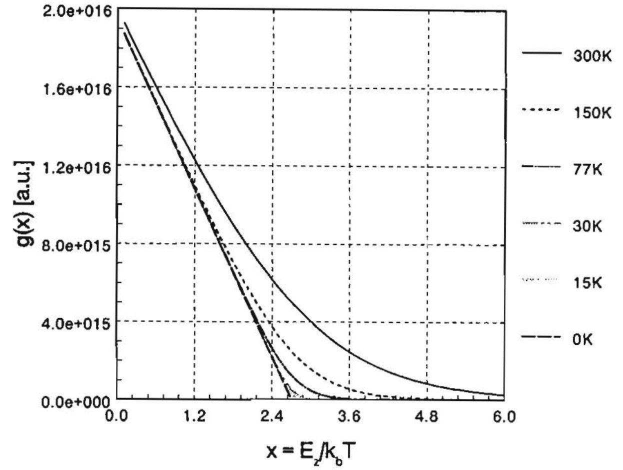


Fig. 4.8 g(E_z) function for E_F = 70 meV and different temperatures.

taken the temperature as a parameter. As observed $g(E_z)$ monotonous decreases with energy and monotonous increases with temperature for all energies. Changing bias conditions $g(E_z)$ is changed because E_F is changed and the emitter quantum well is modified, resulting the change of the position of the quasi-levels, both versus E_C and E_F. As known, the transmission probability (T_{tot}) peaks sharply at electron longitudinal energies corresponding to each quasi-level in the quantum well DBRT structure. Also, the quasi-levels have a certain energetic width (Γ_T), which corresponds to the full energetic width at half maximum of the transmission probability. Γ_T is determined by the intrinsic characteristics of the active region (geometry and materials) (Γ_n) and the scattering phenomena in the well (Γ_s) (see section 2.2). The peak of T_{tot} decreases with bias (the increase of the energetic asymmetry) [11] while Γ_T increases because of the increase of Γ_n. The position of the quasi-levels in the quantum well are mainly determined by its width (narrow quantum well higher position as has already shown in Fig. 2.2a), the barrier thickness and height having less influence in this sense. One can say that Γ_n depends on the effective barrier ($V_{eff}^n = V_B - E_n$) seen by the electron with the energy equal with the corresponding quasi-level (see Eq. 2.11 and Fig.2.2b). Therefore the higher levels in the quantum well have larger Γ_n. Also the increase of well width results in the decrease of Γ_n in connection with the above mentioned phenomenon of lowering of the quasi-level position. Finally, as expressed by Eq. 2.11, Γ_n decreases exponentially with the increase of the barrier thickness. Within this picture we can assert that changing the bias conditions the main peak of the current density occurs when the ground state $E_{A,1}$ of the accumulation region quantum well is aligned with the ground state E_1 or the next level E_2 of the active region quantum well, depending on the device structure [12, 13, 14]. Here we point out that the configuration of the accumulation quantum well depends not only of the emitter layer but also on the barrier region in front of it. The position of the peak current density on the voltage axis of the I-V characteristics depends both of the initial position of the E_1 (E_2) in the DBRT structure quantum well and of the evolution of the accumulation region quantum well with bias, which finally gives the position of the $E_{A,1}$ (both levels are referred to a common energy reference level, for example E_c outside the emitter band bending region).

Experimental results show that the peak current decreases with the increase of the barrier thickness and height [5-9] and the increase of the quantum well width [9,10]. This can be explained taking into account that in all the above situations Γ_n decreases. Now, if we consider the Eq. 2.18, an expression for the transmission probability at resonance ($E = E_n$) can be written in the form:

$$T_{tot}^{n,res} = \frac{\Gamma_n}{\Gamma_T} = \frac{\Gamma_n}{\Gamma_n + \Gamma_s} = \frac{1}{1 + \frac{\Gamma_s}{\Gamma_n}}. \tag{4.3}$$

The above equation shows that a decrease of Γ_n results in a decrease of $T_{tot}^{n,res}$, as well, explaining the peak current density behaviour with the change of the DBRT structure parameters. The current density manifests a very high sensitivity to the change of the barrier thickness and height, as can be seen also in the section 4.4 of this chapter.

The lack of symmetry observed in the I-V characteristics of DBRT devices is attributed to factors connected to the growth conditions. Thus, it is known that the interfaces between GaAs and barrier material (AlGaAs or AlAs) are not chemically abrupt and flat, a transition region of 1-4 monolayers [ML], function of the growth temperature, being observed [4, 15]. In this situation we speak about an *effective thickness* of the barriers and/or quantum well. The difference in the effective thickness of the two barriers represents one of the reasons of the asymmetric I-V characteristics [4]. For example, in the case of AlAs barriers the lower barrier can be several ML thinner than the upper barrier [4], which in fact could be available for AlGaAs as well. Another reason, available in the case of the n^{++}-GaAs substrate, is the segregation of the dopant (usually Si) towards the double barrier region. In these conditions we consider that the I-V characteristic asymmetry of the W488 devices is due to the different effective thickness of the barriers and a correlation between this factor and doping segregation results in the reversed I-V characteristic asymmetry of the W115 and W 309 devices.

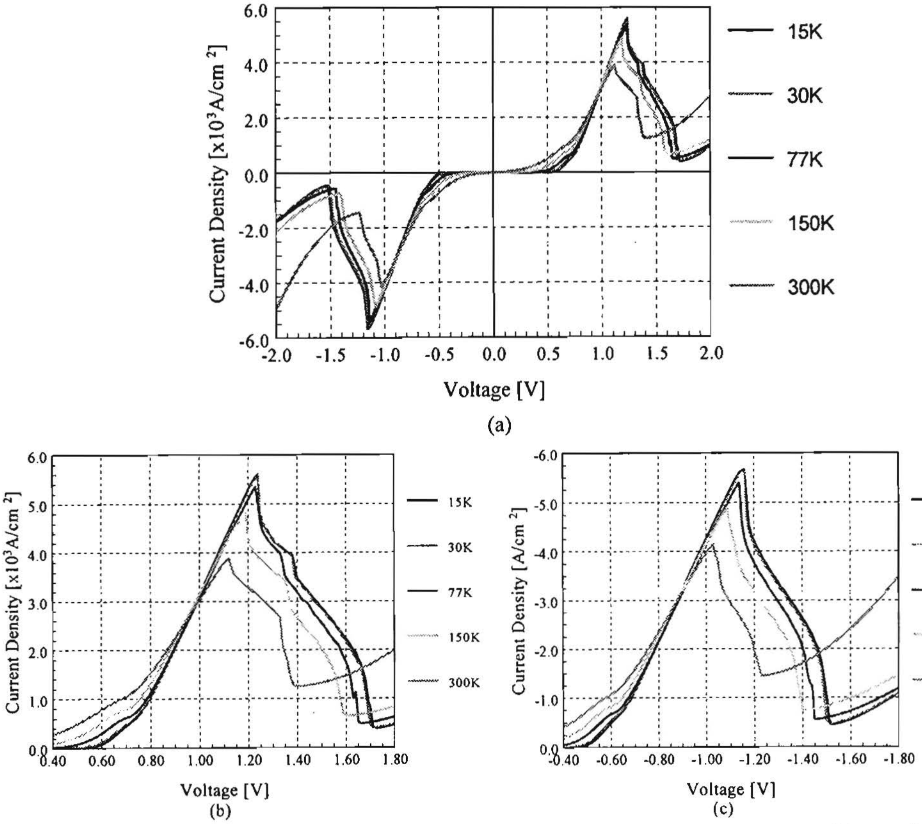
The fact that W115 devices ($Al_{0.4}Ga_{0.6}As$ barriers) manifest in their DC curve a lower J_p than W309 and W488 devices (AlAs barriers) is explained by the much higher barrier thickness (double) of the first. In the same context the larger J_p of the W309 devices versus W488 devices (the same designed barrier thickness) can be attributed to the possible thinner grown barrier thickness of the first - less than 1ML it is enough to get the difference in J_p. Here we have to point out that differences until 1 ML have been observed between the designed and grown barrier thickness [4, 16]. Further on, the same V_p observed for the last two categories of devices in the positive bias regime demonstrate they have the same quantum well width and the lower V_p corresponding to W115 devices shows a different configuration for their accumulation region quantum well.

The undulation observed in the I-V characteristics of all our 2T DBRT devices represents a proof of the quantized levels developed in the accumulation layer under bias conditions [13, 17, 18]. Explicitly, the undulation occurs when the ground quasi-level of the active region quantum well intersects the bottom of the 3D energetic region or the underneath quasi-level of the accumulation layer. This is demonstrated also by its behaviour with temperature (see next paragraph).

b) Temperature Dependent DC Measurements

Our 2T DBRT devices have been DC tested also at temperatures below RT. The measurements have been accomplished at the Institute of Material Physics, Quantum Devices Group from Bucharest. The measurement set-up included a Keithley 236 SMU and a cryogenic system (model LTS-C320) from Cryophysics. The SMU is a programmable system, PC controlled, which can be used both in voltage and current mode. The cryogenic equipment - He circulating system and Lake Shore temperature controller with Si calibrated diode - can reach a temperature of 10K, in medium vacuum conditions (10^{-3}-10^{-4} torr).

We have measured devices processed from W488 sample (AlAs barriers), using the EUT-DU mask set. The devices have been mounted on special packages and wired to the cryostat crossings to ensure the electrical connections. The I-V characteristics for different temperature are shown in Fig. 4.9 and characteristic data are summarized in the Table 4.3. As seen, the devices oscillated



(a)

(b)

(c)

Fig. 4.9 Temperature behaviour of the I-V characteristic for devices processed from W488 sample using EUT-DU mask set.

Temperature	Forward bias (+ on top)					Reverse bias (- on top)				
	J_p x10^3 [A/cm^2]	V_p [V]	J_v x10^3 [A/cm^2]	V_v [V]	$r_{p/v}$	J_p x10^3 [A/cm^2]	V_p [V]	J_v x10^3 [A/cm^2]	V_v [V]	$r_{p/v}$
300	4.13	1.12	1.33	1.41	3.11	4.39	1.03	1.53	1.23	2.87
150	5.16	1.19	0.68	1.6	7.59	5.20	1.09	0.79	1.42	6.58
77	5.67	1.23	0.52	1.66	10.90	5.72	1.14	0.59	1.46	9.69
30	5.96	1.24	0.43	1.72	13.86	5.97	1.16	0.50	1.52	11.94
15	6.00	1.24	0.42	1.73	14.29	6.00	1.16	0.48	1.52	12.50

Table 4.3 Characteristic data of the I-V characteristics at different temperature, for devices processed from W488 sample.

in the NDR region of the I-V characteristics, as in the case of the RT DC measurements, making again difficult the precise determination of J_v and V_v.

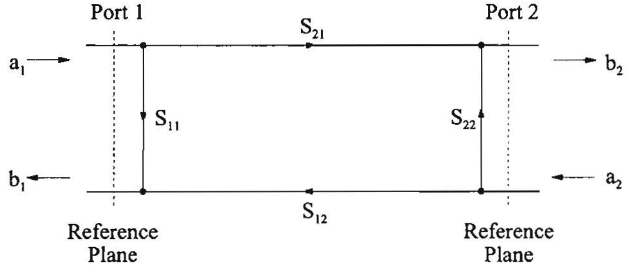
Several important features can be distinguished in the I-V characteristic behaviour with temperature. First, we observe that J_p increases with the temperature lowering and the corresponding V_p shifts to higher voltages. Then, the decrease of the temperature determines lower J_v and higher V_v values. These trends are most evidently in the range 300-77 K. At last, decreasing the temperature, the undulation which precedes the main peak moves to higher voltages and have the tendency to disappear.

The above described evolution of the I-V characteristic with temperature can also be explained in a qualitative manner as we proceed before, in the case of the RT measurements. As we pointed out in chapter 2, carrier scattering phenomena during the tunneling process have an important impact on the transport characteristics of the 2T DBRT devices. Between them, the electron scattering on interface roughnesses, LO phonons and other electrons play the most important role. Scattering processes broaden the width Γ_T of the resonant transmission probability peaks (quantum well quasi-levels), because of the Γ_s contribution, and diminishes their amplitude, $T_{tot}^{n,res}$, (see section 2.2.3). Therefore any change of Γ_s alters the above mentioned two parameters of the transmission probability. The scattering mechanisms are in general very sensitive to the temperature modification. So we expect that the main contribution regarding the temperature behaviour of the DBRT process is due to scattering. The decrease of the temperature means less scattering, i.e. a lower Γ_s, which, considering Eq. 4.3, determines a higher $T_{tot}^{n,res}$. Besides scattering, some other factors can be considered in discussing the temperature behaviour of the I-V curve. Thus, the decrease of the temperature results in the increase of the Fermi level (E_F)[12, 19] and electron m^* [20]. Both of these parameters and the temperature explicitly influence the $g(E_z)$ function (see Eq.4.2), which we have defined in the previous paragraph in connection with the carrier supply in the tunneling direction. Concerning the accumulation layer not only the distribution of the carriers ($\sim g(E_z)$) is changed with temperature but also the configuration of the 2D electron region (see Fig. 4.8): the temperature decrease lowers E_c^{max} and the position of the quasi-levels, versus E_F and E_C outside the emitter band bending region. In these conditions, the total effect of the temperature lowering is that the carrier population decreases drastically in the 3D energetic region of the accumulation layer but increases in the lowest subband $E_{A,1}$ of the 2D region and both the bottom side of the 3D region and the position of $E_{A,1}$ shifts to lower energy. The change of m^* alters also $T_{tot}^{n,res}$ in opposite sense and in a higher degree than $g(E_z)$ [19]. Within the above description we consider that the increase of J_p with temperature decrease is due to the concerted effect of the rise of $T_{tot}^{1,res}$ and the electron population corresponding to $E_{A,1}$ quasi-level in the accumulation layer. The fact that the most change happens between 300 K and 77 K suggests that phonon scattering mechanism is mainly involved in the altering of $T_{tot}^{1,res}$. Then, the position of the main peak shifts to higher biases because of the $E_{A,1}$ descending. Concerning J_v, this decreases at lower temperature mainly due to the decrease of the width of the first quasi-level in the active region quantum well. Finally, the evolution of the undulation is caused by the reduction of the population in the bottom side of the 3D energetic region and/or subsequent quasi-levels and the above mentioned lowering of their position.

4.3.2 AC Measurements

As we mentioned in the first chapter much of the interest in DBRT devices has stemmed from the fact that they can provide a fast room temperature NDR region in their current-voltage characteristic, which can be used as the basis for high-frequency applications. Therefore, the AC measurements play an important role in the device characterization offering information about the intrinsic electrical properties involved in the low and high frequency behaviour. Based on the considerable experience and large facilities existing in the field of microwave measurements within the Electronic Devices Group of EUT a comprehensive AC characterization of our 2T DBRT devices has been carried out, which can not be found in similar approaches communicated until now [21, 22, 23].

A common and direct procedure to determine the AC properties of the 2T (one-port) or 3T (two-port) microwave devices is to measure their S-parameters using a network analyser. A schematic representation of a two-port device measurement and the definition of the S-parameters are presented in Fig. 4.10. S_{11} and S_{22} have the character of reflection coefficients and can be



Forward Parameters:
$S_{11} = b_1/a_1$
$S_{21} = b_2/a_1$

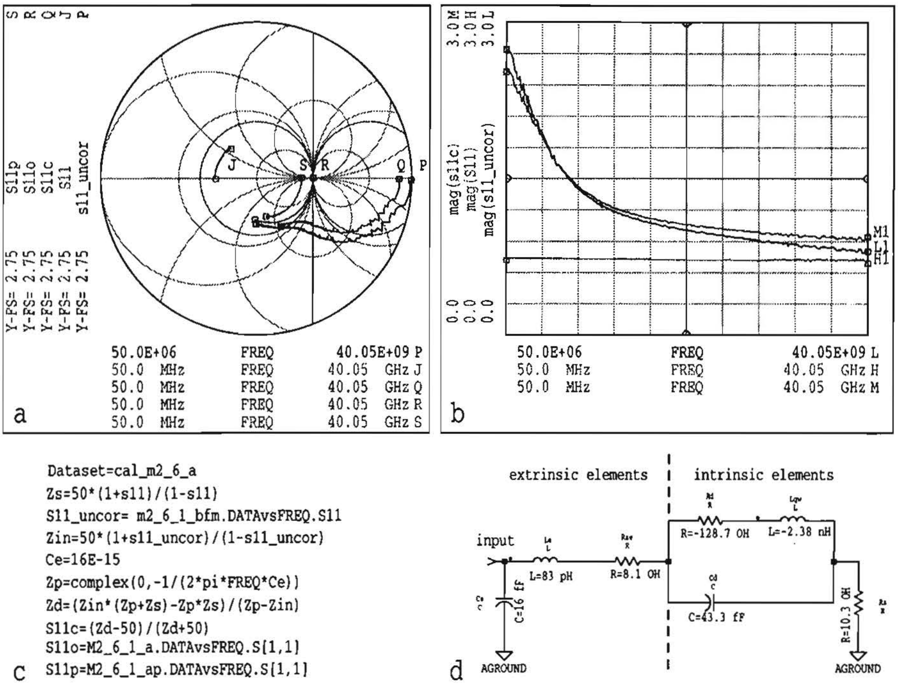
Reverse Parameters:
$S_{21} = b_1/a_2$
$S_{22} = b_2/a_2$

Fig. 4.10 Schematic representation of a two-port microwave device measurement. In the definitions of the S-parameters a_i and b_i represent normalized complex amplitudes.

represented in a Smith chart [24]. The other two S-parameters, S_{12} and S_{21}, have the meaning of transmission coefficients and are often plotted in the same diagram, although this is an improper use of the Smith chart since the impedance coordinates have no meaning. In the case of non-linear devices (ports), as are also the DBRT devices, the S-parameters are used in the small-signal approximation, i.e. small AC quantities are superimposed on the DC quantities [24]. The S-parameters can be converted to impedances or admittances.

The 2T DBRT devices have been tested in microwave conditions using an HP 8510B network analyser, in a one port configuration. In this way only one S-parameter is involved, the reflection coefficient S_{11}. The measurements have been accomplished on-wafer, using Cascade probes, thanks to the suitable design of the pad configuration, both on the EUT-DU and our designed masks (see this chapter and chapter 5). This avoided many complications regarding the subtraction of the parasitic contribution of the package or test fixture used in the past for microwave DBRT device characterization [25]. Even in this situation the measurement of the S-parameters with the network analyser is quite complex and we do not describe it here, being out of the aim of this work. Details about the measurement procedure can be found elsewhere [24, 26]. The network analyser data have been transferred to a MDS® program within which, using an equivalent circuit for the measurement configuration, correction and optimization operations have been accomplished. The most of the measurements have been done on devices processed from the W488 wafer using the EUT-DU masks, with an active area of 6x6 μm^2. The microwave tests have been performed between 50 MHz and 40 GHz.

The entire procedure developed for the AC characterization of our 2T DBRT device involves many steps. For their description we refer to Fig. 4.11. First, the S_{11} has been determined for a shorted pad configuration similar with that of the device under test (DUT) (curve J in Fig. 4.11a). In this way the impedance corresponding to the inductance L_e and resistance R_{se} (see Fig. 4.11d) can be specified (Z_s in the equation description of the Fig. 4.11c). Also the values of these *extrinsic* circuit elements can be determined by matching the experimental S_{11} data over the whole frequency band with those calculated from the equivalent circuit, using the MDS program. As seen in Fig. 4.11d apart L_e and R_{se}, which correspond to the device contact-pad interconnections, the DUT extrinsic circuit also contains a parallel capacitance C_e between the measurements pads (see Fig. 4.2). The value of this capacitance is found in a similar way as that described above for L_e and R_{se} but using a open pad configuration. The values of the extrinsic elements are considered to be constant in all the subsequent DUT measurements, for all the frequencies and applied DC bias ranges. The next step consists in the measure of the S_{11} of the DUT (e.g. curves S, R and P in Fig. 4.11). This has been carried out for an important number of DC bias points (51), covering the whole I-V characteristic of the device, i.e both positive differential resistance (PDR) regions and the NDR region. It is worth to mention that the NDR region was completely stable in the microwave measurement environment, without any oscillations. For each bias, S_{11} has been measured for the entire frequency range, 0.05-40 GHz. In this way a complete map of the device S_{11} for different biases and frequencies has been obtained. These data could be used further on for the optimization procedure. This means that an equivalent circuit for the *intrinsic* part of the DUT (the device itself) has been also considered (see Fig. 4.11d). Then the intrinsic element values corresponding to the DUT equivalent circuit have been calculated in the same way as before in the case of the extrinsic elements, i.e. within MDS, in the above mentioned frequency range, the measured S_{11} have been matched to values calculated for the DUT equivalent circuit, taking



```
Dataset=cal_m2_6_a
Zs=50*(1+s11)/(1-s11)
S11_uncor= m2_6_1_bfm.DATAvsFREQ.S11
Zin=50*(1+s11_uncor)/(1-s11_uncor)
Ce=16E-15
Zp=complex(0,-1/(2*pi*FREQ*Ce))
Zd=(Zin*(Zp+Zs)-Zp*Zs)/(Zp-Zin)
S11c=(Zd-50)/(Zd+50)
S11o=M2_6_1_a.DATAvsFREQ.S[1,1]
S11p=M2_6_1_ap.DATAvsFREQ.S[1,1]
```

Fig. 11 Microwave measurement results of a 2T DBRT device processed from the W488 wafer using EUT-DUT mask set. (a) S_{11}'s plotted on the Smith chart for the shorted pad configuration (curve J) and DUT for different bias points: 0 V / 0 mA (curve R), steepest point in the first PDR region, 0.849 V / 0.497 mA, (curve S) and steepest point in the NDR region, 1.028 V / 0.777 mA, as measured (curve P) and after corrections for the extrinsic contribution of the pads and interconnections (curve Q); (b) S_{11} magnitude corresponding to the shorted pad configuration (curve H) and the steepest point in the NDR region (minimum of the NDR) as measured (curve L) and after corrections (curve M); (c) mathematical procedure used to make the corrections for the contribution of the DUT extrinsic elements; (d) the equivalent circuit of the DUT, used for the optimization and correction procedures - the intrinsic element values correspond to the steepest point in the NDR region.

always constant the extrinsic element values determined previously. Some comments about the intrinsic equivalent circuit of the DUT are necessary. As shown in Fig. 4.11d, apart from the differential resistance R_d due to the resonant tunneling process and the capacitance C_d accounting for the space charge in the device active region (considered also in the circuit modeling of the Esaki tunnel diode), the circuit includes an intrinsic self-inductance L_{qw} in series with R_d. This so called *quantum inductance* is related to the physical parameters by $L_{qw} = \tau_n R_d$, where τ_n is the lifetime of the quantum well quasi-state involved in the tunneling process (for example the ground-state). The above relation is established assuming that the tunneling current varies exponentially with a time constant τ_n when the device bias is suddenly changed [27]. Intuitively, an inductive reactance might be expected in any circuit model of the resonant tunneling devices since the time required for the buildup or decay of the wave function in the quantum well leads to a delay of the current with respect to voltage. It is worth noting that many AC analysis of the 2T DBRT devices have been done without including the quantum inductance [21, 23, 28, 29], just ignoring it or considering that the effect of the quasi-bound state lifetime (L_{qw}) becomes significant when this is higher than the RC time constant (τ_{RC}) of the simple parallel R_d-C_d circuit [27], e.g. for thick barriers. What we want to stress here is that our device analysis in microwave conditions demonstrated clearly that *the use of the L_{qw} results in a perfect match between the measured and calculated S_{11}'s for all the biases and over the whole frequency band (0.05-40 GHz)*, which otherwise can not be obtained [26]. The remaining component in the intrinsic equivalent circuit is the series resistance R_s that arises from ohmic dissipation, due to contact resistance, device structure resistance, spreading resistance, etc..

Fig. 4.12 shows results of the element values for the intrinsic equivalent circuit at different DC biases, obtained after the previous described optimization. On each plot beside the representation of the circuit element versus voltage we have included also the corresponding I-V characteristic that may be helpful in doing easier different useful correlations. In Fig. 4.12a we preferred to represent the differential conductance $G_d = 1/(R_d+R_s)$ versus voltage instead the $R_d(V)$, which presents spiky behaviour in the critical bias points (peak and valley), because of limited number of voltage points used to obtain the I-V characteristic. Further on, we make some remarques concerning the behaviour of these elements with bias. As seen in Fig. 4.12a there is a perfect fit between the differential conductance obtained from the DC I-V characteristic differentiation and that resulted from the optimization procedure using the AC data. This means that we can trust the accuracy of both the network analyser measurements and the optimization procedure within the MDS program. Concerning the $C_d(V)$ curve (see Fig. 4.12b) we mention that the peaked behaviour around the maximum (negative) NDR point, observed when an intrinsic equivalent circuit without L_{qw} is used [23,26], does not appear in our case, the device capacitance being more uniform along the considered bias range. We specify also that the undulation observed in the I-V characteristic, before the main current peak, corresponds to specific behaviours (undulations) in both $G_d(V)$ and $C_d(V)$ curves. If C_d does not change notably with bias, on the other hand L_{qw} demonstrates a strong variation in the NDR region of the I-V characteristic (see Fig. 4.12c), where it takes negative values. This means that in this region the intrinsic impedance shows an obvious inductive character consistent with the physical interpretation introduced previously, $L_{qw} = \tau_n/G_d$ with lifetime values around $\approx$18.5 ps. Inside the NDR L_{qw} goes asymptotical to minus infinity for the voltage points corresponding to peak and valley current and reach a maximum (negative) value when the G_d is minimum (negative). The absence of inductance (very low values) in the PDR regions indicates a much shorter resonant-tunneling traversal time (τ_n) than that in the NDR region. One possible explanation is that the scattering reduces significantly the quasi-level lifetime [30]. The last plot in Fig. 4.12 presents the intrinsic R_s versus voltage. The nonconstant behaviour could have many sources. First, we have to consider that R_s, as mentioned above, represents the result of the contribution of many resistances in series. The contact resistance is not voltage dependent, but we can expect that the change of the space charge profile within the multilayered device structure during biasing may result in a voltage dependence of the corresponding resistance. Then, in the optimisation process R_e is considered constant for all biases. This is not totally true because during the measurements the contact

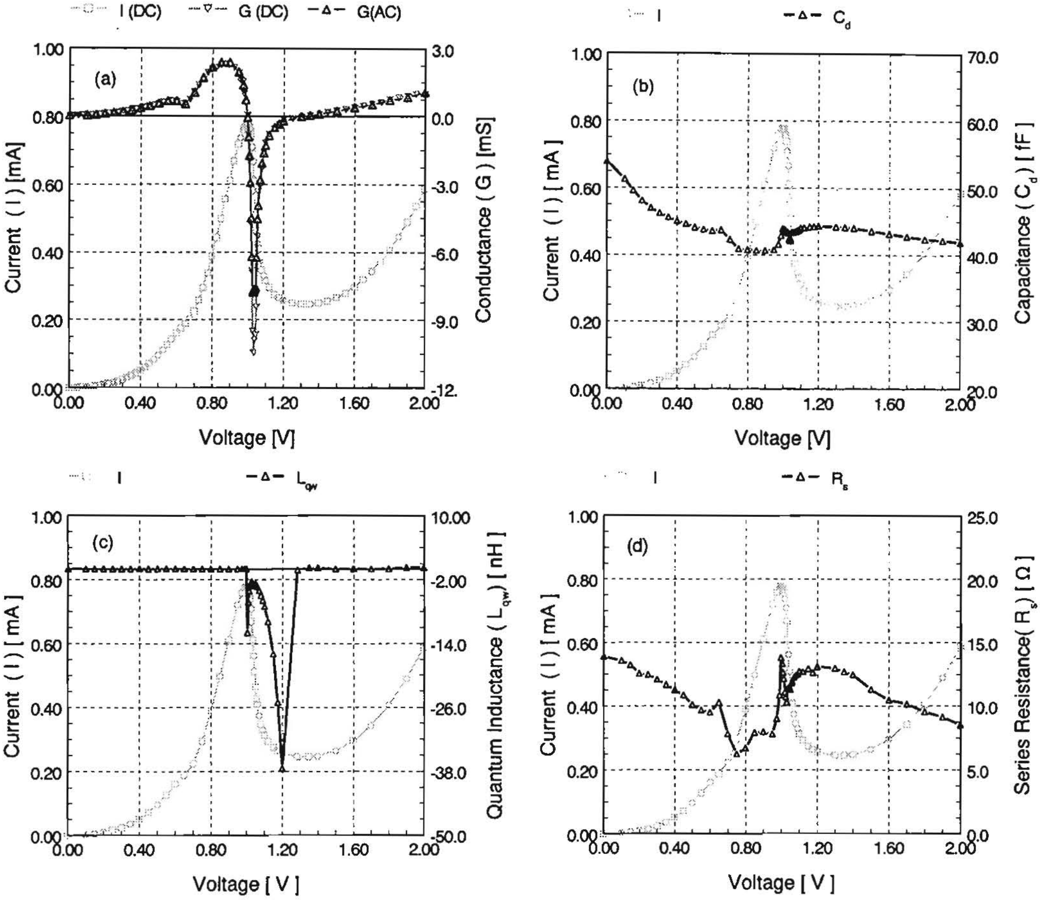


Fig. 4.12 Intrinsic elements of the DUT equivalent circuit, G_d (a), C_d (b), L_{qw} (c), and R_s (d), versus voltage as result of the optimization process using MDS program.

resistance between the pads and the Cascade probes can vary as a result of the different pressure the probes stay on the wafer. At last, in the I-V characteristic regions with low slope R_d is much higher than R_s, resulting in a less accurate determination of the intrinsic series resistance.

From the application point of view an important parameter of the 2T DBRT devices is the maximum oscillation frequency, f_{max}. If such a device is DC biased in the NDR region, it is capable of oscillating up to a frequency at which the magnitude of the reflection coefficient becomes less than one, i.e. the device is not anymore an active electronic component. The maximum oscillation frequency is obtained when the device is biased in the steepest point of the NDR region ($|R_d|$ is minimum). For our DUT we can determine the f_{max} from Fig. 4.11b. In this figure the magnitude of the S_{11} is plotted versus frequency for the extrinsic DUT elements (curve H), DUT (curve L) and after corrections (curve M). The correction procedure consists in the subtraction of the extrinsic element contribution (as seen previously obtained from separate measurements) from the DUT impedance using the equation sequence presented in Fig. 4.11c. In this way the reference plan for the measurements is moved from the pads, where the Cascade probes are positioned, to the device contact level (see Fig. 4.11d). On the corrected plot of the S_{11} magnitude f_{max} corresponds to 26.25 GHz, a value comparable with that reported for a similar device structure [28].

An analytical expression for the limit frequency at which a 2T DBRT device oscillates when biased in the NDR region can be obtained using an equivalent condition with that considered above for the reflection coefficient: the real part of the impedance corresponding to the intrinsic DUT equivalent circuit equals zero. This condition gives for the limit or *resistive cut-off* frequency f_r the following expression:

$$f_r = \frac{1}{2\pi\sqrt{L_{qw}C_d}}\left\{\left(1-\frac{C_d R_d^2}{2L_{qw}}\right)\left(1-\sqrt{1-\frac{1+R_d/R_s}{\left[1-C_d R_d^2/\left(2L_{qw}\right)\right]^2}}\right)\right\}^{1/2} \tag{4.4}$$

If L_{qw}, C_d and R_s are taken in the steepest point in the NDR region (see Fig. 4.11d), i.e. $|R_d|$ is minimum, from Eq. 4.4 results the device f_{max}. In our concrete case, using the intrinsic element values from Fig. 4.11d we get a f_{max} = 24.41 GHz, very nearly to the value obtained from the microwave measurements.

4.4 Modeling

A physical model to calculate the static characteristics of the 2T DBRT devices has been developed within the Electronic Devices Group of EUT [31-35]. The model is based on the steady-state transport picture already described in chapter 2 and in the previous section of this chapter. This means that the device can be considered to consist of the DBRT structure itself and two independent electron reservoirs. Under bias one reservoir will be accumulated and the other one depleted, the electrons tunneling mainly from the first to the second. Although the main feature of the transport through the device is determined by the DBRT structure, the accumulation and depletion regions have a considerable influence on the device characteristics because the former provides the supply of the tunneling electrons andthe latter absorbs most of the bias voltage. In fact the model is constructed from distinct models for the three regions separately. They combine the semi-classical and quantum mechanical treatments. The models are connected together by demanding continuity of the electric field and the electron wavefunctions. In the accumulation region is taken into account that part of the supplied electrons resides in 3D states and part in 2D subbands. The modeling of the barrier region includes the partial incoherent character of the electron tunneling due to interface roughness scattering. Also the effect of the charge build up in the quantum well of the DBRT structure over the potential configuration along all the three regions is selfconsistently considered.

We start this section with a short description of the model, for details making reference to the main sources [23-25]. Then, results of the model which confirm the above analysis of the experimental measurements are presented, completing our understanding over the subject of this chapter.

4.4.1 Model Description

a) The Accumulation Region

As we have pointed out previously, due to the doping profile, under bias conditions, the conduction band in the accumulation region has a maximum ($E_c^{\max}$) and a potential well is formed in front of the first barrier (see Fig. 4.7). Although the situation resembles that in NMOS or HEMT channels, the modeling is more complicated in this case. Here the carrier transport which determines the device characteristics is not parallel with the layer interfaces but perpendicular on them (parallel with the quantization direction). Therefore, the 3D electrons play a much more important role in a DBRT diode than in the other mentioned devices. In the same context, the height and position of the potential energy maximum changes with bias and they have to be determined selfconsistently (for each bias) together with the energetic position and occupation of the subbands. In these conditions the modeling of the accumulation region is based on the extension of an accurate and efficient selfconsistent method which has been developed earlier to study the quantum effects in the inversion layers [36,37], adapted to describe the specific situation of the 2T DBRT devices.

The calculation starts by assuming a value for the Fermi level , taken constant throughout the accumulation region. This is based on the assumption that the tunneling current does not appreciably disturb the electron concentration in this region and consequently it is in quasi-equilibrium conditions. In a first approximation the 2D character of the electrons which could reside in the quasi levels of the accumulation region is neglected and a semi-classical local density approach is used in which Poissons's equation is solved iteratively (successive relaxation method) using the electron concentration calculated as if all of them obeyed a Fermi-Dirac distribution. In this initial *3D analysis* for electrons plane wave functions are used taking into account the actual reflection coefficient at the interface with the barrier region [35]. As soon as the potential profile is known the 3D electron concentration is extracted from the total concentration considering that the classical energetic border line separating the 3D and 2D electrons is at the $E_c^{\max}$ level.

Further on a model potential (analytic approximation) is constructed in which the influence of the 3D electrons and doping effects on the 2D electrons is considered [35]. Using this model potential, the subband energies , the wave functions of the 2D electrons and the 2D electron concentration are found solving together the Schrödinger and Poisson equations, with a reduced self-consistency requirement [35]. All the procedure represents in fact the *2D analysis* within the accumulation model and is similar with the above mentioned approach of the quantum effects in inversion layers [36, 37].

Making distinction between the 2D and 3D electrons, the potential profile and 3D electron concentration are calculated again. This time, however, the local density approach applies only to 3D electrons, whereas the 2D electron concentration substituted in the Poisson equation results from the 2D analysis.

When the above calculations are performed the potential profile, the electron concentration, the subband energies and the wave functions in the accumulation region are obtained. Using these quantities, the barrier and depletion regions can be further on analysed.

b) The Barrier Region

The analysis of this region has to result in the calculation of the transmission probability for a certain electric potential profile along the double barrier structure. The trapezoidal barriers and well are replaced by a small number (1...5) of rectangular steps resulting a piecewise constant potential. Then the calculation is carried out using the transmission matrix method, in the same way which we have described in the Appendix A, section A1. The effect of scattering due to interface roughness is included, the incoherent parameter being extracted from an emirical model described in [33]. Because the electron charge build up in the quantum well changes in its turn the potential profile of the region, increasing the electric field at the interface with the depletion region, a selfconsistent procedure in connection with the analysis of the later is considered. The sheet charge in the well is assumed to be concentrated in the centre of this and is calculated from the square of the coherent and incoherent wave functions integrated over the well.

c) The Depletion Region

In this region there exists a mix of two completely different electron populations: on the one hand electrons are injected from the heavy doped contact layer and diffuse backwards against the electric field and on the other hand there are the electrons that tunnel through the barriers, arriving with high kinetic energies. Therefore a proper analysis of this region could only be done by a Monte Carlo simulation. Fortunately, the density of the tunneling electrons is very small compared with the doping, so their influence on the space charge is quite unimportant. In these conditions the latter is mainly determined by the doping and the electrons diffusing back from the contact, which behave as if they are in an quasi-equilibrium state. The consequence of this situation is that one can treat the depletion region in the same way as it has been proceeded in the first approximation for the accumulation region (3D analysis), i.e. applying a local density approach. Due to the large width of the depletion region there are no quantum effects (2D electrons) in this case.

d) Calculation of the Current-Voltage Characteristic

For a given value of the Fermi level, using the above described models, the current flowing through the device and the voltage drop along the whole structure can be finally calculated. The calculation of the tunnel current density corresponding to the 3D electrons has been discussed previously in chapter 2, section 2.3. Here we only make several specifications. In the model, both the current from the left to the right and that from the right to the left are calculated considering the states where the electrons tunnel empty(the factors $(1 - f_{r,l}(E))$ are 1). Then, the current from the right to the left is significant only at near-zero bias, at higher voltages becoming negligibly small. At last, since the 2D subbands are taken into account in the electron supply region, in the expression of the current density from left to the right the integration has to be carried out over a reduced energy range, from the energy maximum, $E_c^{\max}$, upwards (see Fig. 4.7).

Calculating the tunnel current from the accumulation region quasi-levels poses a problem because a standing wave in the tunneling direction cannot carry any current. In reality the electrons do tunnel, however, and the loss of electrons from the subbands is compensated by scattering of electrons from higher subbands or 3D states. To overcome this problem several approximations are made [35]. First, the 2D wave functions resulted from the 2D analysis are spatial translated over some distance into the barrier. The distance is chosen such that the values and the derivatives of the 2D wave functions match with damped functions in the barrier at the interface. Then, assuming that the tunnel current is relatively small and does not disturb the electron distribution significantly another approximation is introduced replacing the *true 2D wave function*, ψ_i, found in the previous step by an *effective plane wave*, $\psi_{eff,i}$, expressed as linear combination of plane waves, right and left travelling (see for example Eq. A.3 in the Appendix A). The coefficients (amplitudes) of the combination terms are chosen such that the effective plane wave and the true 2D wave function coincide at the barrier and the corresponding effective wave reflection coefficient matches the wave reflection coefficient of the structure, determined in the 2D analysis. In these conditions, the current density contribution of the 2D electrons in the subband i can be written [33]:

$$J_{sub_i} = \frac{2qk_BT}{8\pi\hbar} \cdot \frac{k_{sub_i}}{d_{sub_i}} T_{tot}(E_{sub_i}) \ln\left[1 + \exp\left(\frac{E_F - E_{sub_i}}{k_BT}\right)\right], \tag{4.5}$$

where $d_{sub_i} = \left\{\left[\psi_{eff,i}(z_{interf.})\right]^2 + \left[\psi'_{eff,i}(z_{interf.}) / k_{sub_i}\right]^2\right\}^{-1}$ and can be associated with an equivalent spatial width of the subband i. Here $z_{interf.}$ corresponds to the position of the interface between the accumulation region and the first barrier.

The total voltage on the device is the sum of the voltages corresponding to the main regions considered above (for example in Fig. 2.4 this is $e(V_A + V_W + V_D)$), which are obtained within the analysis of each region.

Repeating the calculation of the current density and total voltage for a range of Fermi energy values the I-V characteristic of the 2T DBRT device can be obtained.

4.4.2 Results

This section presents some results concerning the modeling of the static characteristic of a DBRT diode using the above described physical model. The results refer to aspects considered previously in section 4.3.1 to explain the 2T DBRT device measurements.

First we have checked how the model explains the behaviour of the I-V characteristic with changing the structure parameters, namely the thickness and height of the barriers and the quantum well width. This is shown in Fig. 4.13. In Figs. 4.13a and 4.13b AlAs barriers are considered and in Fig. 4.13c the barrier thickness is 2.5 nm and the well width is 5.0 nm. As a general remarque we can say that the model reproduces, qualitatively, the experimental behaviour

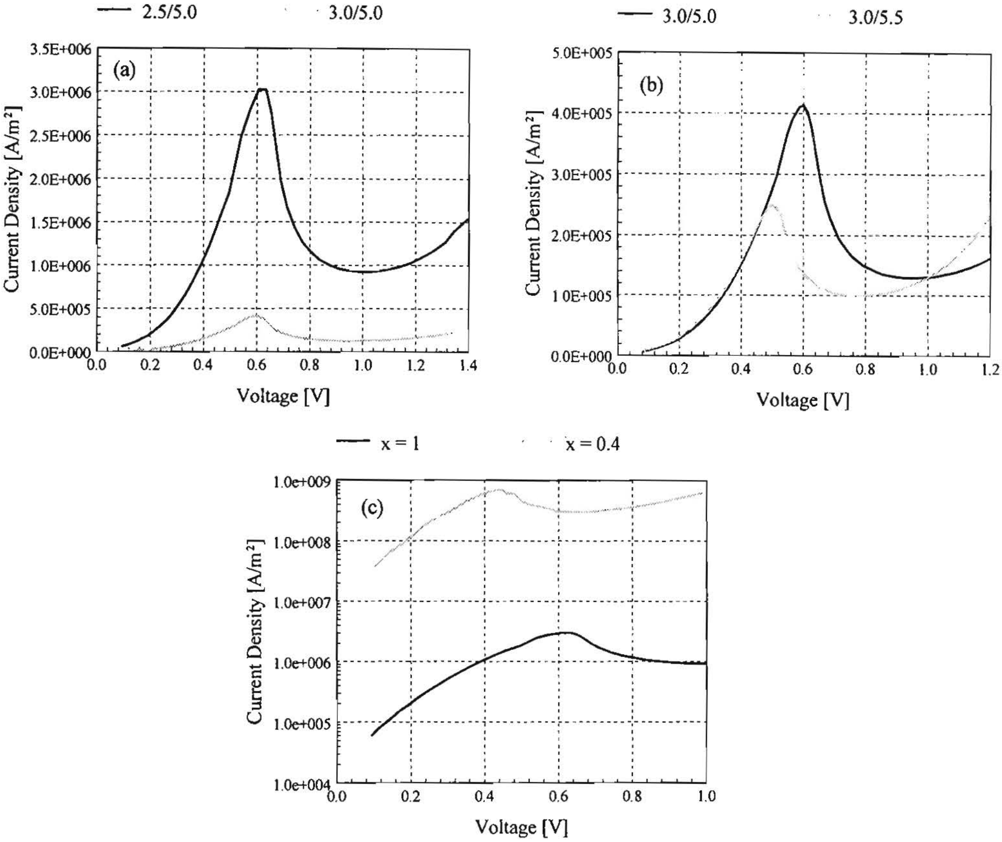


Fig. 4.13 The influence of the structure parameters on the modelled I-V characteristic for the 2TDBRT device. In (a) and (b) the legend indicates "barrier thickness/quantum well width" and AlAs barriers are considered. In (c) the structure has 2.5 nm barriers and 5.0 nm quantum well, x indicating the Al index composition.

in all situations. The modeling results show that the I-V characteristic is very sensitive to the altering of the barrier thickness, whose influence is much higher than the quantum well width modification (see by comparison Figs. 4.13a and 4.13b). We notice that by decreasing the barrier thickness with less than 2 ML, which is in the growth tolerance, the peak current increases with almost one order of magnitude. Also, as expected, the change of the quantum well width is reflected in the shift of the peak voltage because the quasi level position in the well is changed. A small voltage peak shift with barrier thickness change is observed as well, caused most likely by the modification of the accumulation region energetic configuration. The altering of the barrier height has the same high influence on the I-V characteristic behaviour (see Fig. 4.13c). Changing the Al index composition from x = 0.4 (low energetic height barrier) to x = 1 (high energetic height barrier) the peak current decreases with more than 2 orders of magnitude. For this reason

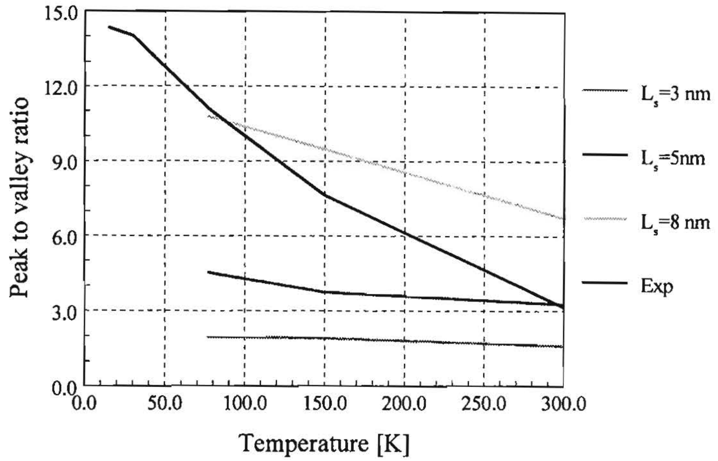


Fig. 4.14 Peak to valley ration versus temperature for different scattering lengths and as obtained from the experiment, for a W488 DBRT structure.

the representation of the corresponding I-V curves is semilogarithmic in Fig. 4.13c. Also, the peak voltage shifts to higher values with the increase of the barrier height, which in this case is determined both by the change of the resonant level position but also by the change of the configuration of the accumulation region quantum well, the latter being dominant.

Another aspect that we considered has been the behaviour of the I-V characteristic with temperature. As we mentioned before the only scattering phenomenon included in the model is the interface roughness scattering. Besides this, of course, other scattering mechanisms have to be taken into account, the most important being the phonon scattering. In this sense, although the roughness scattering is introduced in the model as a temperature dependent phenomenon this is not enough to explain completely the experimental results versus temperature. A fitting parameter used in the roughness scattering model is the scattering length, L_s. Representing the peak to valley ratio $r_{p/v}$ versus temperature for different L_s and in comparison with the experimental results we can already conclude about the existence of other scattering mechanisms (see Fig. 4.14). At room temperature the model curve for L_s = 5 nm (higher scattering) intersects the experimental curve and at 77 K the intersection is happened for L_s = 8 nm (lower scattering), with a more visible temperature dependence of the model curve. From these facts, we can assert that the L_s value corresponding to lower temperature has to be considered and for the complete fitting with the experimental results is necessary to include the phonon scattering. Therefore, taking L_s = 8 nm, we have calculated the I-V characteristics for three temperatures, considering the same structure parameters as those of the real measured device (see section 4.3.1b). The results are shown in Fig. 4.15. Comparing the modeling and experimental results (see also Fig. 4.9) we observe a qualitative agreement concerning the behaviour of the 2T DBRT I-V characteristic versus temperature.

Finally, taken into account the arguments which are used to explain the DC measurements regarding the growth conditions (see section 4.3.1a) we modelled the I-V characteristic of the

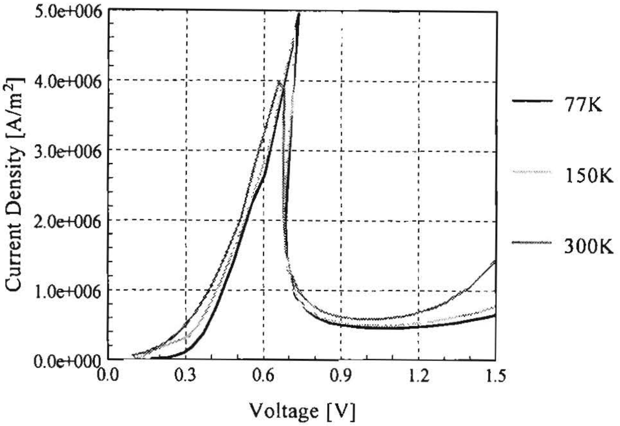


Fig. 4.15 Modelled I-V characteristics for different temperatures for a W488 based structure.

W488 devices (AlAs barriers) considering lower values for the active region layer dimensions and asymmetric barriers, namely the first barrier thickness 1.8 nm, the quantum well width 4.5 nm, and the second barrier thickness 2.1 nm. The modelled and measured results are shown in Fig. 4.16. The agreement between model and experiments is good. The difference observed in the first positive differential resistance region could be caused by the modeling of the accumulation region and the not taking into account of some factors, e.g. phonon scattering, band nonparabolicity, Γ-X coupling, etc.

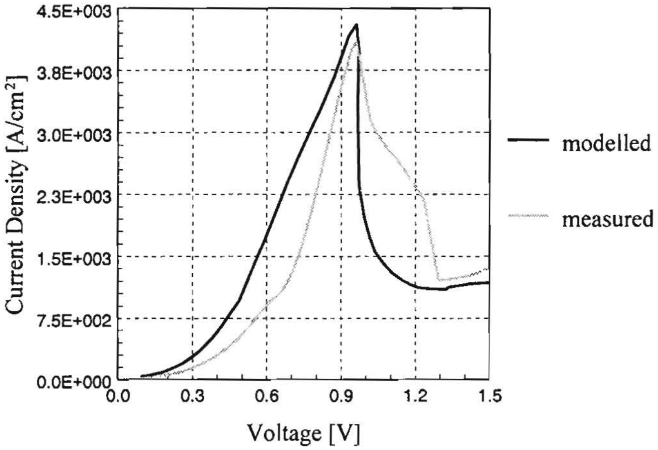


Fig. 4.16 Modelled and measured I-V characteristics for a W 488 based 2T DBRT device. For modeling the active region has the following parameters: first barrier 1.8 nm, quantum well 4.5 nm, and second barrier 2.1 nm (see the text).

4.5 Conclusions

2T DBRT devices are studied comprehensively including fabrication, characterization and modeling. Normal diodes and 2T devices in planar configuration , with both on surface and air bridge interconnections between the top contact and the corresponding pad, have been fabricated from GaAs/AlGaAs MBE grown DBRT structures.

DC measurements carried out on 2T DBRT devices, at room temperature and below, are explained qualitatively, taken into account the impact of several factors: the geometrical and energetical parameters of the device active region, the characteristics of the accumulation region where, energetically, the electrons reside in both 3D and 2D states, the longitudinal electron supply from the emitter and the possible scattering mechanisms involved in the tunneling process. The DC behaviour of the 2T DBRT devices is analysed also using a self-consistent physical model which considers both of carrier types from the accumulation region involved in the device transport process. The results of the model indicate a good description of the I-V characteristic when different parameters of the active region and the temperature are altered.

Accurate AC measurements up to 40 GHz have been made on 2T DBRT devices in stable DC conditions over the whole bias range of the I-V characteristics. It is found that the S-parameter measurements can be described very well, for all biases and over the whole frequency band, by Essaki equivalent circuit model of the device which includes also a *quantum inductance*.

Referring to the 2T DBRT devices processed from GaAs/AlAs based structure, their best room temperature DC performances, $I_p = 4.16\text{x}10^3$ A/cm^2, $V_p = 0.96$ V, $r_{p/v} = 3.42$, and AC performances, $f_{max.}$ =26.25GHz, are comparable with those reported for similar structures.

References

1. H. H. Berger, *Solid-State Electronics* **15**, 145 (1972).
2. R. Williams, "*Modern GaAs processing Methods*", Artech House, Inc., Boston, 1990.
3. G. K. Reeves and H. B. Harrison, *IEEE. Electron Device Lett.*, **EDL-3**, 5 (1982).
4. A. Förster, J. Lange, D. Gerthsen, Ch. Dieker and H. Lüth, *J. Phys. D: Appl. Phys.* **27**, 175 (1994).
5. M. Tsuchiya and H. Sakaki, *Jpn. J. Appl. Phys.* **25**, L185 (1986).
6. M. Tsuchiya and H. Sakaki, *Appl. Phys.* Lett. **50**, 1503 (1987).
7. J. Söderström and T. G. Anderson, *Superlatt. Microstruct.* **5**, 109 (1989).
8. H. Brugger, U. Meiners, J. Schroth, A. Förster, and H. Lüth, private communication.
9. J. S. Wu, C. P. Lee, C. Y. Chang, K. H. Chang, D. G Liu, and D. C. Liou, *Solid-State Electronics* **31**, 577 (1988).
10. M. Tsuchiya and H. Sakaki, *Appl. Phys.* Lett. **49**, 89 (1986).
11. P. N. Racec, T. Stoica, C. Popescu, M. Lepsa, and T. G. van de Roer, to be published.
12. J. Chen, J. G. Chen, C. H. Yang, R. A. Wilson, *J. Appl. Phys.* **70**, 3131 (1991).
13. J. S. Wu, C. Y. Chang, C. P. Lee, K. H. Chang, D. G. Liu, and D. C. Liou,*Appl. Phys.* Lett.

13. J. S. Wu, C. Y. Chang, C. P. Lee, K. H. Chang, D. G. Liu, and D. C. Liou,*Appl. Phys.* Lett. **57**, 2311 (1990).
14. P. Mounaix, J. M. Libberecht, and D. Lippens, *Solid-State Electronics* **38**, 1899 (1995).
15. D. Bimberg, F. Heinrichsdorff, R. K. Bauer, D. Gerthsen, D. Stenkamp, D. E. Mars, and J. N. Miller, *J. Vac. Sci. Technol.* **B10**, 1793 (1992).
16. D. E. Mars, L. Yang, M. R. T. Tan, and S. J. Rosner, *J. Vac. Sci. Technol.* **B11**, 965 (1993).
17. M. J. Paulus, E. T. Koenig, B. Jogai, C. A. Bozada, C. I. Huang, C. E. Stutz, and K. R. Evans, *Superlatt. Microstruct.* **7**, 135 (1990).
18. P. Mounaix, O. Vanbesien, and D. Lippens, *Appl. Phys.* Lett. **57**, 1517 (1990).
19. G. D. Shen, D. X. Xu, M. Willander, and G. V. Hanson, *Phys. Rev.* B **45**, 9424 (1992).
20. J. S. Blakemore, *J. Appl. Phys.* **53**, 9427 (1982).
21. D. Lippens and O. Vanbesien, *Microwave Opt. Technol. Lett.* **2**, 233 (1989).
22. O. Vanbesien, V. Sadaune, D. Lippens, B. Vinter, P. Bois, and J. Nagle, *Microwave Opt. Technol. Lett.* **5**, 351 (1992).
23. T. Wei, S. Stapleton, and O. Berolo, *IEEE Trans. Electron Devices* **42**, 1378 (1995).
24. T. G. van de Roer, "Microwave Electronic Devices", Chapman&Hall, London (1994).
25. M. F. C. Schemmann, H. C. Heyker, J. J. M. Kwaspen, and T. G. van de Roer, *"Mounting and DC to 18 Ghz Characterization of Double Barrier Resonant Tunneling Devices"* Eindhoven University Research Report 89-E-231, ISBN 90-6144-231-1, 1989.
26. J. J. M. Kwaspen, private communication.
27. E. R. Brown, C. D. Parker, and T. C. L. G. Sollner, *Appl. Phys. Lett.* **54**, 934 (1989).
28. T. C. L. G. Sollner, E. R. Brown, W. D. Goodhue, and H. Q. Le, *Appl. Phys. Lett.* **50**, 332 (1987).
29. E. R. Brown, W. D. Goodhue, and T. C. L. G. Sollner, *J. Appl. Phys.* **64**, 1519 (1988).
30. E. R. Brown, *"High-Speed Resonant Tunneling Diodes"* in *"Heterostructure and Quantum Devices"* eds. N. G. Einspruch and W. R. Frensley. Academic Press, London, 1994.
31. T. G. van de Roer, J. J. M. Kwaspen, H. Joosten, H. Noteborn, D. Lenstra, and M. Henini, *Physica B* **175**, 301 (1991).
32. T. G. van de Roer, in *Proc. NATO Adv. Res. Workshop, "Negative Differential Resistance and Instabilities in 2D Semiconductors"*, eds. N. Balkan, B. K. Ridley, and A. J. Vickers, NATO ASI Series Vol. **307**, p. 431, Plenum , New York, 1993.
33. T. G. van de Roer, *Modeling of Double Barrier Resonant Tunneling Diodes: D. C. and Noise Model,* Eindhoven University Research Report 95-E-285, ISBN 90-6144-285-0, 1995.
34. A. M. P. J. Hendriks, *Modellering van het Accumulatiegebied van een Double Barrier Resonant Tunneling Diode*, Thesis EEA 494, Eindhoven University of Technology, the Netherlands (1994).
35. A. M. P. J. Hendriks, W. Magnus, and T. G. van de Roer, *Solid-State Electronics* **39**, 703 (1996).
36. W. Magnus, C. Sala, and K. De Meyer, *J. Appl. Phys.* **63**, 2704 (1988).
37. C. Sala, Ph.D. thesis, Catholic University Leuven, Belgium (1991).

5

Three Terminal DBRT Devices

5.1 Introduction

As we mentioned in the first chapter, one of the main aims of this thesis has been to materialize the idea of fabrication of a three terminal (3T) DBRT device whose third terminal (base) is realized by direct contacting of the resonant tunneling structure quantum well. The main difficulty of making contact to such a thin layer (~ 5nm) as the quantum well has been succesfully surmounted using high selective etching processes. With the help of the third terminal the quantum well potential can be separately altered, resulting in changing of the main I-V characteristic of the device in such a way that important switching properties can be obtained.

One can say that we have approached the problem of the 3T DBRT device fabrication in a total way, based also on the experience accumulated during the DBRT diode processing. In this sense, we have started with the detailed knowledge of the processes intended to be used at the device fabrication, resulted from our own experiments. Afterwards, the design and fabrication of suitable masks have been performed. Finally, the "true" processing, with some inherent corrections has resulted in the obtaining of the 3T DBRT devices. As in the case of the DBRT diode processing now other processes have been optimized, namely the selective etching and the airbridge fabrication using goldplating. All the processes involved in the 3T-DBRT device fabrication have been described in detail in chapter 3. In this chapter we add more a presentation of the mask layout which completes the complex image of the processing. Also, the next section explains concretely the flow of the device fabrication process.

The last section is devoted to device electrical measurements. Mainly, the DC measurements at room temperature and low temperature are presented and analysed. The explanation of the results is done through the device physics point of view. These measurements reveal clearly the positioning of the base contact on the quantum well of the DBRT structure and the influence of the base potential on the I-V characteristic between collector and emitter. Using AC measurements an estimation of the device switching time is accomplished. This kind of determination is much simplified by the presence of the third terminal (quantum well base).

5.2 Mask Layout

For the fabrication of our 3T DBRT devices we have designed and fabricated a specific mask set. The set contains nine layers (masks): three for the mesa definition of the collector (emitter), base and emitter (collector), three for the corresponding contacts, one for the measurement pads and two for the airbridge interconnections between the pads and contacts. A general view of the mask (main cell) layout is shown in Fig. 5.1. On this layout many fields can be distinguished, which are easily identified by specific labels. Thus, besides patterns for the 3T devices there have been included also patterns for the 2T devices, for both AC and DC measurements. The specific design of the pads permits the access with Cascade probes for microwave measurements. For each type of device destined to AC testing, there exist two fields, one with 100 μm distance between

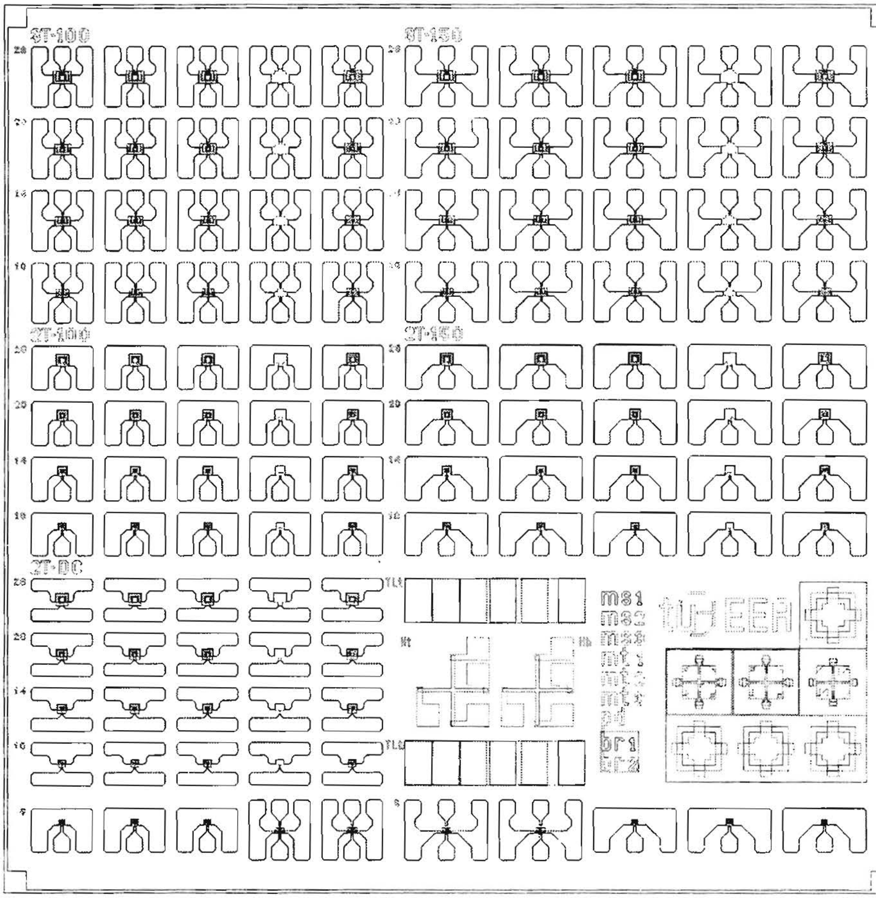


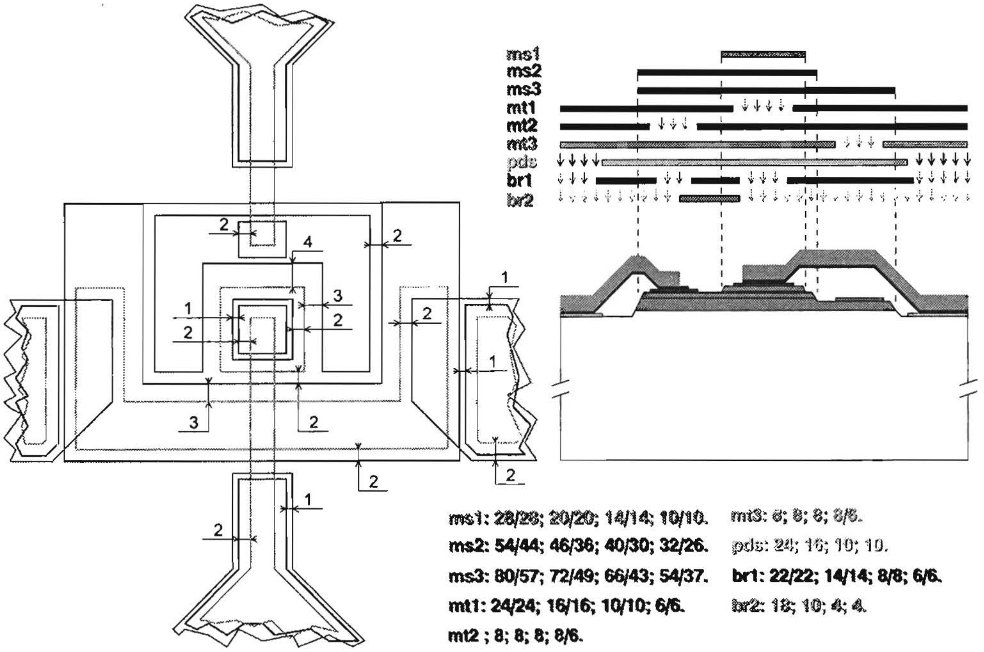
Fig. 5.1 Overview of the mask (main cell) layout for the 3T (2T) DBRT device processing.

the measurement pads and the other one with 150 μm, which corresponds to the Cascade probes available in our laboratory. Both 3T and 2T device fields contain square device patterns with the caracteristic dimensions of 10, 14, 20,and 28 μm. The minimum dimension has been imposed by the use of optical lithography and the airbridge interconnections. Although, in the last row of the layout we have included device patterns with 6x6 μm^2 active area. The first three colons of the device pattern fields contains patterns for active devices and the next two are patterns for test measurements. Besides device pattern fields, the layout contains test patterns for ohmic contact resistance determination and Hall measurements and, of course, the alignment marks patterns.

The layer patterns for a 3T device together with their use during processing are shown in detail, in Fig. 5.2. The data indicate the dimensions (in μm) of the different layer patterns and interspaces.

5.3 Device Fabrication

3T DBRT devices has been processed from the W488 sample (see Table 4.1). The whole processing procedure contains eight main steps. As we did in the case of the 2T DBRT devices we describe these steps in diagrammatic form in Fig. 5.3. Each process in this complex processing scheme has been presented in detail in chapter 3 and the corresponding process recipe can be found in the Appendix B. The proper mask design and the optimization of some other important



ms1: 28/28; 20/20; 14/14; 10/10.
ms2: 54/44; 46/36; 40/30; 32/26.
ms3: 80/57; 72/49; 66/43; 54/37.
mt1: 24/24; 16/16; 10/10; 6/6.
mt2 ; 8; 8; 8; 8/6.
mt3: 8; 8; 8; 8/6.
pds: 24; 16; 10; 10.
br1: 22/22; 14/14; 8/8; 6/6.
br2: 18; 10; 4; 4.

Fig. 5.2 Detail of the layer patterns for a 3T DBRT device. Their use during processing is also sketched. The data are indicated in μm.

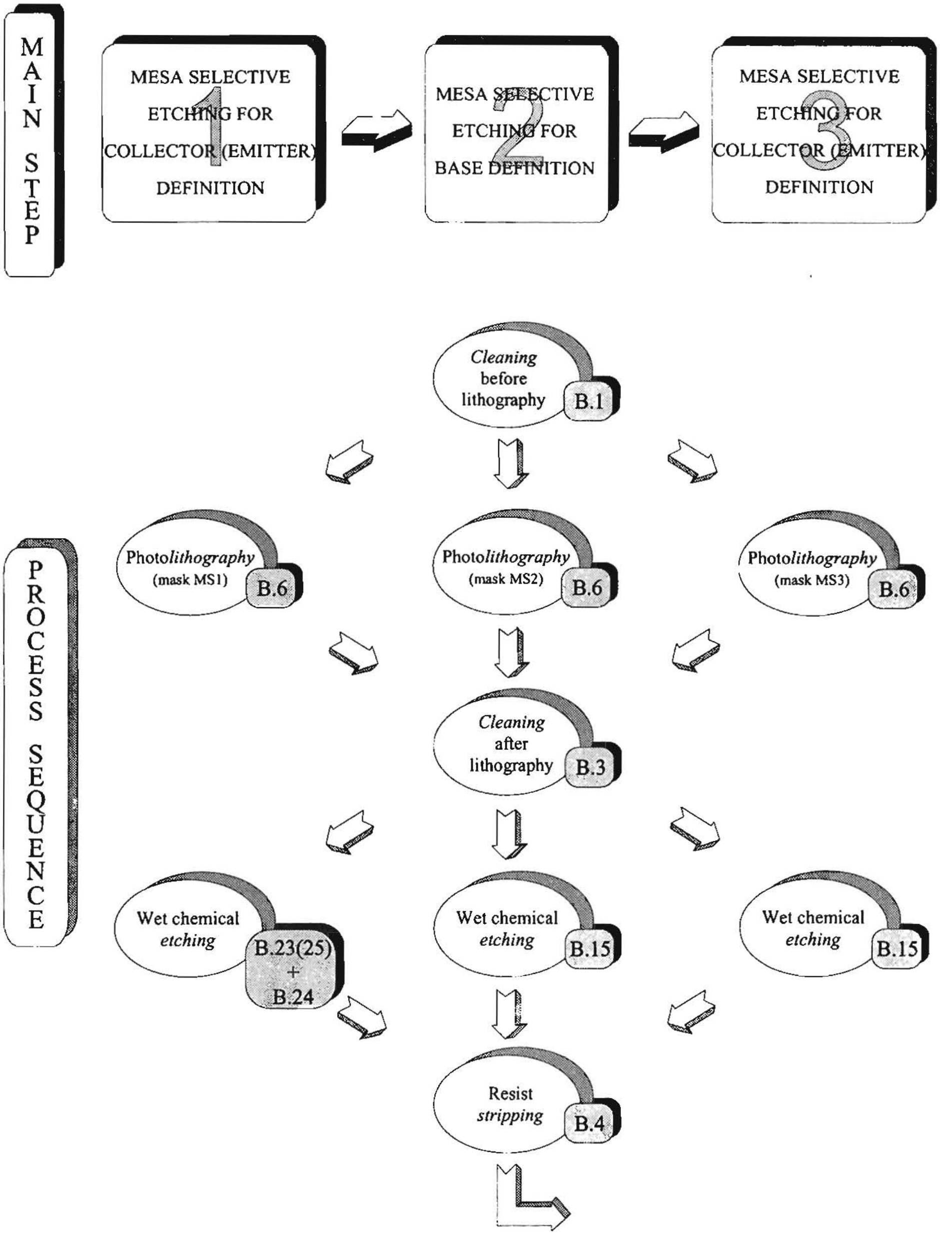


Fig. 5.3 Processsing flow of the 3T DBRT device fabrication. "B.i" denote the process "recipe" detailed in Appendix B.

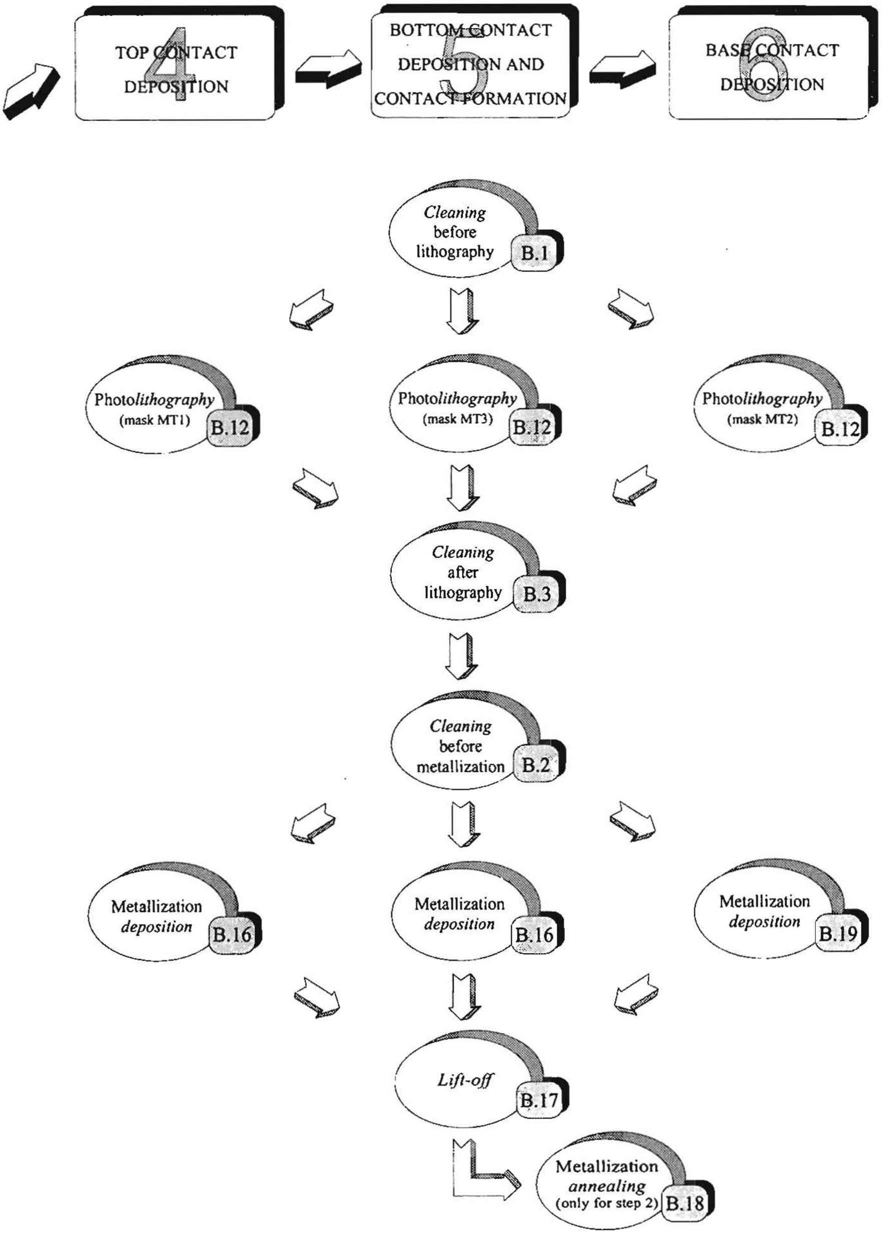


Fig. 5.3 (continued).

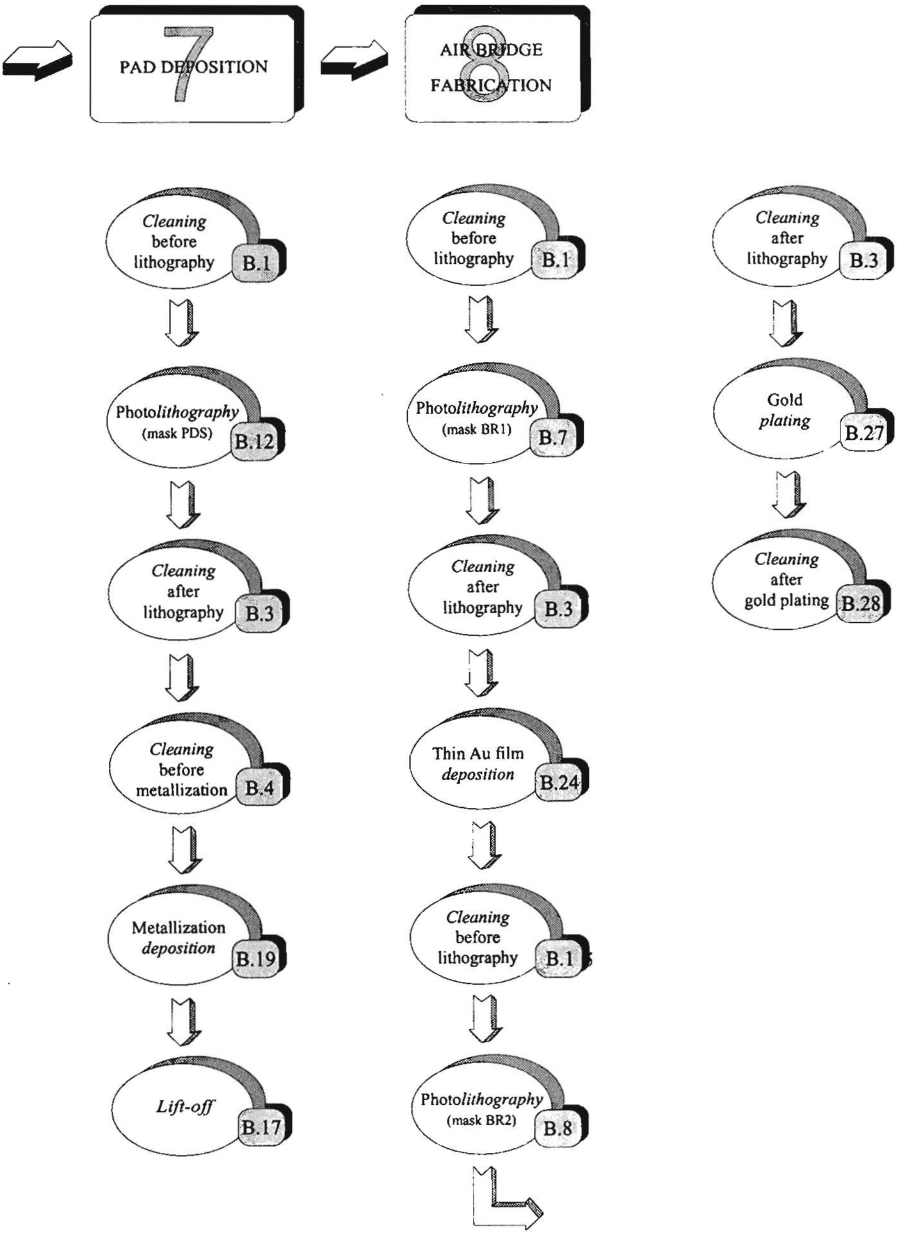


Fig. 5.3 (continued).

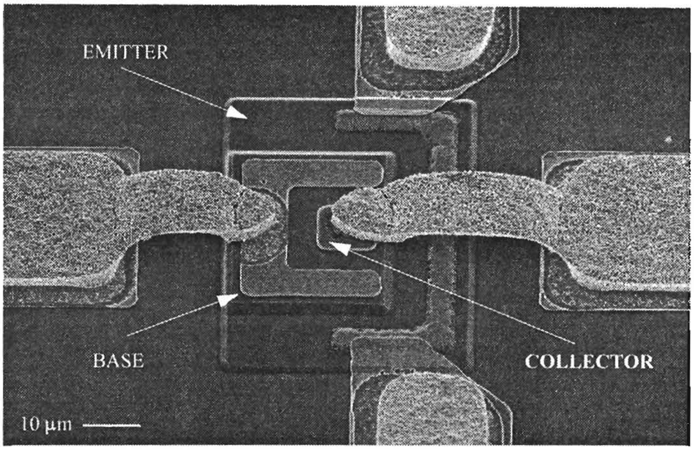


Fig. 5.4 SEM picture of a 3T DBRT device.

processes besides those mentioned at the 2T DBRT fabrication, such as selective etching and air bridge fabrication by gold plating, resulted in the successful processing of our 3T DBRT devices. Fig. 5.4 shows the SEM picture of such a device.

5.4 Measurements

5.4.1 DC Measurements

As in the case of the standard DBRT diodes DC measurements of the 3T DBRT devices represent the most accessible and powerful tool to verify the chosen design and processing options. For a complete investigation of the DC behaviour our 3T DBRT devices have been tested both at room temperature and low temperatures.

The room temperature measurements were accomplished on wafer, with the same measurement setup as for the DC testing of the 2T DBRT devices, e.i. a HP 4141B DC source/monitor system, computer controlled by the IC-CAP 4.3 software package. Each terminal has been connected to a separate source/monitor unit (SMU), thus providing great flexibility, permitting measurements in any device configuration without probe interchange. For low temperature measurements we used the facilities from Duisburg University, Department of Semiconductor Devices. These included a HP 4140B DC source/monitor system, PC controlled, and a vacuum chamber with a microscope table and probes which could be manipulated from outside. The He circulating system connected to the table allowed temperature-dependent measurements between 80 K and room temperature. All the measurements have been done on devices with 20x20 μm^2 active area.

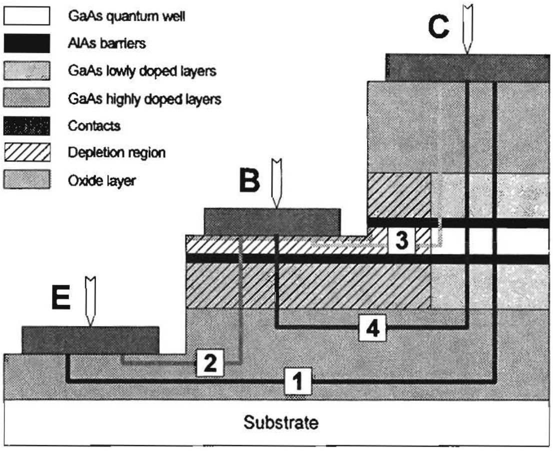


Fig. 5.5 Schematic cross section of the 3T DBRT device. The paths 1, 2, 3 and 4 indicate the possibilities of the current flow.

Fig. 5.5 shows schematically the device cross section, indicating the emitter (E), base (B) and collector terminals. First we analyse separately the two-terminal I-V curves. The biased terminal is indicated by the subscript "bias" and the grounded one by "grd". As we should expect the C_{bias}-E_{grd} I-V characteristic with floating base corresponds to the standard DBRT diode (see Fig. 5.6a).

The discussion of the other two situations, with B_{bias}-E_{grd} and B_{bias}-C_{grd}, necessitates the analysis of the base and underlying regions, before and after the contact deposition. The base surface is reached by selective etching till the quantum well layer of the DBRT structure. With any precautions between the etching step and the contact deposition, we can not avoid the appearance of surface states and the formation of a thin oxide layer. In the case of GaAs the surface states determine the pinning of the Fermi level approximately in the middle of the forbidden band and consequently the formation of a depletion region. In our case the width of the depletion layer is much larger than the base thickness. Moreover, because the low doped layer underneath the base is only 500 Å thick it is very well possible that the depleted zone reaches the following high doped layer. The contact deposition (TiPtAu) alters the number of surface states and their energetic distribution, changing the barrier characteristics. This change is not very drastic and the Schottky behaviour of the contact has little dependence on the used metal. The extension of the depleted region till the high-doped layer has a direct consequence on the barrier properties in electrical conditions. The electric field is constant along it and the barrier width will be independent of the applied bias. We speak in this case of a Mott barrier [8]. The oxide layer, ignored in any other situation , has importance in our case for the explanation of the electrical properties of the base contact. To summarize, we consider that a depletion region is developed under both the contact and the open base surface (see Fig. 5.5) and the contact barrier is the result of three interfacing layers, metal-insulator-semiconductor.

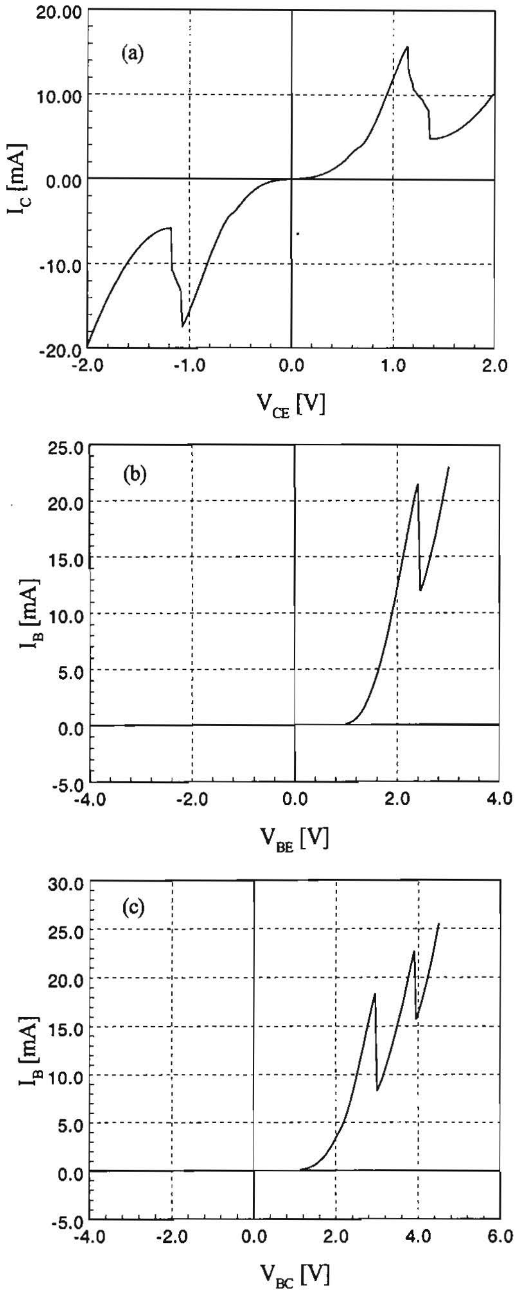


Fig. 5.6 Separate two terminal I-V characteristics of the 3T DBRT device: C_{bias}-E_{grd} (a), B_{bias}-E_{grd} (b) and B_{bias}-C_{grd} (c).

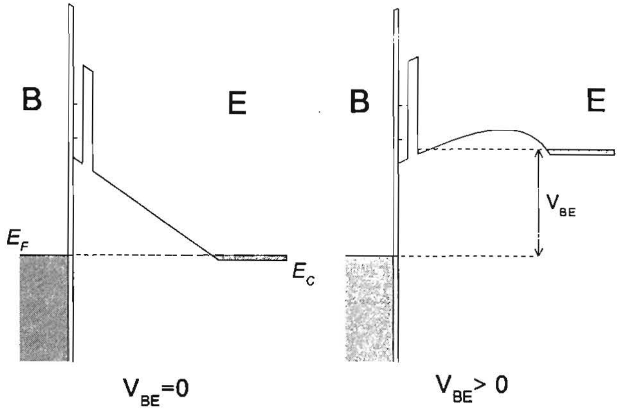


Fig. 5.7 Band diagram representation of the B-E path, without and with bias voltage on the base.

Further we refer to Fig. 5.7 to explain the B_{bias}-E_{grd} I-V characteristics shown in Fig. 5.6b. As one observes in the band diagram, a double barrier structure (DBS) is formed on the top of the Schottky (Mott) barrier, the metal side barrier being formed by the thin oxide layer. We expect then that the I-V curve reflects both the Schottky (Mott) and DBRT characters. This is in fact what is observed. The forward and reverse turn-on voltages of the junction of $V_T^+ \cong 1$ V and $V_T^- \cong -7$ V (not shown), respectively indicate the Schottky (Mott) behaviour of the base contact. Then, the I-V curve presents a negative differential resistance (NDR) region typical for the DBRT structure. One can observe that the peak voltage, measured from the V_T^+, and the peak current have higher values than in the C-E I-V curve. This can be explained by the different characteristics of the "contact" DBS in comparison with those of the "active" DBS from the C-E path. Thus the "contact " DBS has an expected thinner barrier (the oxide layer) and the quantum well is also less thick, which determines a higher position of the resonant levels. Moreover the whole "contact" DBS lies in a depleted region and we also expect the emitter properties to be different.

As one observes in Fig. 5.6c the B_{bias}-C_{grd} I-V characteristic presents at the beginning the same Schottky behaviour and then two NDR regions in the forward direction. An explanation can be found if we refer to the possibilities of the current flow represented in Fig. 5.5 by the paths 3 and 4. The path 3 is practically inaccessible because of the large depleted zone in the direction of the device quantum well (under the open base surface). The other possibility, using path 4, is due to the extension of the electric field till the high-doped layer and having less resistance this will be the preferred way for the current flow. This path corresponds to the situation of two DBRT diodes connected in series, from which results the two NDR regions of the I-V characteristics. The first NDR region is associated with the "active" DBS and the second one with the "contact" DBS.

The low-temperature measurements of the B-E and B-C I-V curves show us that this is not the whole "story" regarding the DC behaviour of the B-C path (see Fig. 5.8). At temperatures of 180 K and below we distinguish three current peaks in the B-C I-V characteristics.

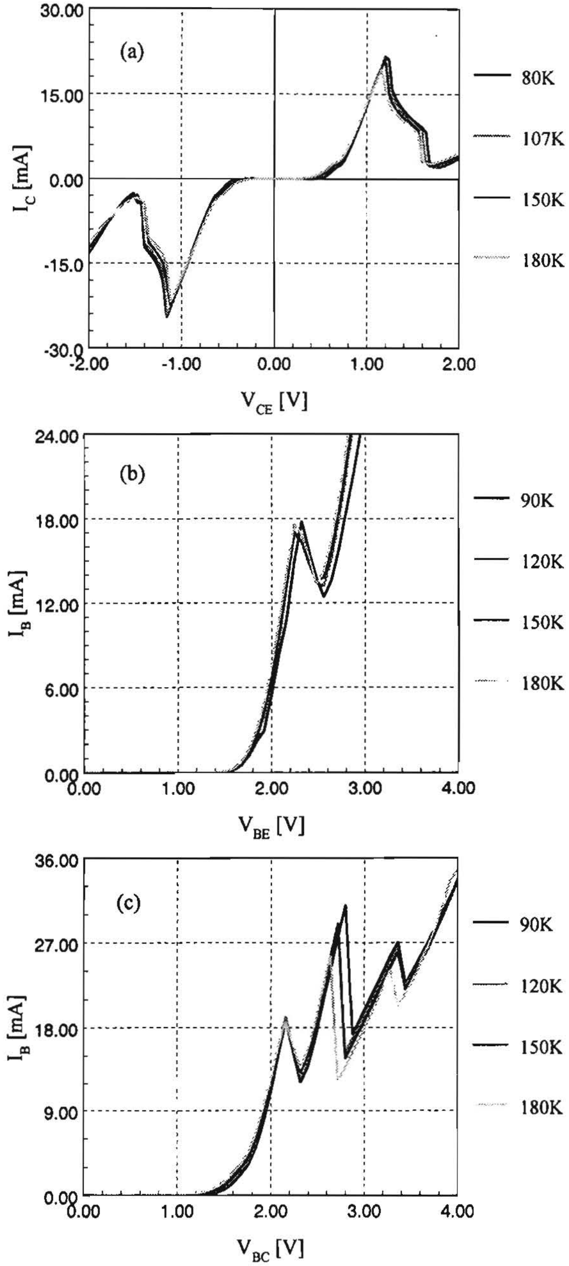


Fig. 5.8 Low temperature I-V characteristics for the two terminal biasing configurations of the 3T DBRT device: C_{bias}-E_{grd} (a), B_{bias}-E_{grd} (b) and B_{bias}-C_{grd} (c).

Of course, we could attribute the new NDR region (the first one) to the possible path 3 for the current flow. Such a situation would correspond to a picture with two DBS's in series (path 4) connected in parallel with a third one (path 3). But why is this not observed at room temperature? A possible explanation is in connection with the temperature behaviour of the I-V curves associated with the two types of DBS's from our 3T device. Figs. 5.9a and 5.9c indicate that the "active" DBS I-V characteristics change visibly with the temperature variation. With increasing the temperature the peak current decreases, the valley current increases and there is significant shift of the peak voltage towards lower values. It is not the same in the case of "contact" DBS. The analysis of Fig. 5.8b shows us a relatively small change of the corresponding I-V curves. With these considerations the disappearance of the first peak at higher T can be explained by the shift of the second peak over it and in this way its masking.

Taking into account the complex nature of the current flow between B-E and especially B-C terminals, the associated I-V characteristics demonstrate without any doubt that the base contact lies on the quantum well of the DBRT structure. In fact this represents the first step in the attempt to control the potential of the quantum well and thus to influence the current flow between emitter and collector. In the following, analysing the current characteristics of the 3T DBRT device, we shall demonstrate that, even with the impediment of a large depletion region between the base contact and the active DBS, this is possible.

Before explaining the common emitter and collector I-V characteristics we demonstrate, in principle, how a 3T DBRT device works, when there is no separation between the controlling and current-carrying carriers. We consider the case in which the base contact is ohmic and lies in the middle of the quantum well width. To illustrate our description we use the band diagram of the E-C path shown in Fig. 5.9. The device operates in the common emitter configuration which means that at the beginning when $V_{CE} = 0$ both E and C are grounded, Fig. 5.9A. A positive bias on the base changes the band diagram as in Fig. 5.9B and determines a current flow between both B-E and B-C terminals. The position of the quasi states in the quantum well lowers versus the emitter and collector quasi Fermi levels. Now, if we apply a positive increasing bias on the collector relative to the emitter the B-C voltage decreases. The base potential remains approximately unchanged until it becomes equal with the collector potential (the corresponding quasi Fermi levels are equal). In this conditions the collector current first decreases to zero and then increases in the opposite sense, and the emitter current increases continuously. When the collector voltage is increased further the corresponding current has the same behaviour as in the case without bias on the base. Finally we obtain an I-V characteristic almost identical with that of the 2T DBRT diode but shifted to the right on the positive voltage axis. The superposition of the two I-V curves, with and without excitation on the base, indicates, for the first one, a small shift to the left on the voltage axis, but because the shift of this to the right is much bigger the resultant effect will be, in fact, the displacement to higher voltages. In terms of the band diagram description this means that the initial change of the position of the quantum well quasi states relative to the emitter quasi Fermi level will be maintained, even though during the increasing of the collector bias this can be diminished. The general behaviour of the I-V characteristic will not depend on interchanging the applied voltage on the emitter and collector. Also the constant current base excitation gives the same result.

For our specific situation the I-V characteristics, in the common emitter and collector configurations are shown in Figs. 5.10 and 5.11, respectively. In both cases the base potential is

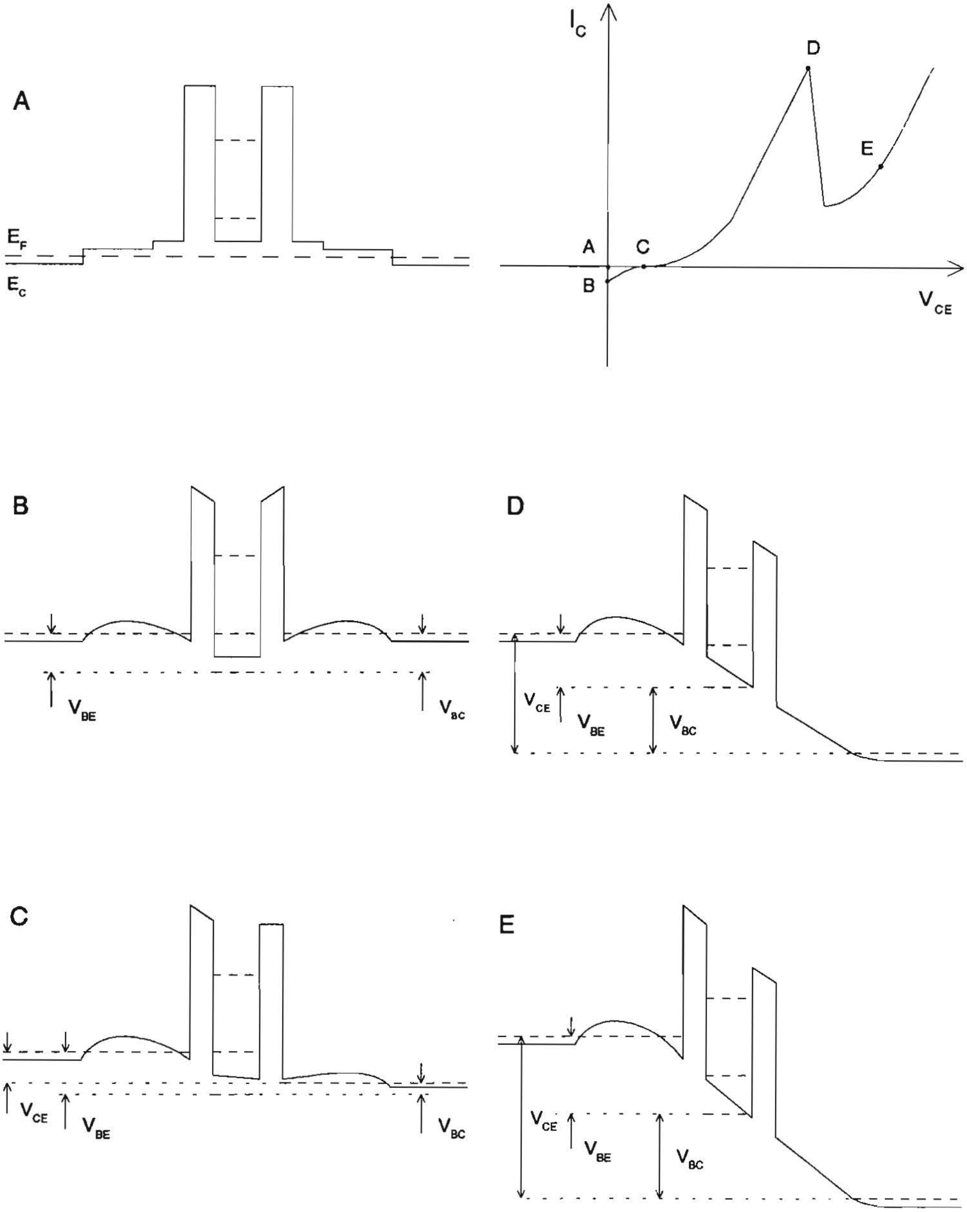


Fig. 5.9 Band diagram illustration of the principle of a 3T DBRT device with no separation between the controlling and current carrying carriers.

altered using a constant current source. From the beginning it comes out that the shift sense of the curves with the base current depends of the configuration and in both situations most of the base current is found again in the emitter current. These aspects can be explained if we take into account the new features introduced in our structure during the processing, which make it different from the ideal case presented above. As we pointed out previously, when we discussed the two terminal I-V characteristics, between the base and the quantum well of the "active" DBS there is a large depletion region. We have seen that in this situation the B-C current flow is almost entirely along the path 4 indicated in Fig. 5.5. The effect will be that in both configurations the base current changes the potential along the whole "active" DBS, the highest potential value being in the emitter side (with our notation convention from Fig. 5.5). This is equivalent with an ideal situation in which the base contact is positioned underneath the DBS (at the emitter side) and the base current alters directly the emitter triangular quantum well potential. Thus, the shift of the I-V curves with the base current increasing is to the right for the common emitter configuration and to the left for the common collector one. It is not yet clear why the base current flows almost entirely through the emitter terminal because the turn on voltages are almost the same for both the B-E and B-C I-V characteristics (see Fig. 5.6). In connection with this the B-C current is small and the base influence on the C-E I-V curve is not so high. Even in these conditions the transfer characteristic, for instance $I_C = f(I_B)$ in the common emitter configuration, demonstrates an important switching property of our 3T DBRT device (see Fig. 5.12). As one observes, applying a fixed collector voltage near the valley (peak) point of the forward (reverse) principal I-V characteristics, and varying the base current, an abrupt variation of the collector current is registered. In the cases shown in Fig. 5.12 the effect is disturbed by the instability of the device in the NDR region. This effect is also analysed in the next paragraph where a switching time is estimated.

The behaviour of the 3T device I_C-V_{CE} characteristic does not change with temperature (see Fig. 5.13a). Comparing Figs. 5.13b and 4a it is interesting to point out that the influence of the base excitation on the principal device current is practically independent of temperature.

The static characteristics of the 3T DBRT device presented in this section demonstrate that the idea of the DBS controlled by the direct action on the quantum well potential is feasible. Of course, with our device the optimum is not yet achieved. The main problem remains the reduction (elimination) of the depletion region under the open base surface and the base contact.We believe that, in comparison with the few trials concerning the same basic idea from literature [2, 3, 4, 5], our results show more clearly the technological way to follow in this direction, even if important difficulties still have to be overcome.

5.4.2 AC Measurements

Microwave measurements have been carried out on the 3T DBRT devices in a two port configuration (see Fig. 4.10), using the same HP 8510B network analyzer as for the 2T devices. Again the frequency range has been between 50MHz and 40 GHz. Maintaining the same convention concerning the device terminals as in the case of the DC measurements (see Fig. 5.5), during the microwave testing the device emitter has been grounded. The measurements have been done on devices with 20x20 μm^2 active area.

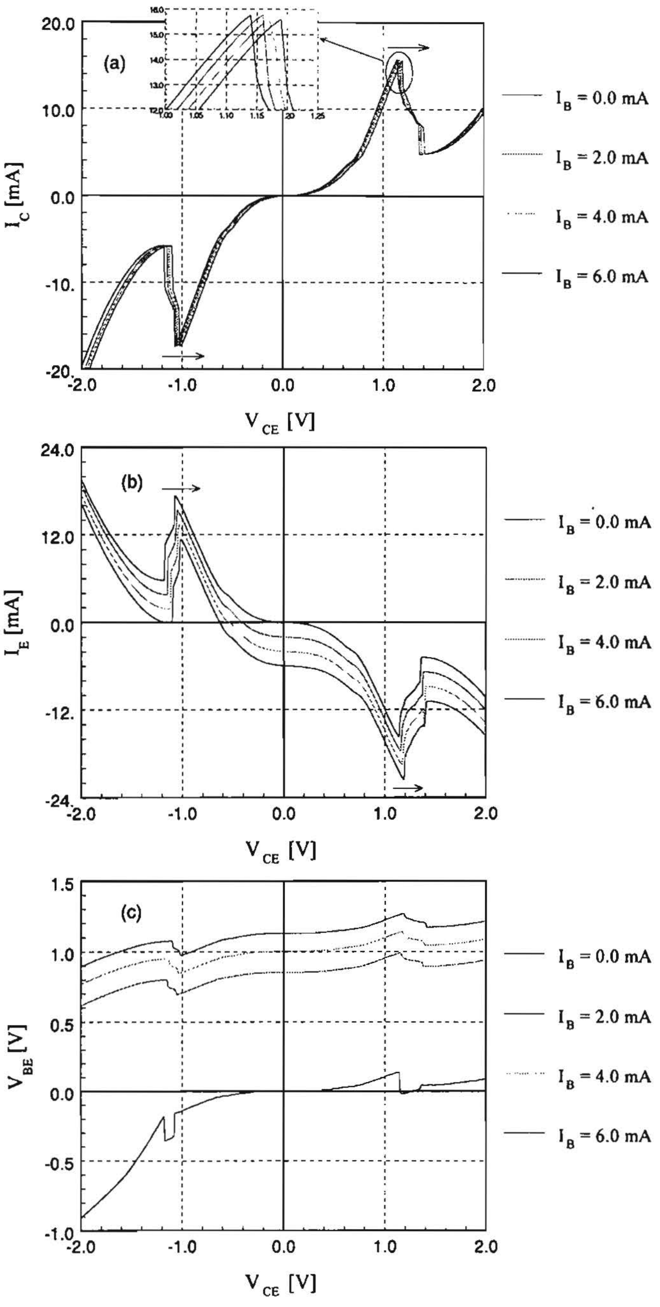


Fig. 5.10 Constant base current I-V characteristics in the common emitter configuration: I_C -V_{CE} (a), I_E -V_{CE} (b) and V_{BE} -V_{CE} (c).

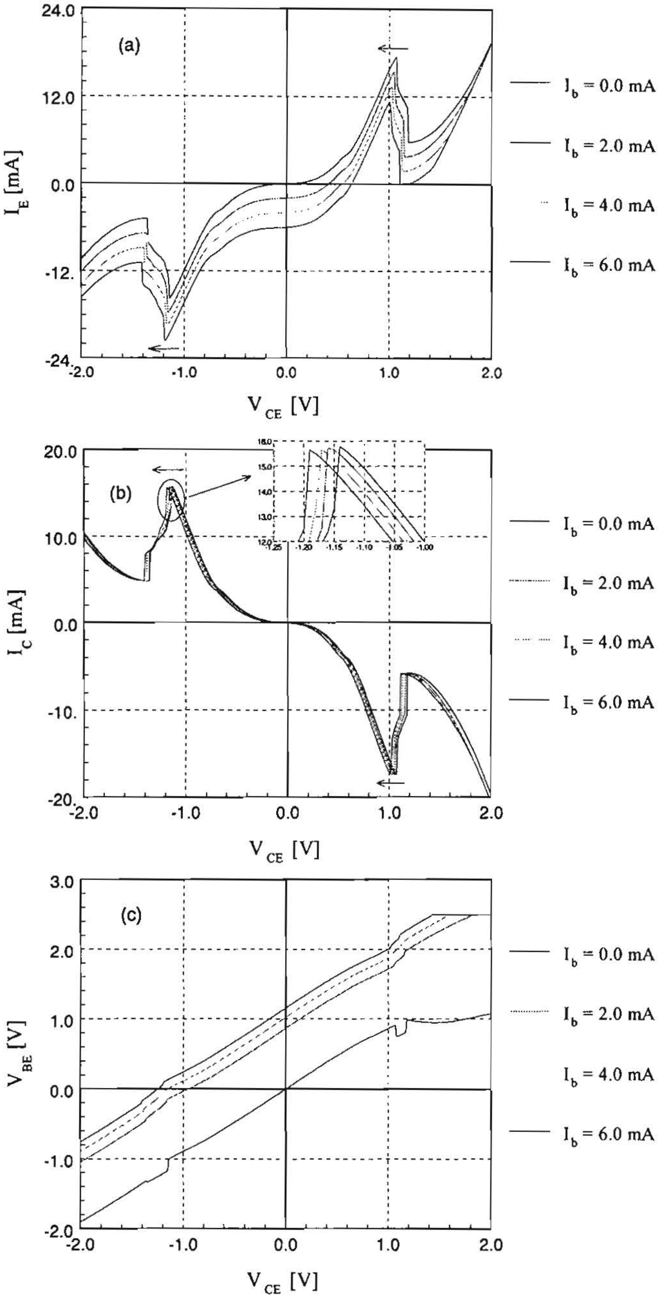
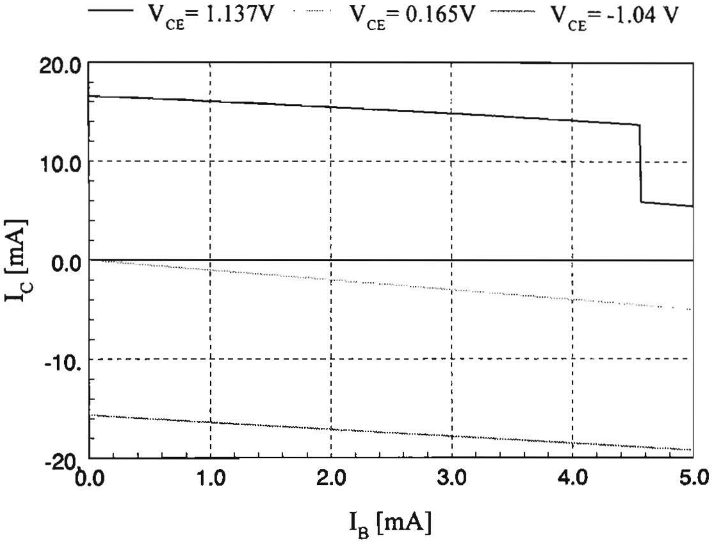


Fig. 5.11 Constant base current I-V characteristics in the common collector configuration: I_E -V_{EC} (a), I_C -V_{EC} (b) and V_{BC} -V_{EC} (c).



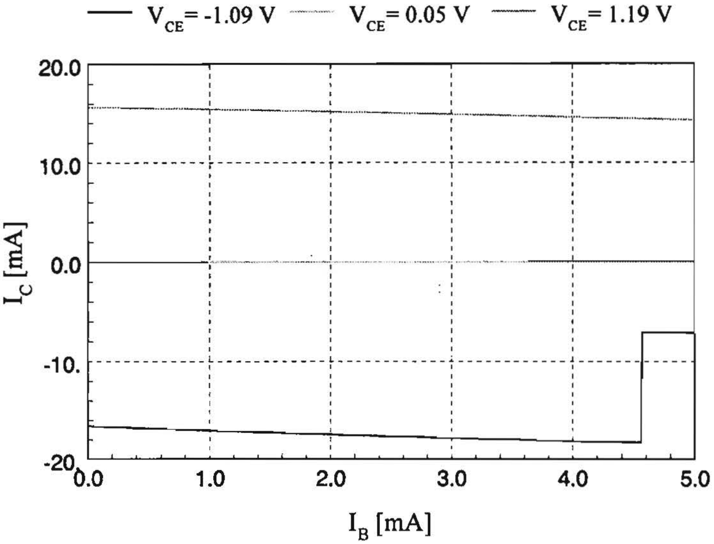


Fig. 5.12 Transfer characteristics for different collector voltages. For V_{CE} of 1.37 V and -1.09 V the collector current has abrupt jumps.

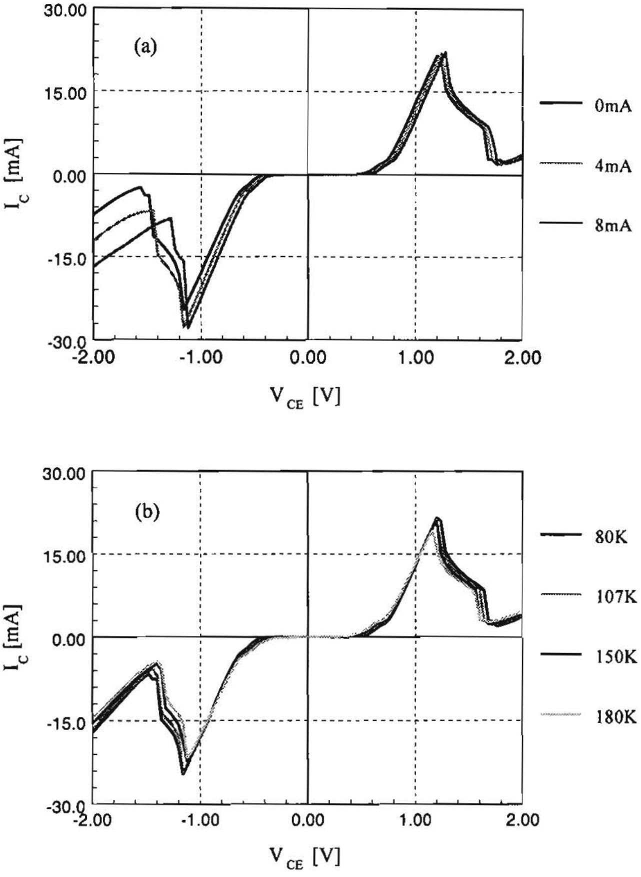
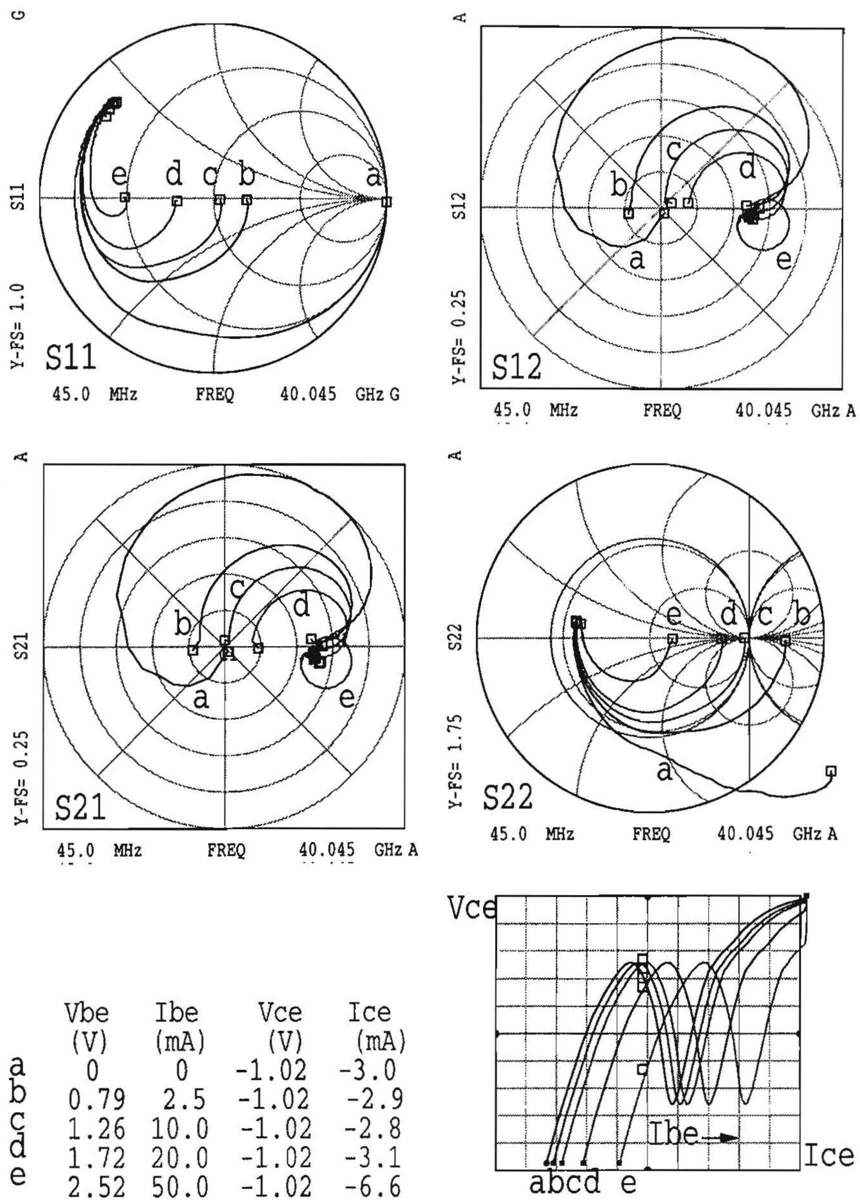


Fig. 5.13 Temperature dependence of the 3T DBRT device I-V characteristics : I_C-V_{CE} at 80 K (a) and I_C-V_{CE} at different temperatures for $I_b = 4$ mA (b).

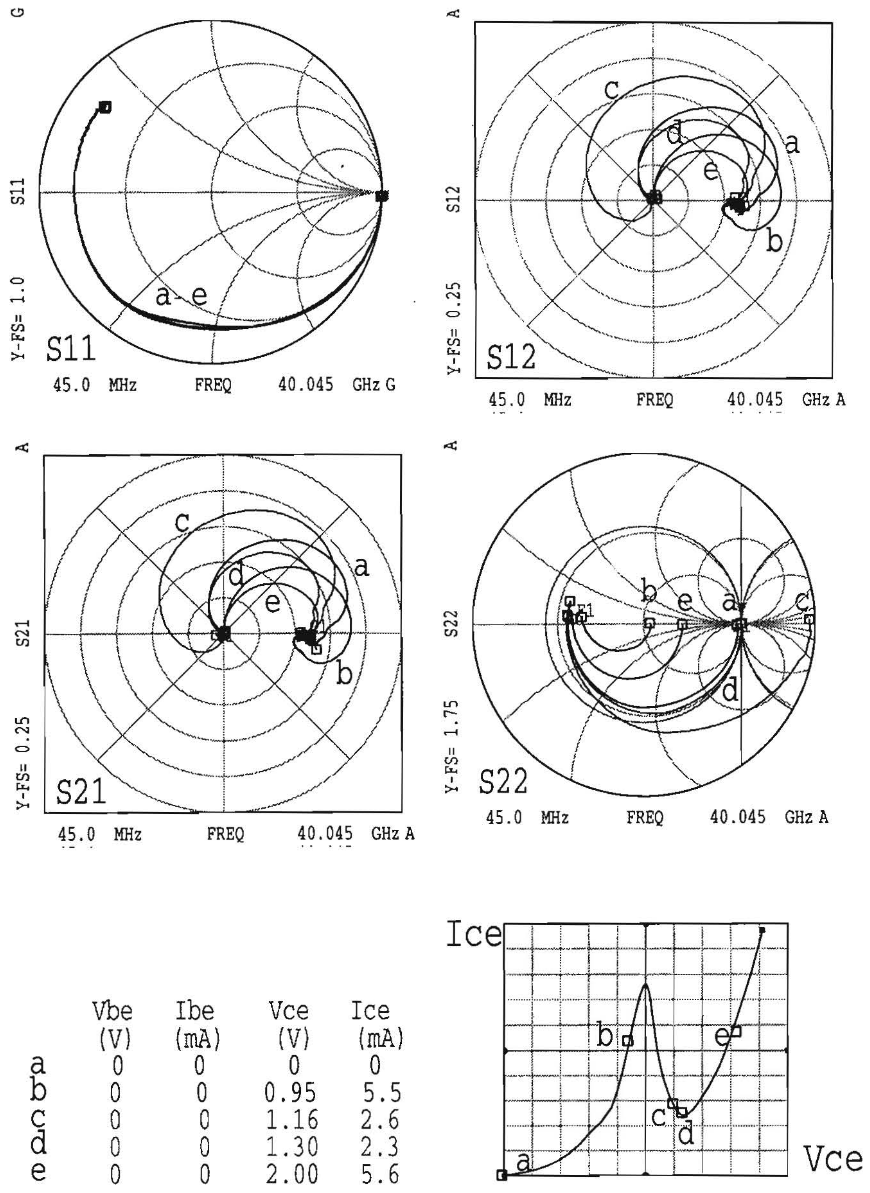
Two sets of measurements have been considered. In the first set the base-emitter DC bias has been kept constant, at zero value, and the S-parameters for different collector-emitter DC biases have been determined. Fig. 5.14 shows the results in a self explanatory way. As expected, S_{11} reflection coefficient corresponding to the B-E port, does not change with C-E DC bias. Also, S_{22} at the C-E port shows the usual behaviour of a simple 2T DBRT device when the DC bias is changed. Both the transmission coefficients, S_{21} and S_{12}, are identical for a certain C-E DC bias.

In the second measurement set the S-parameters are measured for a fixed C-E DC bias when the B-E DC bias is varied. The results are presented in Fig. 5.15. S_{11} shows the known diode behaviour when the DC bias (current) is increased. Instead, S_{22} behaviour demonstrates the DC influence of the base excitation on the C-E DC I-V characteristic, as we have observed during the measurements. Again the transmission coefficients have the same behaviour.



	Vbe (V)	Ibe (mA)	Vce (V)	Ice (mA)
a	0	0	-1.02	-3.0
b	0.79	2.5	-1.02	-2.9
c	1.26	10.0	-1.02	-2.8
d	1.72	20.0	-1.02	-3.1
e	2.52	50.0	-1.02	-6.6

Fig. 5.15 S-parameter measurements on a 3T DBRT device, for five B-E bias points and constant C-E bias (-1.02 V).



	Vbe (V)	Ibe (mA)	Vce (V)	Ice (mA)
a	0	0	0	0
b	0	0	0.95	5.5
c	0	0	1.16	2.6
d	0	0	1.30	2.3
e	0	0	2.00	5.6

Fig. 5.14 S-parameter measurements on a 3T DBRT device, for constant B-E bias (zero) and five C-E bias points.

Besides the possibility to produce high frequency oscillations (see section 4.3.2), the NDR region of the DBRT device I-V characteristic is also the basis for switching from the peak to the valley region and vice versa. The switching process occurs by DC biasing the device near the peak point through a load resistance R_{load} that is consistent with two possible DC bias quiescent points, one above and one below the NDR region. This resistance must satisfy the condition $R_{load} > \Delta V/\Delta I$. In these conditions, a slight increase of the bias voltage will result in a switch to the valley point or beyond. The switching time of the resonant tunneling devices represents an important characteristic parameter demonstrating once more that these are very fast [6, 7].

For the standard DBRT diode the switching time has been measured in the past using the electro-optic sampling technique [6, 7]. With 3T DBRT devices as we fabricated the measurement procedure of the switching time becomes much simpler. After the DC biasing of the device between C-E near the peak point, the small additional signal (bias) can be directely applied on the quantum well base, which, as we have seen before, produces the shift of the main I-V characteristics. The measurements have been carried out with the setup presented in Fig. 5.16. As in the case of the microwave mesurements we made use of Cascade probes to access the device terminals, with the emitter grounded. Also, the tested device had a 20x20 μm^2 active area. A negative bias near the peak V_{CE} = -1.097 V (I_{CE} = -16.1 mA) has been applied between C-E and the base has been positively biased with V_{BE} = 1.30 V (I_{BE} = 0.51 mA). The results are shown in Fig. 5.17, indicating a switching time t_{switch} = 55.4 ps. Taking into account the corrections needed because of the measurement procedure [8] we finally get the value t_{switch} = 42.1 ps. We have to

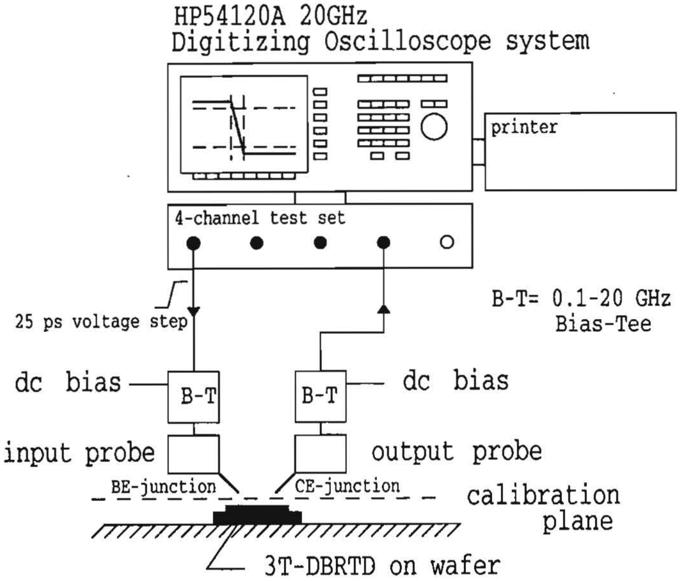
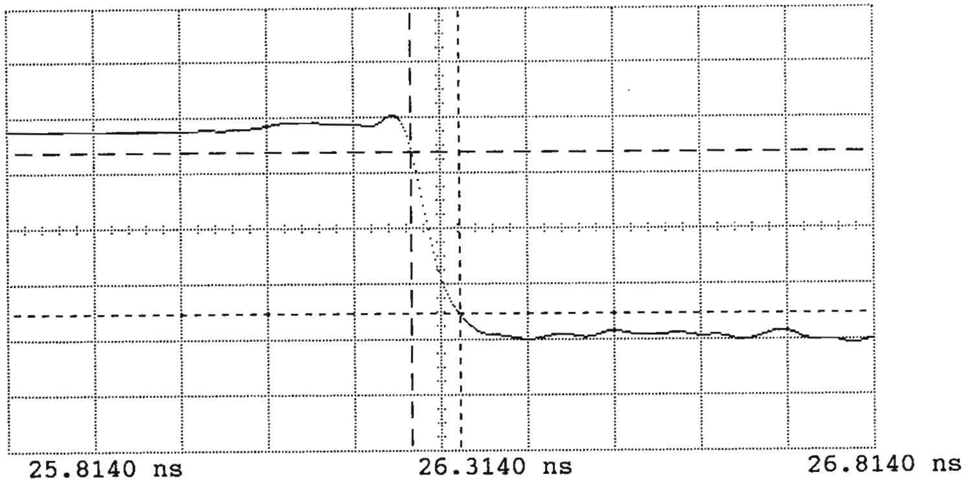


Fig. 5.16 Measurement setup for the switching time of the DBRT structure using a 3T device.



```
Timebase   =  100   ps/div                          Delay      =   25.8140 ns
Memory 2   =  80.00 mVolts/div                      Offset     =  -150.0 mVolts
Timebase   =  100   ps/div                          Delay      =   25.8140 ns
Mem 2 Parameters                                    Fall Time  =   55.4    ps
```

Fig. 5.17 Switching time measuring with the digitizing oscilloscope.

mention that in our experiment the device switched not between the peak and valley points but between the peak and an intermediate point in the NDR region, because the device was not totally stable in the measuring circuit (see also the DC results describing the same problem). A theoretical evaluation of t_{switch} can be obtained using the approximate formula [9] $t_{switch} \approx 4.4 (\Delta V/\Delta I)C$, where $\Delta V = V_p - V_v$, $\Delta I = I_p - I_v$, and C the static capacitance of the DBRT structure. In our particular conditions this expression gives $t_{switch} \approx 42.5$ ps which is quite close to the measured result.

5.5 Conclusions

Using high selective etching processes we have fabricated for the first time a 3T DBRT device based on the direct contacting of the quantum well. This is demonstrated without doubts by the I-V characteristics between the B-E and B-C terminals. The base excitation determines a shift of the collector-emitter I-V characteristics on the voltage axis. Consequently the transfer characteristic, collector current as a function of the base current or voltage, shows important switching properties. The switching time of the DBRT structure used in the 3T DBRT devices has been measured, indicating a value of 42.1 ps that is quite close to the estimation obtained from an approximate theoretical formula used in the literature.

The electrical properties of our 3T DBRT devices can be used in switching applications and for tunned oscillators.

References

1. H. K. Henisch, "*Semiconductor Contacts - An approach to ideas and models*" Clarendon Press, Oxford, 1984.
2. G. I. Haddad, R. K. Mains, U. K. Reddy, and J. R. East, *Superlatt. Microstruct.* **5**, 437 (1989).
3. U. K. Reddy, I. Mehdi, R. K. Mains and G.I. Haddad, *Solid-State Electronics* **32,** 1377 (1989).
4. J. Genoe, *Ph. D. Thesis*, Leuven, 1994
5. W. C. B. Peatman, E. R. Brown, M. J. Rooks, P. Maki, W. J.Grimm and M. Shur, IEEE *Electron Device Lett.* **15**, 236 (1994).
6. J. F. Whitaker, G. A. Mourou, T. C. L. G. Sollner, and W. D. Goodhue, *Appl. Phys. Lett.* **53**, 385 (1988).
7. S. K. Diamond, E. Özbay, M. J. W. Rodwell, D. M. Bloom, Y.C. Pao, E. Wolak, and J. Harris, Jr., in "*OSA Proceedings on Picosecond Electronics and Optoelectronics*", Vol. **4**, p. 101, Optical Society of America, Washington, D.C., 1989.
8. J. J. M. Kwaspen, private communication.
9. E. R. Brown, "Resonant-Tunneling Transistors", in *Heterostructures and Quantum Devices*, eds. N. G. Einspruch and W. R. Fresnley, Academic Press, San Diego, 1994.

Conclusions

In this work 2T DBRT devices have been studied and a new 3T DBRT device has been fabricated and characterized. All the devices were processed from MBE grown GaAs/AlGaAs based structures.

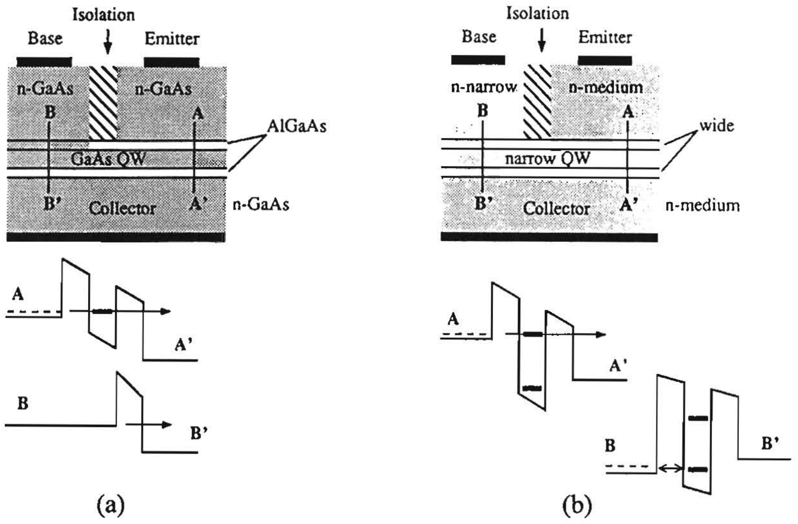
The technological processes used for the fabrication of the devices have been investigated minutely. A detailed description of these is presented (chapter 3 and Appendix B), the basic idea of the presentation being the understanding of the process mechanism and their direction of evolution. Important results have been obtained in the development and optimization of the dual resist photolithography, wet chemical etching, selective etching of GaAs and AlGaAs, and gold plating for airbridge interconnections. In the same field of the technology, a mask set for the 3T DBRT device processing has been designed and fabricated.

In a separate section (chapter 2) the physical basics of the resonant tunneling phenomenon are presented in the Breit-Wigner formulation of the process, an approach used in quantum mechanics to describe the resonant scattering phenomena in general. Also, main elements necessary for the measurement interpretation and modeling are introduced.

GaAs/AlAs and $Al_{0.4}Ga_{0.6}$ As based 2T DBRT devices, processed in planar configuration, have been DC and AC characterized (chapter 4). DC results, at room temperature and below, are dicussed and explained from the point of view of device physics in a qualitative manner and using a self consistent physical model. These descriptions are based on the theoretical considerations developed in chapter 2. In both approaches one emphasizes the role played by the accumulation region (where the electrons are distributed both in 3D and 2D states)in the DC behaviour. An accurate AC analysis reveals that the most proper equivalent circuit representation of our DBRT structures has to include besides the Esaki model elements also a quantum inductance. For the used active region configuration, the DC and AC performances of our 2T DBRT devices are situated at the superior limit, indicating a high quality fabrication and accurate characterization.

GaAs/AlAs based 3T DBRT devices with direct contacting of the quantum well have been fabricated for the first time, using highly selective etching processes and airbridge interconnections (chapter 3). DC measurements indicate clearly the positioning of the base contact to the quantum well and the shift of the collector-emitter I-V characteristics on the voltage axis when the base potential is altered. The devices show an important switching property which characteristic, the switching time, has been determined in AC conditions.

Further developments concerning 2T DBRT devices can be done in two directions. One of them refers to the fabrication of devices with high electrical performance necessary in high speed electronics, an aspect in which the theoretical limits have not been attained. In these sense the problem must be approached by using other material systems for the device structure, some of them and their advantages being mentioned in chapter 1, e. g. InGaAs/InAlAs, InAs/AlSb or



(a) (b)

Fig. C.1 3T DBRT devices with the base contact on the quantum well: (a) unipolar device using the "classical" GaAs/Al$_x$Ga$_{1-x}$As DBRT structure; (b) bound-state resonant tunneling transistor (BSRTT) [1].

InAs/GaSb/AlSb.

The other direction of investigation regards the improvement of the physical model developed within the EUT Electronic Devices Group, that because of its accessibility can become an efficient tool for the design of different DBRT structures.

The investigation of the 3T DBRT devices can be continued as well. Here we also think of two directions. The first one refers to the improvement of the approach developed in this dissertation. On the one hand this can be done by reduction of the open base surface using the actual mask set in a different device processing. On the other hand, in a much more elaborated fabrication procedure, one can try the realization of a 3T DBRT device with hanging base using a two side sample processing.

The second direction has in view the fabrication of a 3T DBRT transistor. To illustrate this we refer to Fig. C.1. For the fabrication of our 3T DBRT devices we made use of a simple DBRT structure. In such a configuration the equilibrium electrons in the quantum well base occupy the lowest quasi-states. On the other hand, as known, in the DBRT structure the lowest quasi-states are typically also the electron resonant tunneling states (path AA' in Fig. C.1a). Therefore, there is no separation between the base controlling and collector-emitter current-carrying electrons, as it is required for transistor action. Moreover, the base-collector tunnel barrier (path BB') is also not sufficient to confine the electrons to the base, as it must be thin enough to permit resonant tunneling (this could be overcome by the "hanging base" solution suggested above). Consequently, in this kind of device transistor action is not obtained. A configuration which overcomes the above drawbacks, separating the controlling and current-carrying electrons is shown in Fig. C.1b. In this bound-state resonant tunneling transistor (BSRTT) a narrow-bandgap quantum well is used to hide the lowest-lying electron state from the tunneling electrons. The base electrons are supplied to the ground state by direct electrical connection to the well (base contact) and confined by the heterojunction collector (see section BB'). Resonant tunneling base electrons are supplied to the ground state by direct electrical connection to the well (base contact) and confined by the heterojunction collector (see section BB'). Resonant tunneling occurs when the quantum well excited states align with the emitter electron distribution, an alignment which is controlled by the base-emitter bias. In the practical realization of this solution the quantum well was not directly contacted but using other technological modalities [2, 3]. Besides the unipolar devices described above bipolar DBRT transistors can be also conceived. This solution has been materialized in several variants [4-8], contacting of the well not being realized in a direct way. So, combining our realization with the solutions which ensure transistor action an important progress in the field of the 3T DBRT devices could be registered.

Apart from this developments aiming for fabrication of devices destined to high speed applications one can not ignore the possibility of fundamental research of some new experimental situations encountered in the fabrication of our devices such as metal-2D semiconductor layer junction.

References

1. A. C. Seabaugh and M. A. Reed, "Resonant-Tunneling Transistors", in *Heterostructures and Quantum Devices*, eds. N. G. Einspruch and W. R. Fresnley, Academic Press, San Diego, 1994.
2. J. N. Schulman and M. Walder, *J. Appl. Phys.* **63**, 2859 (1988).
3. G. I. Haddad, R. K. Mains, U. K. Reddy, and J. R. East, *Superlatt. Microstruct.* **5**, 437 (1989).
4. M. A. Reed, W. R. Frensley, R. J. Matyi, J. N. Randall, and A. C. Seabaugh, *Appl. Phys. Lett.* **54**, 1034 (1989).
5. A. C. Seabaugh, W. R. Frensley, J. N. Randall, M. A. Reed, D. L. Farrington, and R. J. Matyi, *IEEE Trans. Electron Devices* **36**, 2328 (1989).
6. T. Waho, K. Maezawa, and T. Mizutami, *Jpn. J. Appl. Phys.* **30**, L2018 (1991).
7. T. Waho, K. Maezawa, and T. Mizutami, *IEEE Electron Device Lett.* **14**, 202 (1993).
8. A. C. Seabaugh, Y.-C. Kao, W. R. Frensley, J. N. Randall, and M. A. Reed, *Appl. Phys. Lett.* **59**, 3413 (1989).

Appendix A

A.1 Transmission Probability in a General DBRT Structure

For the structure shown in Fig. A1 the transmission probability is found , first for the coherent tunneling and afterwards considering scattering by introducing an "incoherent parameter". We use for this task the well known transmission matrix method [1, 2].

In the effective-mass approximation the motion of an electron along the structure is described by the following time-independent Schrödinger equation in one dimension

$$\left[-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m^*(z)}\frac{\partial}{\partial z}+V(z)\right]\psi(z)=E\psi(z) \tag{A.1}$$

where $\psi(z)$ is the envelope function, $m^*(z)$ is the effective mass of the electron and $V(z)$ is the potential energy, i.e. the energy in the minimum of the conduction band.

The probability to find the electron in a certain place z is proportional to $|\psi(z)|^2$ and the current density per electron is

$$J_z=\frac{q\hbar}{m^*}\operatorname{Im}\left(\psi^+\frac{\partial\psi}{\partial z}\right) \tag{A.2}$$

where q is the electron charge and we adopted the '+' notation for the complex conjugate.

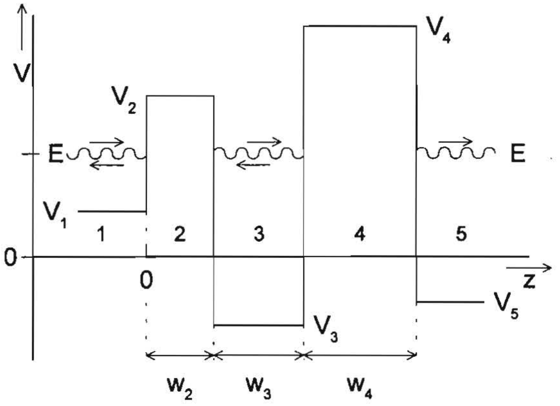


Fig. A.1 Schematic band diagram for a general DBRT structure. The electron wave function corresponding to the energy $E > V_1$ is incident on the double barrier from the left. The wave reflects at the barriers and transmits through the structure to the right.

For each region ' j ', where $V(z)$ is constant, the general solution of Eq. A.1 may be written as a combination of plane waves, right and left travelling, or growing and decaying exponentials, depending of the relative magnitude of E versus $V(z)$,

$$\psi = a_j e^{ik_j z} + b_j e^{-ik_j z} \tag{A.3}$$

where

$$k_j = \sqrt{\frac{2m_j^*}{\hbar^2}(E - V_j)}, \qquad E \geq V_j \tag{A.4a}$$

$$= \pm i\sqrt{\frac{2m_j^*}{\hbar^2}(V_j - E)} = \pm i\alpha_j, \quad E < V_j \tag{A.4b}$$

Here, k_j and α_j denote the electron wave vector and attenuation coefficient, respectively.

The different solutions on both sides of an interface are matched each other with the boundary conditions:

$$\psi\ (z) \text{ continuous and } \frac{1}{m^*}\frac{d}{dz}\psi(z) \text{ continuous [3].}$$

Referring to the junction from Fig. A1.2 the solution to the Schrödinger equation in immediate vicinity of $z = z_i$ is:

$$\psi = \begin{cases} a_l e^{ik_l z} + b_l e^{-ik_l z}, & z \leq z_i \\ a_r e^{ik_r z} + b_r e^{-ik_r z}, & z \geq z_i. \end{cases} \tag{A.5}$$

The application of the continuity conditions yields a_l, b_l in terms of a_r, b_r and vice versa. Because the matching conditions are linear, a_l, b_l must be linearly related to a_r, b_r, i.e. there exists a 2x2 matrix M^{z_i} satisfying,

$$\begin{aligned} \begin{pmatrix} a_l \\ b_l \end{pmatrix} &\equiv \begin{pmatrix} \text{incoming amplitude on left} \\ \text{outgoing amplitude on left} \end{pmatrix} \\ &= M^{z_i}\begin{pmatrix} \text{outgoing amplitude on right} \\ \text{incoming amplitude on right} \end{pmatrix} \\ &\equiv M^{z_i}\begin{pmatrix} a_r \\ b_r \end{pmatrix}. \end{aligned} \tag{A.6}$$

The matrix M^{z_i} is called the transfer or transmission (T) matrix and can easily be put in the following useful form

$$M^{z_i} = \frac{1}{2}\begin{pmatrix} (1+\gamma)e^{i(k_r - k_l)z_i} & (1-\gamma)e^{-i(k_r + k_l)z_i} \\ (1-\gamma)e^{i(k_r + k_l)z_i} & (1+\gamma)e^{-i(k_r - k_l)z_i} \end{pmatrix}, \tag{A.7}$$

with $\gamma = \dfrac{m_l^* k_r}{m_r^* k_l}$. If k_l is imaginary the expression of M^{z_i} continues to be valid provided that this is replaced by $i\alpha_l$.

The T-matrix of a chain of heterojunctions is equal to the product of those of the components. So, in this T-matrix approach a potential barrier can be represented by the matrix product of two

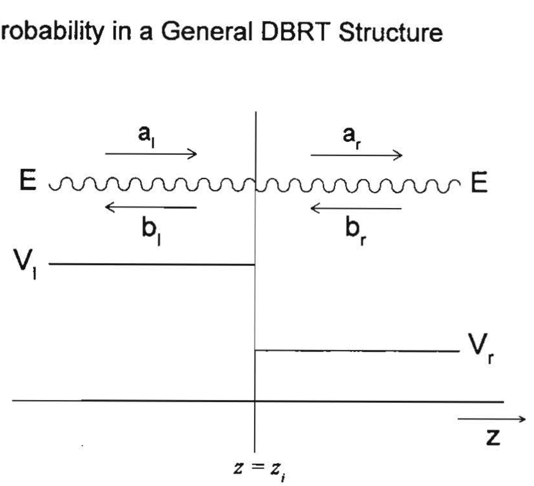


Fig. A.2 Energy band diagram of a heterointerface. It is assumed that $E > V_l > V_r$, so that there is physical electron wave propagation on both immediate sides of $z = z_i$. a_l, b_l and a_r, b_r are the incoming and outgoing wave amplitudes on the left and right side, respectively.

interface matrices, and a double barrier structure by the product of two single barrier matrices. Consequently, for the structure represented in Fig. A.1 we can write

$$M^{tot} = M^{B_1} M^{B_2},$$

with

$$M^{B_1} = M^0 M^{w_2} \text{ and } M^{B_2} = M^{w_2+w_1} M^{w_2+w_3+w_4},$$

in self-explanatory notation. M^{B_1} and M^{B_2} can be expressed in a condensed form as

$$M^{B_1} = \frac{1}{4}\begin{pmatrix} z_{12}z_{23}\chi_2\xi_2 + z_{12}^{+}z_{23}^{+}\chi_2\xi_2^{-1} & z_{12}z_{23}^{+}\chi_2^{+}\xi_2 + z_{12}^{+}z_{23}\chi_2^{+}\xi_2 \\ z_{12}^{+}z_{23}\chi_2\xi_2 + z_{12}z_{23}^{+}\chi_2\xi_2^{-1} & z_{12}^{+}z_{23}^{+}\chi_2^{+}\xi_2 + z_{12}z_{23}\chi_2^{+}\xi_2^{-1} \end{pmatrix}, \tag{A.8a}$$

and

$$M^{B_2} = \frac{1}{4}\begin{pmatrix} z_{34}z_{45}\chi_3^{+}\chi_4\xi_4 + z_{34}^{+}z_{45}^{+}\chi_3^{+}\chi_4\xi_4^{-1} & z_{34}z_{45}^{+}\chi_3^{+}\chi_4^{+}\xi_4 + z_{34}^{+}z_{45}\chi_3^{+}\chi_4^{+}\xi_4 \\ z_{34}^{+}z_{45}\chi_3\chi_4\xi_4 + z_{34}z_{45}^{+}\chi_3\chi_4\xi_4^{-1} & z_{34}^{+}z_{45}^{+}\chi_3\chi_4^{+}\xi_4 + z_{34}z_{45}\chi_3\chi_4^{+}\xi_4^{-1} \end{pmatrix}, \tag{A.8b}$$

where $z_{12} = 1 + i\dfrac{m_1^*\alpha_2}{m_2^*k_1} = 1 + i\gamma_1$, $z_{23} = 1 - i\dfrac{m_2^*k_3}{m_3^*\alpha_2} = 1 - i\gamma_2$, $z_{34} = 1 + i\dfrac{m_3^*\alpha_4}{m_4^*k_3} = 1 + i\gamma_3$ and $z_{45} = 1 - i\dfrac{m_4^*k_5}{m_5^*\alpha_4} = 1 - i\gamma_4$; $\chi_2 = e^{ik_3w_2}$, $\chi_2 = e^{ik_3(w_2+w_3)}$, and $\chi_3 = e^{ik_5(w_2+w_3+w_4)}$; $\xi_2 = e^{\alpha_2 w_2}$;and $\xi_4 = e^{\alpha_4 w_4}$.

It turns out that $M_{22}^{B_i} = (M_{11}^{B_i})^+$ and $M_{21}^{B_i} = (M_{12}^{B_i})^+$, $i = 1, 2$, which is generally valid for any barrier

[4]. The total T-matrix of the double barrier is then calculated using the single barrier matrix elements:

$$M^{tot} = \frac{1}{16}\begin{pmatrix} M_{11}^{B_1} M_{11}^{B_2} + M_{12}^{B_1} M_{21}^{B_2} & M_{11}^{B_1} M_{12}^{B_2} + M_{12}^{B_1} M_{22}^{B_2} \\ M_{21}^{B_1} M_{11}^{B_2} + M_{22}^{B_1} M_{21}^{B_2} & M_{21}^{B_1} M_{12}^{B_2} + M_{22}^{B_1} M_{22}^{B_2} \end{pmatrix}. \tag{A.9}$$

Any barrier or chain of barriers are characterized by the wave transmission and reflection coefficients. We define these coefficients first for a single barrier and then extend the definition for the whole structure. Considering, for example, the first barrier of our structure, for a wave incident from the left the corresponding coefficients are:

- the wave transmission coefficient $t_{B_1} \equiv \left.\frac{a_3}{a_1}\right|_{b_3=0} = \frac{1}{M_{11}^{B_1}},$ (A.10a)

- the wave reflection coefficient $r_{B_1} \equiv \left.\frac{b_1}{a_1}\right|_{b_3=0} = \frac{M_{21}^{B_1}}{M_{11}^{B_1}}.$ (A.10b)

and for an incident wave from the right they are:

- the wave transmission coefficient $t'_{B_1} \equiv \left.\frac{b_1}{b_3}\right|_{a_1=0} = M_{22}^{B_1} - \frac{M_{12}^{B_1} M_{21}^{B_1}}{M_{11}^{B_1}},$ (A.11a)

- the wave reflection coefficient $r'_{B_1} \equiv \left.\frac{b_1}{a_1}\right|_{b_3=0} = -\frac{M_{12}^{B_1}}{M_{11}^{B_1}}.$ (A.11b)

Concerning the whole structure we are interested to express these coefficients only for the situation described in the Fig. A.1, where one considers the incident wave travelling from the left. In this case the wave transmission coefficient is

$$t \equiv \left.\frac{a_5}{a_1}\right|_{b_5=0} = \frac{1}{M_{11}^{tot}} = \left[M_{11}^{B_1} M_{11}^{B_2} + M_{12}^{B_1} M_{21}^{B_2}\right]^{-1} = \frac{t_{B_1} t_{B_2}}{1 - r'_{B_1} r_{B_2}} \equiv t_{B_1}(1 - r_{B_2} r'_{B_1})^{-1} t_{B_2}, \tag{A.12a}$$

and the wave reflection coefficient has the form

$$r \equiv \left.\frac{b_1}{a_1}\right|_{b_5=0} = \frac{M_{21}^{tot}}{M_{11}^{tot}} = \frac{r_{B_1} + r_{B_2}(t'_{B_1} t_{B_1} - r'_{B_1} r_{B_1})}{1 - r'_{B_1} r_{B_2}} \equiv r_{B_1} + t_{B_1} r_{B_2}(1 - r'_{B_1} r_{B_2})^{-1} t'_{B_1}. \tag{A.12b}$$

If in Eqs. (A.12a) and (A.12b) $(1 - r'_{B_1} r_{B_2})^{-1}$ is expanded in a power series the transmission and reflection coefficients can be written as

$$t = t_{B_1} t_{B_2} + t_{B_1}(r_{B_2} r'_{B_1}) t_{B_2} + t_{B_1}(r_{B_2} r'_{B_1})^2 t_{B_2} + \ldots + t_{B_1}(r_{B_2} r'_{B_1})^n t_{B_2} + \ldots \tag{A.13a}$$

and

$$r = r_{B_1} + t_{B_1} r_{B_2} t'_{B_1} + t_{B_1} r_{B_2}(r'_{B_1} r_{B_2}) t'_{B_1} + t_{B_1} r_{B_2}(r'_{B_1} r_{B_2})^2 t'_{B_1} + \ldots + t_{B_1} r_{B_2}(r'_{B_1} r_{B_2})^n t'_{B_1} + \ldots \tag{A.13b}$$

The above form of *t* and *r* reflects very well the physical mechanism of the resonant tunneling process through a double barrier structure. Thus, the *n*+1 term in the expression of *t* expresses the fact that first the tunneling through the first barrier takes place, then the wave is reflected *n* times, back and forth, in the well, and finally this escapes through the second barrier. Similarly, the reflection coefficient includes, besides the first barrier reflection, terms which take into account the multiple reflections between the barriers, in the quantum well. These multiple reflections encountered in the transmission and reflection of the electron wave function can interfere constructively or destructively leading to resonances in the transmission.

The wave transmission and reflection coefficients can be further used to calculate the transmission and reflection probabilities. These probabilities are defined in connection with the probability current densities, as we already pointed out in section 2.2.2, for the transmission probability. Evident they can refer to a single barrier or multiple barriers. For our necessities we first introduce the transmission and reflection probabilities of the first barrier of the double barrier structure. Thus we have on the left hand side

$$T_{B_1} \equiv \frac{J_{r\to}}{J_{l\to}}\bigg|_{B_1} = \frac{m_1^* k_3}{m_3^* k_1}\left|t_{B_1}\right|^2 = \frac{m_1^* k_3}{m_3^* k_1}\frac{1}{\left|M_{11}^{B_1}\right|^2}, \tag{A.14a}$$

$$R_{B_1} \equiv \frac{J_{l\leftarrow}}{J_{l\to}}\bigg|_{B_1} = \left|r_{B_1}\right|^2 = \frac{\left|M_{21}^{B_1}\right|^2}{\left|M_{11}^{B_1}\right|^2}, \tag{A.14b}$$

and on the right hand side

$$T'_{B_1} \equiv \frac{J_{l\leftarrow}}{J_{r\leftarrow}}\bigg|_{B_1} = \frac{m_3^* k_1}{m_1^* k_3}\left|t'_{B_1}\right|^2 = \frac{m_3^* k_1}{m_1^* k_3}\frac{\left|\det M^{B_1}\right|^2}{\left|M_{11}^{B_1}\right|^2}, \tag{A.15a}$$

$$R'_{B_1} \equiv \frac{J_{r\to}}{J_{r\leftarrow}}\bigg|_{B_1} = \left|r'_{B_1}\right|^2 = \frac{\left|M_{12}^{B_1}\right|^2}{\left|M_{11}^{B_1}\right|^2}, \tag{A.15b}$$

in self-explanatory notation. Because the diagonal elements of the barrier T-matrix are complex conjugate each other and $\det M^{B_1} = \frac{m_1^* k_3}{m_3^* k_1}$ one verifies $T_{B_1} = T'_{B_1}$, $R_{B_1} = R'_{B_1}$ and $T_{B_1} + R_{B_1} = T'_1 + R'_{B_1} = 1$. The last equalities guarantee in fact the current continuity which we could expect because the Schrödinger equation together with the continuity conditions conserve the probability current density. In a similar way, we can now define the transmission and reflection probabilities for the double barrier structure and, using Eqs. (A.12a) and (A.12b), write

$$T \equiv \frac{J_{r\to}}{J_{l\to}} = \frac{m_1^* k_5}{m_5^* k_1}\left|t\right|^2 = \frac{T_{B_1} T_{B_2}}{1 + R_{B_1} R_{B_2} - 2\sqrt{R_{B_1} R_{B_2}}\cos\phi} \tag{A.16a}$$

and

$$R \equiv \frac{J_{r\to}}{J_{l\to}} = \left|r\right|^2 = \frac{R_{B_1} + R_{B_2} - 2\sqrt{R_{B_1} R_{B_2}}\cos\phi}{1 + R_{B_1} R_{B_2} - 2\sqrt{R_{B_1} R_{B_2}}\cos\phi}, \tag{A.16b}$$

where $\phi = \arg(r'_{B_1}) + \arg(r_{B_2}) = -\pi + \arg\left(\frac{M_{12}^{B_1}}{M_{11}^{B_1}}\right) + \arg\left(\frac{M_{21}^{B_2}}{M_{11}^{B_2}}\right)$. Again $T + R = 1$ is valid , reflecting the continuity of the current density for the double barrier resonant tunneling in the coherent case.

Further on we consider that scattering processes occur during the electron tunneling. In this situation the coherence of the wave function is lost and we speak about the incoherent case. Following [5], we model this introducing an incoherence parameter ζ, without specifying the nature of the scattering process. ζ is specific for a DBRT structure and it reflects the loss of coherence during every internal reflection. For one bouncing back and forth of the wave function, with one reflection on each barrier, the expression for ζ can be taken as $\zeta = exp(-2w_{3eff}/\lambda_c)$, where $w_{3eff} = w_3 k/k_3$ represents the effective width of the quantum well, with k the total wave vector in the well, and $\lambda_c = \tau_s v_e$ is the coherence length, with v_e the electron velocity and τ_s the scattering time. Using this ζ, now the expressions (A.13) of the transmission and reflection coefficients in the coherent case are replaced by

$$t(\zeta) = t_{B_1}\zeta^{1/2}t_{B_2} + t_{B_1}\zeta^{1/2}(r_{B_2}r'_{B_1}\zeta)t_{B_2} + t_{B_1}\zeta^{1/2}(r_{B_2}r'_{B_1}\zeta)^2 t_{B_2} + \ldots + t_{B_1}\zeta^{1/2}(r_{B_2}r'_{B_1}\zeta)^n t_{B_2} + \ldots \quad \text{(A.17a)}$$

and

$$r(\zeta) = r_{B_1} + t_{B_1}r_{B_2}\zeta t'_{B_1} + t_{B_1}r_{B_2}\zeta(r'_{B_1}r_{B_2}\zeta)t'_{B_1} + t_{B_1}r_{B_2}\zeta(r'_{B_1}r_{B_2}\zeta)^2 t'_{B_1} + \ldots + t_{B_1}r_{B_2}\zeta(r'_{B_1}r_{B_2}\zeta)^n t'_{B_1} + \ldots, \quad \text{(A.17b)}$$

which written in a concentrated form become

$$t(\zeta) = \zeta^{1/2}\frac{t_{B_1}t_{B_2}}{1-\zeta r'_{B_1}r_{B_2}} \quad \text{(A.18a)}$$

and

$$r(\zeta) = r_{B_1} + \zeta\frac{t_{B_1}r_{B_2}t'_{B_1}}{1-\zeta r'_{B_1}r_{B_2}}. \quad \text{(A.18b)}$$

With $t(\zeta)$ and $r(\zeta)$ thus expressed we find for the transmission and reflection probabilities

$$T(\zeta) = \frac{\zeta T_{B_1}T_{B_2}}{1+\zeta^2 R_{B_1}R_{B_2} - 2\zeta\sqrt{R_{B_1}R_{B_2}}\cos\phi} \quad \text{(A.19a)}$$

and

$$R(\zeta) = \frac{R_{B_1} + \zeta^2 R_{B_2} - 2\zeta\sqrt{R_{B_1}R_{B_2}}\cos\phi}{1+\zeta^2 R_{B_1}R_{B_2} - 2\zeta\sqrt{R_{B_1}R_{B_2}}\cos\phi}. \quad \text{(A.19b)}$$

The sum of $T(\zeta)$ and $R(\zeta)$ no longer equals unity. But it represents only the fraction of electrons that tunnel coherently and consequently the difference $1 - T(\zeta) - R(\zeta)$ is interpreted as the fraction of electrons that are stuck inside the well and leave it in a separate process, i.e., incoherently. The incoherent electrons tunnel to the left and to the right in the same ratio as the transmission probabilities of the left and right barriers, that is the corresponding probabilities are $T_{B_1}/(T_{B_1}+T_{B_2})$ and $T_{B_2}/(T_{B_1}+T_{B_2})$ and respectively [6]. Hence the total transmission probability is obtained as

$$T_{tot} = T(\zeta) + [1 - T(\zeta) - R(\zeta)]\frac{T_{B_2}}{T_{B_1}+T_{B_2}}, \quad \text{(A.20)}$$

or, in worked out form

$$T_{tot} = \left(1 + \frac{(1-\zeta)(1+\zeta R_{B_2}}{\zeta(T_{B_1}+T_{B_2})}\right)T(\zeta). \quad \text{(A.21)}$$

The expression (A.21) of the transmission probability in the general, incoherent case can be used at the calculation of the current density and charge in the well for a particular DBRT structure if we specify the form of ζ, corresponding to different scattering mechanisms involved in the resonant tunneling process.

A.2 Breit-Wigner Form of the Transmision Probability in a Symmetrical DBRT Structure

Following [7] we show how the transmission probability for a symmetrical DBRT structure (see Fig. 2.1), obtained with the T-matrix formalism, can be expressed in a Breit-Wigner form at energies near resonances.

We start from Eq. (A1.12a) of the transmission coefficient, which is detailed here using Eqs. (A.8a) and (A.8b):

$$\begin{aligned} t(E)^{-1} &= M_{11}^{B_1} M_{11}^{B_2} + M_{12}^{B_1} M_{21}^{B_2} = \\ &= \frac{1}{16}\Big[\left(z_{12}z_{23}\chi_2\xi_2 + z_{12}^{+}z_{23}^{+}\chi_2\xi_2^{-1}\right)\left(z_{34}z_{45}\chi_3^{+}\chi_4\xi_4 + z_{34}^{+}z_{45}^{+}\chi_3^{+}\chi_4\xi_4^{-1}\right) + \\ &\quad + \left(z_{12}z_{23}^{+}\chi_2^{+}\xi_2 + z_{12}^{+}z_{23}\chi_2^{+}\xi_2\right)\left(z_{34}^{+}z_{45}\chi_3\chi_4\xi_4 + z_{34}z_{45}^{+}\chi_3\chi_4\xi_4^{-1}\right)\Big] \end{aligned} \qquad (A.22)$$

For the symmetrical DBRT structure $k_1 = k_3 = k_5 = k$, $\alpha_2 = \alpha_4 = \alpha$, $m_1^* = m_3^* = m_5^*$, $m_2^* = m_4^*$, $d_2 = d_4 = L_B$ and $d_3 = L_W$ from which results $\gamma_1 = \gamma_3 = \frac{m_W^*\alpha}{m_B^* k} = \gamma$ and $\gamma_2 = \gamma_4 = \frac{m_B^* k}{m_W^*\alpha} = \frac{1}{\gamma}$; $z_{12} = z_{34} = 1 + i\gamma$ and $z_{23} = z_{45} = 1 - \frac{i}{\gamma}$; $\chi_2 = e^{ikL_B}$, $\chi_3 = e^{ik(L_B+L_W)}$, and $\chi_4 = e^{ik(2L_B+L_W)}$; $\xi_2 = \xi_4 = e^{\alpha L_B}$. In these conditions Eq. (A.22) reads:

$$\begin{aligned} t(E)^{-1} &= \frac{1}{16}e^{2\alpha L_B}e^{ik(2L_B+L_W)}\Bigg\{\left[(1+i\gamma)\left(1-\frac{i}{\gamma}\right)+(1-i\gamma)\left(1+\frac{i}{\gamma}\right)e^{-2\alpha L_B}\right]^2 e^{-ikL_W} + \\ &+\left[(1+i\gamma)\left(1+\frac{i}{\gamma}\right)+(1-i\gamma)\left(1-\frac{i}{\gamma}\right)e^{-2\alpha L_B}\right]\left[(1-i\gamma)\left(1-\frac{i}{\gamma}\right)+(1+i\gamma)\left(1+\frac{i}{\gamma}\right)e^{-2\alpha L_B}\right]e^{ikL_W}\Bigg\} \end{aligned} \qquad (A.23)$$

After ordinary but elaborate algebraic manipulations (calculations) we obtain the folowing condensed form for the transmission coefficient:

$$t(E) = \left\{Q\,\frac{1}{\gamma}(1+i\gamma)(1-\frac{i}{\gamma})\left[L + 2\frac{1-i\gamma}{1+i\gamma}P\exp(-2\alpha L_B) + \left(\frac{1-i\gamma}{1+i\gamma}\right)^2 N\exp(-4\alpha L_B)\right]\right\}^{-1}, \qquad (A.24)$$

where

$$Q = \frac{\exp[ik(2\alpha L_B + L_W)]\exp(2\alpha L_B)}{8}, \qquad (A.25a)$$

$$L = (\gamma^2 - 1)\sin(kL_W) + 2\gamma\cos(kL_W),$$
$$P = (\gamma^2 + 1)\sin(kL_W), \tag{A.25b}$$
$$N = (\gamma^2 - 1)\sin(kL_W) - 2\gamma\cos(kL_W),$$

As observed from Eq. (A.24) the poles $\overline{E_n^0}$ of $t(E)$ generally are complex. They can be associated with the eigenvalues of the energy in the quantum well, i.e.:

$$\overline{E_n^0} = E_n - i\frac{\Gamma_n}{2} \tag{A.26}$$

where E_n and Γ_n are positive.

We analyse the behaviour of t(E) at real energies in the vicinity of $\overline{E_n^0}$ from which we obtain the Breit-Wigner form of the transmission probability. The poles of the right side of Eq.(A.24) in the complex energy plane are the roots of the following equation:

$$F(E) = L + 2\frac{1-i\gamma}{1+i\gamma}P\exp(-2\alpha L_B) + \left(\frac{1-i\gamma}{1+i\gamma}\right)^2 N\exp(-4\alpha L_B) = 0. \tag{A.27}$$

Using the transformation $1 + i\gamma = (1+\gamma^2)^{1/2}\, e^{i\eta}$ with η = arctgγ and separating the real and imaginary part Eq.(A.27) takes the form:

$$F(E) = U - iW = 0, \tag{A.28}$$

where

$$U(E) = L + 2M\exp(-2\alpha L_B)\cos(2\eta) + N\exp(-4\alpha L_B)\cos(4\eta),$$
$$W(E) = 2M\exp(-2\alpha L_B)\sin(2\eta) + N\exp(-4\alpha L_B)\sin(4\eta). \tag{A.29}$$

We consider that the individual barrier penetrabilities in the structure are small, otherwise any quasistable state will very rapidly decay through the barriers. In this "thick barrier" limit $\alpha L_B \gg 1$ and, therefore, $e^{-\alpha L_B} \ll 1$. It follows that $W(E) \ll U(E)$ except at values of E near the roots of $U(E) = 0$, which we call E_r. $U(E)$, as seen from Eqs.(A.29) and (A.25), is a linear combination of sin(kL_W) and cos(kL_W) with real coefficients for real E. For this reason, in general, its roots, E_r are real. Now, for any root E_r of $U(E)$=0, the corresponding root $\overline{E_n^0}$ of Eq. (2.8) is given to very good approximation by Newton's Rule:

$$F(\overline{E_n^0}) = F(E_r) + (\overline{E_n^0} - E_r)F'(E_r) + \ldots = 0, \tag{A.30a}$$

from which results:

$$\overline{E_n^0} \cong E_r - \frac{F(E_r)}{F'(E_r)} = E_r + i\frac{W(E_r)}{F'(E_r)}. \tag{A.30b}$$

It follows, comparing Eqs. (A.30b) and (A.26) that in the approximation of "thick barriers" the quantity E_r satisfying $U(E) = 0$ is the real part E_n of the associated complex zero $\overline{E_n^0}$ of $F(E)$ and that the imaginary part of $\overline{E_n^0}$,- $i\Gamma_n/2$ equals $iW(E)\,/\,F'(E)$, calculated for $E = E_r = E_n$. This last correspondence gives for Γ_n the expression:

$$\Gamma_n = \left.\frac{-4Pe^{-2\alpha L_B}\sin(2\eta)}{L'(E)}\right|_{E=E_n=E_r}, \tag{A.31}$$

where using the same "thick barrier" approximation, the term involving N has been dropped in $W(E)$ and we have kept from $F'(E_n)$ only $L'(E_n)$. In the frame of the same approximation the values of E_r to be used in Eq. (A.31) are given sufficiently accurately by solving the equation $L(E) = 0$ instead of the equation $U(E) = 0$ that originally specified E_r. In this sense, one observes that the energies E_r satisfying $L(E_r) = 0$ occur when the equations

$$\sin kL_W = \pm\frac{2\gamma}{\gamma^2+1},$$
$$\cos kL_W = \mp\frac{\gamma^2-1}{\gamma^2+1} \tag{A.32}$$

are satisfied. The above equations help us to obtain $P(E)$ and $L'(E)$, for $E = E_r$, as $P(E_r) = \pm 2\gamma$ and

$$L'(E_r) = \mp\frac{m_W^*}{\hbar^2 k}\left[L_W\left(\gamma^2+1\right)+\left(\gamma^2\frac{m_B^*}{m_W^*}+1\right)\frac{2}{\alpha}\right]. \tag{A.33}$$

Thus, the expression of Γ_n from Eq. (A.31) can be worked out thereby and the final result is:

$$\Gamma_n = \left.\frac{16\hbar^2 k\gamma^2\exp(-2\alpha L_B)}{m_W(1+\gamma^2)\left[(1+\gamma^2)L_W+\left(1+\frac{m_B}{m_W}\gamma^2\right)2\alpha^{-1}\right]}\right|_{E=E_n}. \tag{A.34}$$

Eqs.(A.24) and (A.27) imply that at any real energy $t(E)$ reads:

$$t(E) = \left[Q\frac{1}{\gamma}(1+i\gamma)(1-\frac{i}{\gamma})F(E)\right]^{-1}. \tag{A.35}$$

At real energies near the complex zero $\overline{E_n^0}$ of $F(E)$,

$$F(E) = F(\overline{E_n^0})+(E-\overline{E_n^0})F'(\overline{E_n^0})+\ldots\cong\left(E-E_n+i\frac{\Gamma_n}{2}\right)L'(E_n), \tag{A.36}$$

The factors other than $F(E)$ in Eq. (A.35) are slowly varying near $E = E_n = E_r$, and at such energies can be computed at $E = E_n$. Using Eqs. (A.25a), (A.36), and (A.33) in Eq. (A.35) we see that at real energies E near E_n

$$t(E) = \mp i\exp\left[-ik(2L_B+L_W)\right]\exp(-i\eta)\frac{\frac{1}{2}\Gamma_n}{E-E_n+i\frac{\Gamma_n}{2}} \tag{A.37}$$

and

$$T(E) = |t(E)|^2 = \frac{\left(\frac{1}{2}\Gamma_n\right)^2}{(E-E_n)^2+\left(\frac{\Gamma_n}{2}\right)^2}. \tag{A.38}$$

Eq. (A.38) represents the Breit-Wigner form of the transmission probability through a symmetrical DBRT structure at energies near those corresponding to resonances.

References

1. E. O. Kane, "Basic Concepts of Tunneling", in "*Tunneling Phenomena in Solids*", ed. E. Burnstein and S. Lundqvist, Plenum Press, New York,1969.
2. E. Merzbacher, "Quantum Mechanics", 2nd ed., chaps. 6 and 7, John Wiley&Sons, New York, 1970.
3. D. J. Ben Daniel and C. B. Duke, *Phys. Rev.* **152**, 683 (1966).
4. H. Joosten, "*Numerical Study of Coherent Tunneling in Double Barrier Resonant Tunneling Device*", Ph.D. Thesis, Dept. of Electrical Engineering, Tampere University of Technology, publ. no. 81, ISBN 951-721-755-2, 1991.
5. M. Jonson and A.Grincwajg, *Appl. Phys. Lett.* **51**, 1729 (1987).
6. M. Büttiker, *IBM J. Res. Develop.* **32**, 63 (1988).
7. E. Gerjuoy and D. D. Coon, *Superlatt. and Microstruct.* **5**, 305 (1989).

Appendix B

Details about the Technological Processes Involved in the Resonant Tunneling Device Fabrication

In this section a complete description of the processing steps used for the fabrication of the resonant tunneling devices is presented. The presentation contains also information about the chemicals involved in different stages of the processing chain, equipments and special requirements. The step recipes are grouped in the frame of the same subdivision introduced in chapter 3 for the technological processing.

All the processes, except gold plating, have been accomplished in a clean room with high level of cleanliness. Thus, in the spaces destinated for sample handling in different processing steps - lithographic processes, wet chemical etching, loading for metallization, annealing or plasma processes, etc.- a class-100 is ensured. The other spaces have a class < 10,000. In the photolithographic (yellow) room there is a strict control of the temperature and humidity: 21 °C ± 1 °C and 40% ± 5%, respectively. The temperature is controlled also in the chemistry room (21 °C ± 2 °C) and in the spaces destinated to contact fabrication and plasma etching and deposition (21 °C ± 2 °C). Where necessary, for cleaning and drying purposes, nitrogen of high purity (5N) is delivered, using pistol tools.

The solutions used in different steps are in general obtained from many components, so their composition is specified indicating the proportion in volumes from each component. After each process a visual inspection is performed with a Polyvar optical microscope, which permits a maximum magnification of x 1000. The dimensions of the patterns transferred into the resist as well as of those subsequently defined by the subtractive and additive operations have been checked using the same microscope or, when necessary, the scanning electron microscope (SEM), Jeol JSM 6400F. Also, the thickness of the resist films and of the different deposited layers , or the etching depths were measured with a Tencor step profiler or the SEM.

B.1 Cleaning Processes

Some useful properties of solvents, used in cleaning processes and other several processing steps, are listed in Table B.1. The data are from reference [31], chapter 3, and sheet specifications of the producers.

The isothermal control of the hot plates, used for heating or drying, has been of ±1 °C. Rinse has been performed first in reverese osmosis (RO) water and than in deionized (DI) water, with resistivity higher than 6 MΩ. In both cases, the water has been kept flowing.

Property / *Solvent*	*Chemical Formula*	*Boiling Point [°C]*	*Density [g/ml]*	*Water Solubility*
2-Propanone (Acetone)	CH_3COCH_3	56.2	0.7899	100 %
2-Propanol (Isopropyl alcohol)	$CH_3CH(OH)CH_3$	82.4	0.7855	100 %
Trichloromethane (Chloroform)	$CHCl_3$	61.7	1.4832	<1%
Methanol (Methyl alcohol)	CH_3OH	64.96	0.791	100 %
Ethanol (Ethyl alcohol)	C_2H_5OH	78.5	0.7893	100 %
Trichloroethene (Trichloroethylene)	C_2HCl_3	87.2	1.45	< 1 %

Table B.1 Physical properties of the solvents used in resonant tunneling device processing.

❑ Cleaning process before photolithography:

- Immersion (with agitation) in boiling *acetone*: 2 min;
- Immersion (with agitation) in boiling *2-propanol*: 1 min;
- Immersion (with agitation) in boiling *chloroform*: 2 min;
- Hot plate drying: 105°C / 5 min;
- Visual inspection.

(B.1)

❑ Cleaning process before metallization:

- Immersion in *HCl*:H_2O (1:4) solution: 20°C / 30 s;
- Rinse in *RO water*: 1 min;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

(B.2)

❑ Cleaning process after photolithography:

- *Oxygen plasma* etching: O_2flow = 30 sccm/p = 20mTorr/ $P_{\mu w}$ = 100 W / DC bias = -167 V/ $\succ$1 min.
- Visual inspection.

(B.3)

❑ Cleaning process for removing (stripping) the resist after an etch process:

- Immersion in boiling *acetone*: 5 min
- Immersion (with agitation) in *2-propanol*: 1 min.
- Rinse in *RO water*: 30 s.
- Rinse in *DI water*: 2 min.
- *Nitrogen* blow drying.
- Visual inspection.

(B.4)

B.2 Photolithographic Processes

Helpful properties of the AZ type positive and dual image resists and developer, used in our lithographic processings, are shown in Table B.2. The data proceed mainly from sheet specifications.

Property *Resist*	*Solids Content (%)*	*Water Content (%)*	*Solvent Type*	*Thickness [μm] at 3000/5000 rpm*	*Spectral Sensitivity [nm]*
AZ 1505	17.7	max. 0.5	PGMEA	0.58/0.45	310/440
AZ 1350J	29.5	max. 0.5	EGMEA + Xylene	2.08/1.61	365-436
AZ 4330A	34.5	max. 0.5	EGMEA + Xylene	3.81/2.95	310/440
AZ 4620A	39.5	max. 0.5	EGMEA + Xylene	7.16/5.54	310/440
AZ 5206E (IR)	18.0	max. 0.5	PGMEA	0.69/0.64	310/420
AZ 5214E (IR)	28.3	max. 0.5	PGMEA	1.63/1.26	310/420

(a)

(continued)

Property *Developer*	*Alkali Type*	*Density (20 °C)* *[kg/l]*	*Normality (conc.)*
AZ Dev.	$Na^{\oplus}$ based	1.05 ± 0.01	0.600 ± 0. 005 N

(b)

Table B.2. Main properties of the AZ type positive and dual image photoresists (a) and developer (b), used for the lithographic processings. In (a) "IR" stands for Image Reversal and indicates that the corresponding resist is a dual image one.

The application of the resist film has been performed with a Headway Research spinner, having continuos speed variation between 0-10,000 rpm and variations in the constant speed regime of ±100 rpm.

The SUSS MJB 3 UV 400/300 mask aligner, used for our lithographic necessities, is equipped with a 350 W mercury short-arc lamp, providing primary exposure wavelengths in the range 350-450 nm. This is endowed with a high precision alignment stage allowing alignment accuracies to 0.1 microns. We have worked with it in the contact printing technique, clamping the sample against the mask with vacuum, hard contact mode (HCM), or mechanically, soft contact mode (SCM). The exposures around UV h-line (405nm) and i-line (365) can be performed using specific filters, UV 400 and UV 300, respectively. Normally, with a good lamp, the light intensity for the two UV lines, in each case of the two filters, is approximately:

- UV 400 : 19 mW/cm^2 for h-line and 8mW/cm^2 for i-line;
- UV 300 : 3 mW/cm^2 for h-line and 8mW/cm^2 for i-line

☐ Photolithographic process using the AZ 1505 resist:

- Application of the resist film by spinning: low speed ➢ 1 s / 3000 rpm ➢ 30 s;
- Soft bake on the hot plate: 95°C ➢ ~5 min;.
- Exposure : HCM / UV 400 ➢ 4 s;
- Development in *AZ Dev:DI water* (1:1): MA ➢20 s; (B.5)
- Rinse in *RO water*: 1 min;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

☐ Photolithographic process using the AZ 1350J resist:

- Application of the resist film by spinning: low speed ➢ 3 s / 5000 rpm ➢ 30 s;
- Soft bake on the hot plate: 95°C ➢ 10 min;. (B.6)
- Exposure: HCM / UV 400 ➢ 12 s;

(continued)

- Development in *AZ Dev:DI water* (1:1): MA ➢ ~50 s;
- Rinse in *RO water*: 1 min;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

❑ Photolithographic process using the AZ 4330A resist (specific for air bridge fabrication by plating):

- Application of the resist film by spinning: low speed ➢ 5 s / 5000 rpm ➢ 30 s;
- Soft bake on the hot plate: 95°C ➢ 15 min;.
- Exposure : SCM / UV 400 ➢ 25 s;
- Development in *AZ Dev:DI water* (1:1): MA ➢ ~120 s;
- Rinse in *RO water*: 1 min;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Hard bake: 120°C ➢ 120 min;
- Visual inspection.

(B.7)

❑ Photolithographic process using the AZ 4620A resist (specific for air bridge fabrication by plating):

- Application of the resist film by spinning: low speed ➢ 5 s / 5000 rpm ➢ 30 s;
- Soft bake on the hot plate: RT ➢ 10 min / 95°C ➢ 15min / RT ➢ 30 min;
- Exposure : HCM / UV 400 ➢ 45 s;
- Development in *AZ Dev:DI water* (1:1): MA ➢ ~4 min;
- Rinse in *RO water*: 1 min;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Hard bake: 110 °C ➢ 30 min;
- Visual inspection.

(B.8)

❑ IR photolithographic process using the AZ 5206E resist, suitable for lift-off metallization technique:

- Application of the resist film by spinning: low speed ➢1 s / 3000 rpm ➢ 30 s;
- Prebake on the hot plate: 100°C ➢ 3 min;.
- Exposure : HCM / UV 300 ➢ 6 s;
- Post exposure bake on the hot plate: 115°C ➢ 2 min;
- Flood exposure: UV 400 ➢ 8 s;
- Development in *AZ Dev.:DI water* (1:1): MS / 20°C ➢ ~30 s (visual survey);
- Rinse in *RO water*: 30 s;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

(B.9)

☐ IR photolithographic process using AZ 5214E resist, suitable for lift-off metallization technique:

- Application of the resist film by spinning: low speed ➢ 3 s / 5000 rpm ➢ 30 s;
- Prebake on the hot plate: 100°C ➢ 3 min;.
- Exposure : HCM / UV 300 ➢ 10 s;
- Post exposure bake on the hot plate: 115°C ➢ 2 min;
- Flood exposure: UV 400 ➢ 12 s;
- Development in *AZ Dev.:DI water* (1:1): MS / 20°C ➢ ~50 s (visual survey);
- Rinse in *RO water*: 30 s;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

(B.10)

☐ Direct photolithographic process using AZ 5206E resist, suitable for lift-off metallization technique:

- Application of the resist film by spinning: low speed ➢ 1 s / 3000 rpm ➢ 30 s;
- Prebake on the hot plate: 95°C ➢ 5 min;.
- Flood exposure: UV 300 ➢ 2.5 s;
- Post exposure bake on the hot plate: 105°C ➢ 5 min;
- Exposure: HCM / UV 300 ➢ 14 s;
- Development in *AZ Dev.:DI water* (1:1): MS / 20°C ➢ ~50 s (visual survey);
- Rinse in *RO water*: 30 s;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

(B.11)

☐ Direct photolithographic process using AZ 5214E resist, suitable for lift-off metallization technique:

- Application of the resist film by spinning: low speed ➢ 3 s / 5000 rpm ➢ 30 s;
- Prebake on the hot plate: 95°C ➢ 5 min;.
- Flood exposure: UV 300 ➢ 3.5 s;
- Post exposure bake on the hot plate: 105°C ➢ 5 min;
- Exposure: HCM /UV 300 ➢ 24 s;
- Development in *AZ Dev.:DI water* (1:1): MS / 20°C ➢ ~1 min 30 s (visual survey);
- Rinse in *RO water*: 30 s;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

(B.12)

☐ Direct photolithographic process using AZ 5214E resist, suitable for the obtaining of very small deviations from the mask dimensions and specific for the dielectric resist support:

- Application of the HDMS film by spinning: 3000 rpm ➢ 30s;
- Application of the resist film by spinning: low speed ➢ 3 s / 5000 rpm ➢ 30 s;
- Prebake on the hot plate: 100°C ➢ 5 min;.
- Flood exposure: UV 300 ➢ 1.2 s;
- Post exposure bake on the hot plate: 115°C ➢ 1 min;
- Exposure: HCM / UV 300 ➢ 7 s;
- Development in *AZ Dev.:DI water* (1:1): MS / 20°C ➢ ~ 50 s (visual survey);
- Rinse in *RO water*: 30 s;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

(B.13)

B.3 Wet Chemical Etching Processes

Table B.3 shows the initial concentrations for the reagents used in our cleaning and wet chemical etching processes.

Reagent	*Percentage of Weight*
HCl	37
NH_4OH	29 (as NH_3)
H_2SO_4	98
H_3PO_4	85
H_2O_2	30
CH_3OH	99.9
HF	48

Table B.3. Initial concentration of the reagents used in cleaning and wetchemical etching processes.

In the following etching processes the indicated etch rate (r^{sol}_{etch}) is for GaAs material which is prevalent in our DBRT structures (see sections 3.2.3 and 4.2).

❑ Wet chemical etching process using H_3PO_4:H_2O_2:CH_3OH (1:1:3) solution:

- Preparation of the solution: 40 ml H_3PO_4 + 40 ml H_2O_2 + 120 ml CH_3OH;
- Cooling of the solution in a thermostatic bath to reach the constant *etch temperature* : setpoint 0°C ➢ ~ 1 h;
- Etching process: 0°C / *without agitation* / $\boxed{r^{sol}_{etch} = 0.34\ \mu m/min}$;
- Rinse in *RO water*: 1 min;

(B.14)

(continued)

- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection;
- Etch depth determination.

☐ Wet chemical etching process using $H_3PO_4:H_2O_2:H_2O$ (6:3:100) solution:

- Preparation of the solution: 12 ml H_3PO_4 + 6 ml H_2O_2 + 200 ml H_2O;
- Waiting time for solution homogenization and thermal equilibrium (RT): 20 min;
- Etching process: RT / *with* slow magnetic *stirring* / $\boxed{r_{etch}^{sol} = 0.110\ \mu m/min}$; (B.15)
- Rinse in *RO water*: 1 min;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection;
- Etch depth determination.

B.4 Metallization Processes

Metal depositions for device contacts, pads and plating necessities have been made in the single wafer load-locked Leybold - L 560 UV - evaporator. This equipment has two electron beam evaporators and an ion gun, which serves to semiconductor surface cleaning purposes. The base pressure obtained in the deposition chamber is around 3.0×10^{-8} mbar. Between the sample holder and the e-gun there exists a 50 cm distance to minimise heating up of the resist. The metal deposition rate (r_{dep}) is adjusted by changing the power supplied to e-gun. All the operations can be manually or automatically controlled. The automatic control is ensured by specific software working under the Windows program, on a PC. Metallization annealing has been performed in a Rapid Thermal Annealer - AST SHS 100. As in the case of the evaporator, the annealing process can be controlled both manually and automatically. In general, the process is accomplished in specific atmosphere, using gases (hydrogen, nitrogen or argon) with a high purity level (5N).

☐ Metal deposition process for n-type ohmic contacts on GaAs:

- Loading of the sample and establishing of the *high vacuum* conditions: 3×10^{-8} mbar;
- Evaporation of *specific metallization*: 20 nm Ge ➢ r_{dep} = 12 nm/min /
  15 nm Ni ➢ r_{dep} = 12 nm/min /
  200 nm Au ➢ r_{dep} = 12 nm/min the first 10nm (B.16)
  and
  24 nm/min the rest;
- Unloading of the sample;
- Visual inspection.

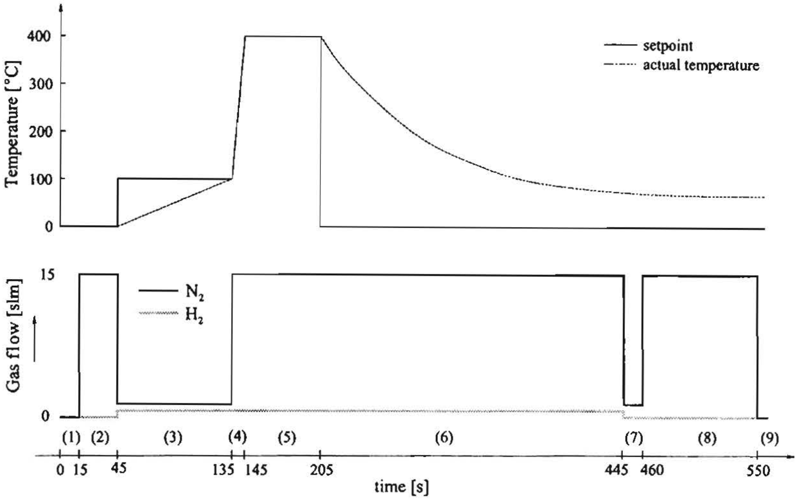


Fig. B.1 Schematic overview of the annealing process used for the Ge/Ni/Au ohmic contact formation.

❑ Lift-off process:

- Immersion (with agitation) in boiling *acetone*: 2 min;
- Immersion (with agitation) in boiling *2-propanol*: 1 min;
- Rinse in *RO water*: 30 s.
- Rinse in *DI water*: 2 min.
- *Nitrogen* blow drying.
- Visual inspection.

(B.17)

For the annealing of the above metallization, the specific process shown diagrammatically in Fig. B1 have been used. Details of the process are presented in B.18.

❑ Annealing process of the Ge/Ni/Au metallization:

- Chamber evacuation (1): 15 s;
- Nitrogen purge (2): 15 slm flow ≻ 30 s;
- Preheating in N_2/H_2 atmosphere (3): setpoint 100°C/ N_2 flow = 1 slm + H_2 flow = 0.5 slm ≻ 90 s;
- Ramp up to the annealing temperature (4): setpoint 400°C/N_2 flow = 1 slm + H_2 flow = 0.5 slm ≻ 10 s;
- Stable at the annealing temperature (5): setpoint 400°C/N_2 flow = 1 slm + H_2 flow = 0.5 slm ≻ 60 s;
- Ramp down (6): setpoint 0°C/N_2 flow = 1 slm + H_2 flow = 0.5 slm ≻ 240 s;
- Chamber evacuation (7): 15 s;
- Nitrogen purge (8): 15 slm ≻ 90 s;
- End (9): 3sec.

(B.18)

☐ Metal deposition process for p-type ohmic contacts and Schottky barriers on GaAs:

- Loading of the sample and establishing of the *high vacuum* conditions: $3\text{x}10^{-8}$ mbar;
- Evaporation of *specific metallization*: 50 nm Ti ➢ r_{dep} = 12 nm/ min /
  20 nm Pt ➢ r_{dep} = 12 nm/min /
  200 nm Au ➢ r_{dep} = 12 nm/min the first 10nm (B.19)
  and
  24 nm/min the rest;
- Unloading of the sample;
- Visual inspection.

B5. Thin Film SiO_2 Plasma Deposition and Etching Processes

☐ SiO_2 deposition process by PECVD:

- Loading of the sample;
- SiO_2 deposition: 300°C /p = 100 mTorr / $P_{\mu w}$ = 200 W / $\boxed{r_{dep}^{PECVD} = 60\ \mathrm{nm\ /\ min}}$;
  / upper manifold: N_2O flow = 40 sccm /
  N_2 flow = 90 sccm/ (B.20)
  / lower manifold: SiH_4 flow = 4 sccm/
  N_2 flow = 30 sccm;
- Unloading of the sample;
- Visual inspection.

☐ Annealing process of the SiO_2 film:

It is similar with the process B.18 except the steps 4 and 5:
- Ramp up to the annealing temperature (4): setpoint 450°C/N_2 flow = 1 slm +
  H_2 flow = 0.5 slm ➢ 10 s; (B.21)
- Stable at the annealing temperature (5): setpoint 450°C/N_2 flow = 1 slm +
  H_2 flow = 0.5 slm ➢ 600s;

☐ SiO_2 dry etching process using ECR plasma:

- Loading of the sample;
- SiO_2 dry etching: RT / p = 0.6 mTorr / $P_{\mu w}$ = 200 W / $\boxed{r_{etch}^{ECR} = 20\ \mathrm{nm\ /\ min}}$;
  /upper manifold: SF_6 flow = 10 sccm /
  CFH_3 flow = 2 sccm /
  Ar flow = 2 sccm / (B.22)
  / upper magnet current = 140 A /
  / lower magnet current = 80 A /
  / DC selfbias = 24 V /
- Unloading of the sample;
- Visual inspection.

B.6 Selective Etching Processes

☐ Selective wet chemical etching process for GaAs over AlAs using $C_6H_8O_7$:H_2O_2 (4:1) solution:

- Preparation of the $C_6H_8O_7$ solution: 200 g $C_6H_8O_7 \bullet H_2O$ + 200 ml H_2O /
/ dissolving of the $C_6H_8O_7 \bullet H_2O$ using magnetic stirring ➢ at least 2 h;
- Preparation of the *selective etching* solution: 160 ml $C_6H_8O_7$ solution + 40ml H_2O_2;
- Waiting time for solution homogenization and thermal equilibrium (RT): 20 min;
- Etching process: RT / *with* slow magnetic *stirring*; (B.23)
- Rinse in *RO water*: 1 min; $\boxed{r_{etch}^{sol} = 0.450\ \mu m / min}$ for GaAs
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection; $\boxed{\text{selectivity GaAs / AlAs} \cong 1500}$;
- Etch depth determination.

☐ Selective wet chemical etching process for AlAs over GaAs using *HF* solution, destinated to remove the first barrier of the resonant tunneling structure:

- Preparation of the selective *etching solution*: 2 ml conc. *HF* (48 %) + 998 ml *DI water*;
- Cooling of the solution in a thermostatic bath to reach the constant *etch temperature* :
setpoint 0°C ➢ ~ 1 h;
- Etching process: 0°C / *short deep* of the sample in the solution; (B.24)
- Rinse in *RO water*: 1 min;
- Rinse in *DI water*: 2 min;
- *Nitrogen* blow drying;
- Visual inspection.

☐ Selective RIE process for GaAs over AlAs using pure CCl_2F_2:

- Loading the sample;
- Etching process: RT / p = 6.7 mTorr / $P_{\mu w}$ = 24W /
/ CCl_2F_2 flow = 12 sccm/ $\boxed{r_{etch}^{RIE} = 0.2\ \mu m / min}$ for GaAs;
/ DC bias = - (29-30) V / (B.25)
- Unloading the sample;
- Visual inspection; $\boxed{\text{selectivity GaAs / AlAs} \cong 450}$;
- Etch depth determination.

B.7 Air Bridge Interconnection Fabrication Processes

☐ Au layer deposition process for plating current conduction:

It is the same as the process B.17 but only a 100 nm Au layer is deposited. (B.26)

The optimised operating conditions for the PUR-A-GOLD 202B bath, which we have used for the plating process are specified in the Table B4, as they are indicated by the producer.

Parameter	*Value*
○ Metallic gold content	12 g/l or more
○ pH electrometric	6.0 (5.5 - 6.5) at 60°C
○ Specific gravity	Minimum 17° Baumé
○ Temperature	60°C (55 - 70 °C)
○ Agitation	Uniform solution agitation
○ Anode to cathode ratio	2 to 1 or higher
○ Cathode current density	0.3 A/dm^2 (0.1 - 0.6 A/dm^2)
○ Deposition rate	112 mg/(A·min) or higher

Table B.4 Operating conditions for PUR-A-GOLD 202B plating bath, produced by Enthone-Omi.

☐ Gold plating process for air bridge interconnections:

- Preparing of the *solution*: 10 g/l Au / pH = 5.7 ;
- Gold *plating* process: 40°C / magnetic stirring ➢ 600 rot/min / J = 0.155 A/dm^2 /
- Rinse in *RO water*;
- Rinse in *DI water*; $\boxed{r_{dep}^{plt} = 0.06\mu m / min}$; (B.27)
- *Nitrogen* blow drying;
- Visual inspection;
- Deposition thickness determination.

☐ Removing of resist layers after gold plating:

- Exposure without mask: UV 400 ➢ 360 s;
- Development in AZ Dev.:DI water (3:1) ➢ 2 min;
- Rinse in *RO water*: 1min; (B.28)
- Rinse in *DI water*: 2min;
- Removing of the thin gold layer (etching) in Degussa® solution;
- Removing of the second resist layer in boyling DMS solution ➢ 2min.

B.8 Design of Masks

In Table B5 an estimation of the dimensional and misalignment deviations for the resists used in our DBRT device processing is presented. Further on, Table B6 shows the underetch produced by the wet chemical processes involved in the device fabrication.

Resist	*δ [μm]*	*Δx (Δy) [μm]*
AZ 1505	0.2	0.2
AZ 1350J	0.8	0.4
AZ 4330A	1.5	1
AZ 4620A	2.5	1.5
AZ 5206E (IR)	0.3	0.2
AZ 5214E (IR)	0.6	0.3

Table B.5 Dimensional and misalignment deviations for the photoresits used to the lithographic processes.

Process	*UnderEtch [μm]*
B.15	~ 75% from etch depth (AZ 1500 resist)
B.16	~ etch depth (AZ 1350J resist)

Table B.6 Estimated underetch produced in the wet chemical processes.

Summary

Double barrier resonant tunneling (DBRT) devices belong to the new generation of (opto) electronic devices, made from III-V compound semiconductors, whose properties originate in quantum effects. They have developed due to the progress registered in the modern epitaxial growth techniques, namely Molecular Beam Epitaxy (MBE) and Metal Organic Vapour Phase Epitaxy (MOVPE), and represent a real touchstone for the quality of the layers grown with these techniques. The increased interest in DBRT devices is stimulated by potential applications in high-speed electronics, e.g. high-frequency oscillators and fast switching devices (sampling, triggering, and logic applications). During our work we have studied some physical aspects of the two terminal (2T) DBRT device and proposed a solution for a three terminal (3T) DBRT device with a base contact to the quantum well. To realize this, a comprehensive approach of the problems has been developed, from the technological aspects of the device processing to the electrical measurements and then to the physical modeling, with possible and inherent interconnections between them.

The structures used for the processing of both 2T and 3T DBRT devices have been grown by MBE at the Solid State Physics Department of the Eindhoven University of Technology, from the GaAs/AlGaAs material system. All the technological processes involved in the device fabrication have been investigated in detail and optimised for our clean room conditions and specific design characteristics of the devices. We paid special attention to the dual image resist photolithography for lift-off technique, wet chemical etching, selective etching of GaAs and AlGaAs, and gold plating for air bridge interconnections. For the 3T DBRT device processing we have designed a specific mask set which permits the fabrication of 2T DBRT devices as well. The mask design provides planar processing of devices with several dimensions and air bridge interconnections between the upper contacts and the measurement pads. The pads have special geometry suitable for on-wafer microwave measurements using Cascade probes.

2T DBRT devices have been fabricated in planar configuration from GaAs/AlAs and GaAs/$Al_{0.4}Ga_{0.6}As$ materials, with both on-surface and air bridge interconnections between the top contact and the corresponding pad. DC measurements have been accomplished at room temperature and below, down to 15 K. The room temperature measurements indicate similar performances of the GaAs/AlAs based devices to those reported in the literature with the same designed structure. They have been used for a comparative analysis of the two types of processed devices. The measurements at low temperatures could be interpreted in connection with possible tunneling mechanisms. We made use also of the physical modeling of the DBRT diode I-V characteristic to describe the DC results, qualitatively and quantitatively. On-wafer microwave measurements have been performed and analysed in correlation with the DC device properties.

3T DBRT device fabrication and characterization represent the main goal of our work. For

this kind of devices only GaAs/AlAs based structures have been used. The technological solution is based on highly selective etching processes to reach the plane of the very thin quantum well (5.0 nm) for base contacting. Concerning the selective etching, we have investigated both dry and wet processes. The devices have been tested in DC conditions at room temperature and below, down to 80 K. The results demonstrate clearly that the base contact lies on the level of the DBRT quantum well and the collector-emitter I-V characteristic is influenced by the potential of the base terminal. They are explained, qualitatively, making use of the physical picture (energy band diagram) of the DBRT structure. The device transfer characteristic (collector current function of base current) shows abrupt switching properties. Using microwave measurements, the conclusions established under DC conditions have been confirmed. We have used the third device terminal to measure the switching time of our DBRT structure in a much more simple way than previously reported efforts. The experimental value for this parameter (~ 42 ps) is in accordance with the theoretical estimation using the DC properties of the device. All the results show that direct contacting of the DBRT quantum well is possible. The optimization of our procedure will result in new possibilities for the investigation of the resonant tunneling phenomena and development of important applications in high-speed electronics.

Samenvatting

Double Barrier Resonante Tunneldioden (DBRT's) behoren tot de nieuwe generatie van (opto-) elektronische componenten gefabriceerd van III-V verbindings halfgeleidermaterialen, waarvan de eigenschappen bepaald worden door quantumeffekten. Deze konden dankzij de vooruitgang in moderne epitaxiale groeitechnieken, met name de Moleculaire Bundelepitaxie (MBE) en de Metaalorganische Gasfase-epitaxie (MOVPE) worden ontwikkeld en representeren een toets voor de kwaliteit van de lagen die met behulp van deze technologieNn worden vervaardigd. De toegenomen interesse in DBRT's wordt mede bepaald door de potentiNle toepassingen bij hoogfrequente elektronica, zoals oscillatoren en schakelende bouwstenen (sampling, triggeren en logische toepassingen). Tijdens het promotie onderzoek werden enkele fysische aspekten van de DBRT-diode met twee contacten (2T - DBRT) bestudeerd en een voorstel uitgewerkt voor de realisatie van een DBRT-device bouwsteen met drie contacten (3T - DBRT) met de quantumput als basiscontact. Om deze bouwsteen te kunnen realiseren zijn uitgebreide studies op het gebied van de halfgeleidertechnologie, de elektrische karakterisatie en het fysische modelleren uitgevoerd, mede met aandacht voor mogelijke en inherente afhankelijkheden.

De epitaxiale GaAs/AlGaAs-strukturen voor de 2T en 3T DBRT's zijn gegroeid met behulp van de MBE-faciliteit bij de Leerstoel Vaste Stof Fysica, Faculteit Technische Natuurkunde, TUE. Alle technologische aspekten die gerelateerd zijn aan de bouwsteenfabricage zijn in detail onderzocht en geoptimaliseerd in de "cleanroom" van de Leerstoel Elektronische Bouwstenen, Faculteit Elektrotechniek, TUE. Met als centraal aspekt het ontwerpen van de componenten zijn in het bijzonder de fotolithografie voor lift-off processen, natchemische etstechnieken voor het selectief etsen van GaAs resp. AlGaAs en de elektrolytische gouddepositie voor luchtbruggen onderzocht. De resultaten zijn gebruikt om een maskerset te realiseren waarmee planaire 2T en 3T bouwstenen met verschillende afmetingen kunnen worden vervaardigd. De contact-verbinding met behulp van een luchtbrug maakt deel uit van een van de ontwerpen. De contactgeometrie staat toe om microgolfmetingen op de chip uit te voeren (on-wafer design).

Planaire 2T GaAs/AlAs en GaAs/$Al_{0.4}Ga_{0.6}As$ DBRT's zijn gerealiseerd zowel met conventionele contacten, als ook met luchtbrugtechnologie. Bij kamertemperatuur (en lager) zijn DC metingen uitgevoerd. De resultaten van metingen bij 300 K aan GaAs/AlAs DBRT's zijn, bij vergelijkbare geometrie, in overeenstemming met literatuurwaarden. De gegevens vormden de basis voor de vergelijkende analyse van de twee typen van componenten. De metingen bij lagere temperatuur (tot 15 K) zijn benut voor de studie naar mogelijke tunnelmechanismen. Om de resultaten van de DC metingen kwalitatief en kwantitatief te verklaren,is gebruik gemaakt van de resultaten van het fysisch modelleren van de I-V karakteristiek van de DBRT's. On-wafer microgolfmetingen zijn eveneens uitgevoerd en geanalyseerd.

De realisatie en karakterisatie van 3T DBRT's was het hoofddoel van dit werk. Voor dit type bouwsteen zijn alleen epitaxiale GaAs/AlAs-structuren gebruikt. De technologische realisatie kon alleen bewerkstelligd worden door de inzet van zeer selectieve etsprocessen, aangezien het contacteren van de uiterst dunne quantum put (5 nm) essentieel is voor het basiscontact. Hiertoe is zowel de natchemische- als ook de beschikbare droge etstechnologie onderzocht. Ook deze bouwstenen zijn in het temperatuurbereik (80 - 300 K) onderzocht. De verkregen resultaten tonen zonder enige twijfel aan, dat de basis gecontacteerd is aan de quantum put en de collector-emitter I-V-karakteristiek beïnvloed wordt door de elektrische potentiaal aan de basis. De resultaten konden kwalitatief verklaard worden met behulp van het energieschema van de DBRT-structuur. De overdrachts-karakteristiek (de collectorstroom als funktie van de basisstroom) toont de voorspelde, abrupte schakelkarakteristiek. Met behulp van microgolfmetingen konden de conclusies uit de DC metingen worden bevestigd. De schakeltijd van de DBRT kon met behulp van het basis contact, op een veel eenvoudiger manier bepaald worden dan eerder in de literatuur aangegeven is. De experimentele waarde (~ 42 ps) is in overeenstemming met de theoretisch geschatte waarde.

De conclusie is, dat het rechtstreeks kontakteren van DBRT quantum putten dus praktisch mogelijk is. De resultaten van dit onderzoek kunnen benut worden voor verder onderzoek naar resonante tunnelfenomenen en bijdragen tot de ontwikkeling van belangrijke toepassingen op het gebied van snelle elektronika.

Curriculum Vitae

Mihail Lepsa was born in October 18, 1955 in Braşov, Romania. He has graduated from the Lyceum Andrei Şaguna from Braşov in 1974.

In 1975 he started to study at the Bucharest University, Faculty of Physics. In 1980 he obtained the Master of Science degree in the field of Biophysics on the subject of charge transfer in biological systems.

Between 1980-1983 he was teacher of Physics, high school level, in Bucharest, Romania.

Starting from 1984 he became a researcher in the Research Institute for Material Physics, Bucharest. Here he was involved mainly in the technology and characterization of semiconductor laser diodes grown by liquid phase epitaxy.

In 1992 he spent half a year at Eindhoven University of Technology, Electronic Devices Group, within an EC-Tempus project, on the subject of double barrier resonant tunneling diodes. One year later, in February 1993 he started as a PhD student in the same group. He worked in the field of the fabrication and characterization of two and three terminal double barrier resonant tunneling devices.

From February 1997 he is back at the Institute for Material Physics where he is continuing the PhD activity in the field of quantum devices.

Stellingen

behorende bij het proefschrift van
Mihail Ion Lepsa

Eindhoven, 12 November 1997

I

Besides careful experimental development, scrupulous attention to detail, and/or stringent process control, theoretical knowledge of process mechanisms is very important to guarantee the success in the field of semiconductor device processing.

II

The great discoveries in science which marked the spectacular progress of the human civilization demonstrate that the imagination is more important than knowledge*. However this must be always connected to the reality.

*Albert Einstein

III

After more than sixty years of different dictatorship regimes, the recovery of the Romanian civil society depends on the drastic change of the civic mentality and conscience and the activation of the civic potential energy of the civil society as a whole and of each member as an active and aware part of it.

IV

Until now the evolution of the compound semiconductor devices has been inevitably marked by an important and insurmountable restriction: the one-side sample processing. Despite the fantastic progress obtained under the frame of this concept, it seems that the limits have been attained in this direction and a new device fabrication principle is necessary.

V

The opposite side processing of the thin semiconductor film containing the device structure represents a realistic alternative for further developments in the device technology.

*Serge Luryi, *IEEE Trans. Electron Devices* **41**, 2241 (1994).

VI

In everybody both inferiority and superiority complexes are manifested. This is more evident for those who lived under a political dictatorship.

VII

As the Netherlands after the second war, Rômania can recover by hard work and external help. The external help has to be not only financial or economical, but equally political and of integration, as recognition of Romania as a by right member of the European countries family.

VIII

We can not have in the same time all what we had in the past and what we have in the present, but we may wish it. What is important is to know at what to renounce.

IX

All the friends quarrel but the good ones reconcile.

X

As a physicist I am envious how the poet express in such a wonderful manner abstract concepts:

Time present and time past
Are both perhaps present in time future,
And time future contained in time past.
If all time is eternally present
All time is unredeemable.
What might have been is an abstraction
Remaining a perpetual possibility
Only in a world of speculation.
What might have been and what has been
Point to one end, which is always present.

T. S. Eliot "*Four Quartets*" Faber and Faber, London, 1974.