THESE

Présentée à l'Université des Sciences et Technologies de Lille

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE
Spécialité: ELECTRONIQUE

par

Adel MIKHAIL

Simulation des effets tunnel résonnant dans les nanostructures de semiconducteurs

Soutenue le: Mai 1998 devant la commission d'examen

Membres du Jury: M. R. Fauquembergue
M. D. Lippens — Directeur de thèse
M A. de Lustrac — Rapporteur
M J. C. Vaissière — Rapporteur
M J.C. De Jaeger

INTRODUCTION

General introduction :

The resonant tunnelling effect in semiconductors is known since the pioneering work of L.Esaki in 1958 followed by the first experimental demonstration of Negative Differential Conductance (NDC) effect at liquids Nitrogen. Now, with the recent advance in epitaxial growth and modern lithography techniques a wide variety of devices exhibiting such NDC effect are fabricated. With respect to conventional NDC devices including the so called Esaki tunnel diode the major advantages stem from quantum resonance effect. Indeed, it is well known that by means of this quantum resonance the transmission probability can reach unity with a selective character. Basically, the observation of a resonant tunnelling effect needs to fabricate a quantum-sized system which can be supplied or probed by a tunnelling mechanisms. For this very general view point, the Double Barrier Heterostructure (DBH) appears as the generic nano-structure. By using hetero-junction rather than homo-junction we take advantage of numerous degrees of freedom in device design. So that band-gap engineering can be carried-out. Therefore with the goal to use these quantum devices in electronic and optoelectronic applications, a large variety of relevant figure of merits can be optimised depending on the targeted application. For instance, the intrinsic response time in direct connection with the life time of carriers within the quantum well is extremely short provided the various barriers are sufficiently thin.

The NDC effect which results from a resonant tunnelling device gained the attention of many functional component designers. The ultra-high speed of resonant tunnelling effect is the primary motivation. On the other hand such an effect allows them to integrate complex functions into a single functional device. Also should be added that one can implement a single or a multiple NDC-region device or circuit in order to realise a multi-functional device.

On the other hand, the current voltage characteristics can be tailored particularly at convenience. In this context, there is an important need for accurate simulation tools with the goal to computer aided design of these specific quantum structures. In a real device, the conduction mechanisms are much more complicated than the simple view of tunnelling electron through the quantum level of the DBH. In the emitter region there is obviously a spread in the injection energy. As a consequence, only some electrons take advantage of resonant tunnelling transfer of charge while the others are accumulated close to the heterointerface. Such an accumulation yields to the formation of an electrostatic potential. This electrostatic quantum well modifies significantly the charge transfer through the injection condition. In other words, the injection process involves transitions between to quantum systems. This issue will be an important part of the effect investigated in this thesis. On the other hand, with respect to the ideal picture of a pure coherent process, experiment shows dominant impact of disorder and scattering assisted tunnelling effects. However, for accounting for such spurious or parasitic effects which tends to smear out the relevant NDC phenomena are also of prime concern. This specific problem will concern an important part of the present work. At last, it seems interesting to generalise the simulation treatment to three terminal devices which appeasers a key component in digital applications. In that case the generic structure is the dual gate FET fabricated on nano-scale. With respect to the issues addressed previously we will see the necessity to solve the Boltzmann equation along with the Schroedinger equation. Some resonant tunnelling features will be demonstrated with the counterpart of a very low temperature operation.

In the first chapter briefly, we address the general issue of the NDC effect. We will discuss the various physical mechanisms responsible for negative differential conductance in the current voltage characteristics of a two terminal device. In this introductory chapter we discuss in short the conventional along with resonant tunnelling structures. To give an overall picture into the other types rather than tunnelling transfer effect, the reciprocal transfer as well as the real space transfer effects are outlined. The tunnelling effect through supper-lattices is also introduced

as an analogus to the reciprocal-space transfer one. To go further into the problem of realisation of high functionality devices, we will review some of the most promising options by discussing, on one hand, the possibility of direct access into the quantum well by an Omic or Shottky contact control and on the other hand, the co-integration of resonant tunnelling structure with active devices. This includes, Resonant Tunnelling Bipolar Transistor (HBRT), Resonant Hot Electron Transistor, Inverted base-collector tunnel transistor (Stark effect transistor) and Resonant Tunnelling Field Effect Transistor (RTFET). From the other hand, the Field induced tunnelling structures along with the solutions combine optical control gate are finally discussed.

In the second chapter, we investigate the effect of two-dimensional injection on purely coherent resonant tunnelling. We shall discuss the development of a new model which defines a two dimensional supply function. Concerning this issue, we will start from the notion of local density of states. Special attention will be paid to the wave function normalisation and some preliminary results will be also discussed by introducing the effect of non-parabolicity which influences the effective mass in the accumulation zone. Finally, to support the validity of the new formalism we will report a comparison between the simulation results carried out in the present study and the experiments.

In chapter three, we will address the effects of phonon scattering on the current-voltage characteristics. Our device under test will be an AlGaAs/GaAs resonant tunnelling diode fabricated and characterised in our group of research. We carried out a self consistent solution based on the fact that the broadening in the transmission probability is linked to the probability of scattering. The second issue is the inclusion of the phase breaking not only into the transmission probability but also into the 2D supply function. Besides, The interface roughness scattering will be addressed along with the impact of the space charge on the I-V curve. To validate the new formalism, we experimentally characterised the DBH diode and compare the experimental data with the results of the theory.

On the other hand, the current voltage characteristics can be tailored particularly at convenience. In this context, there is an important need for accurate simulation tools with the goal to computer aided design of these specific quantum structures. In a real device, the conduction mechanisms are much more complicated than the simple view of tunnelling electron through the quantum level of the DBH. In the emitter region there is obviously a spread in the injection energy. As a consequence, only some electrons take advantage of resonant tunnelling transfer of charge while the others are accumulated close to the heterointerface. Such an accumulation yields to the formation of an electrostatic potential. This electrostatic quantum well modifies significantly the charge transfer through the injection condition. In other words, the injection process involves transitions between to quantum systems. This issue will be an important part of the effect investigated in this thesis. On the other hand, with respect to the ideal picture of a pure coherent process, experiment shows dominant impact of disorder and scattering assisted tunnelling effects. However, for accounting for such spurious or parasitic effects which tends to smear out the relevant NDC phenomena are also of prime concern. This specific problem will concern an important part of the present work. At last, it seems interesting to generalise the simulation treatment to three terminal devices which appeasers a key component in digital applications. In that case the generic structure is the dual gate FET fabricated on nano-scale. With respect to the issues addressed previously we will see the necessity to solve the Boltzmann equation along with the Schroedinger equation. Some resonant tunnelling features will be demonstrated with the counterpart of a very low temperature operation.

In the first chapter briefly, we address the general issue of the NDC effect. We will discuss the various physical mechanisms responsible for negative differential conductance in the current voltage characteristics of a two terminal device. In this introductory chapter we discuss in short the conventional along with resonant tunnelling structures. To give an overall picture into the other types rather than tunnelling transfer effect, the reciprocal transfer as well as the real space transfer effects are outlined. The tunnelling effect through supper-lattices is also introduced

as an analogus to the reciprocal-space transfer one. To go further into the problem of realisation of high functionality devices, we will review some of the most promising options by discussing, on one hand, the possibility of direct access into the quantum well by an Omic or Shottky contact control and on the other hand, the co-integration of resonant tunnelling structure with active devices. This includes, Resonant Tunnelling Bipolar Transistor (HBRT), Resonant Hot Electron Transistor, Inverted base-collector tunnel transistor (Stark effect transistor) and Resonant Tunnelling Field Effect Transistor (RTFET). From the other hand, the Field induced tunnelling structures along with the solutions combine optical control gate are finally discussed.

In the second chapter, we investigate the effect of two-dimensional injection on purely coherent resonant tunnelling. We shall discuss the development of a new model which defines a two dimensional supply function. Concerning this issue, we will start from the notion of local density of states. Special attention will be paid to the wave function normalisation and some preliminary results will be also discussed by introducing the effect of non-parabolicity which influences the effective mass in the accumulation zone. Finally, to support the validity of the new formalism we will report a comparison between the simulation results carried out in the present study and the experiments.

In chapter three, we will address the effects of phonon scattering on the current-voltage characteristics. Our device under test will be an AlGaAs/GaAs resonant tunnelling diode fabricated and characterised in our group of research. We carried out a self consistent solution based on the fact that the broadening in the transmission probability is linked to the probability of scattering. The second issue is the inclusion of the phase breaking not only into the transmission probability but also into the 2D supply function. Besides, The interface roughness scattering will be addressed along with the impact of the space charge on the I-V curve. To validate the new formalism, we experimentally characterised the DBH diode and compare the experimental data with the results of the theory.

The lateral structures resulting in an electrostatic induced potential whose space variation depends at least on two directions will be addressed in chapter four. Before investigating the tunnelling transport in these lateral structures, a special attention will be paid to find out the most realistic potential function, the relevant Fermi-level and the injection-mode.

The analysis will be conducted over by considering: firstly, the perfect adaptation of a full two-dimensional hydrodynamic energy model to simulation purposes at very low temperatures as well as in the ambient conditions. The second step, will be devoted onto a novel physical picture including both the lateral and the vertical confinement effects into the transmission along with the injection conditions. On this basis we introduce the necessary modifications to simulate the lateral resonant-tunnelling dual gate MODFET at liquid Helium. This concerns the validity of shallow Ohmic contact for such a structure along with a novel approach to model the deep Ohmic contact. Also the degeneracy correction into the hydrodynamic energy model will be extensively investigate. Finally, we will address the lateral tunnelling process under various assumptions.

CHAPTER 1

1. EFFECTS OF NEGATIVE DIFFERENTIAL CONDUCTANCE AND TRANSCONDUCTANCE

1.1 Negative differential conductance and high functionality :

Negative differential conductance (NDC) which results from a resonant tunnelling effect gained the attention of many functional component designers not only as it is an extremely high speed phenomenon but also as it allows them to integrate complex functions into a single functional device. For example, frequency multiplier and four bit parity generator have already been elaborated using a single device [1], inverters, OR gates and flip-flops have been fabricated using double-barrier GaAs/AlGaAs resonant tunnelling diodes [2]. It is to be added that one can implement a single or a multiple NDC-region device or circuit in order to realise a multi-functional device, NOR [3].

1.2 Physical effects based on localisation and tunnelling phenomena :

1.2.1 Conventional tunnelling structures :

The first acknowledged paper on tunnel devices discussed the tunnel diode, also referred to as Esaki diode, and was written [4][5] by L.Esaki in 1958. Basically the conventional tunnelling structures depend on the band to band tunnelling phenomena. Figure 1.1 shows a schematic energy band diagram of a conventional tunnel structure in thermal equilibrium. Because of the high doping concentration the Fermi-level is located within the allowed bands. The degree of degeneracy denoted by V_p and V_n, is typically a few $k\theta$ where k is Boltzmann' constant and θ is the absolute temperature. The depletion layer width is of the order of 100 Å or less, which is relatively narrower than the conventional p-n junction.

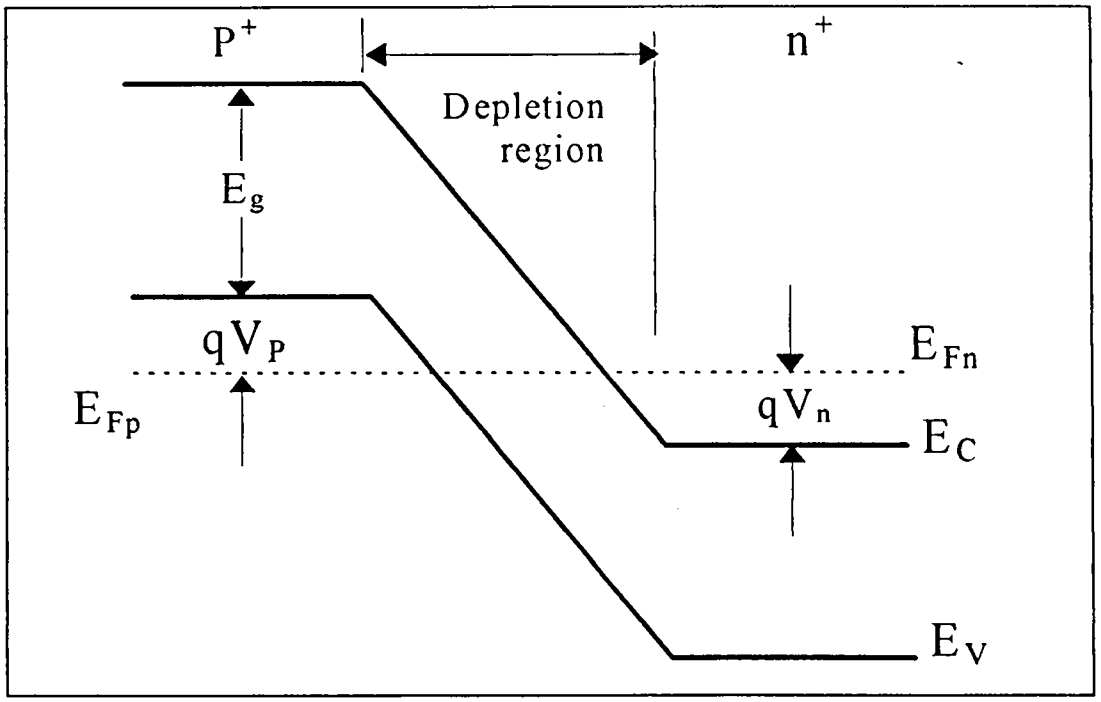


Figure 1.1 Energy-band diagram of a conventional tunnel diode in thermal equilibrium.

Figure 1.2a shows a typical static current-voltage characteristic of a tunnel diode. In the forward bias direction, the current first increases to a maximum value (peak current or I_p) at a voltage V_p, then decreases to a minimum value I_v at a voltage V_v.. For a voltage higher than V_v, the current increases exponentially with the voltage. The static characteristic is the result of three current components illustrated in

Figure 1.2b, band- to-band tunnelling current, excess current, and thermal current respectively.

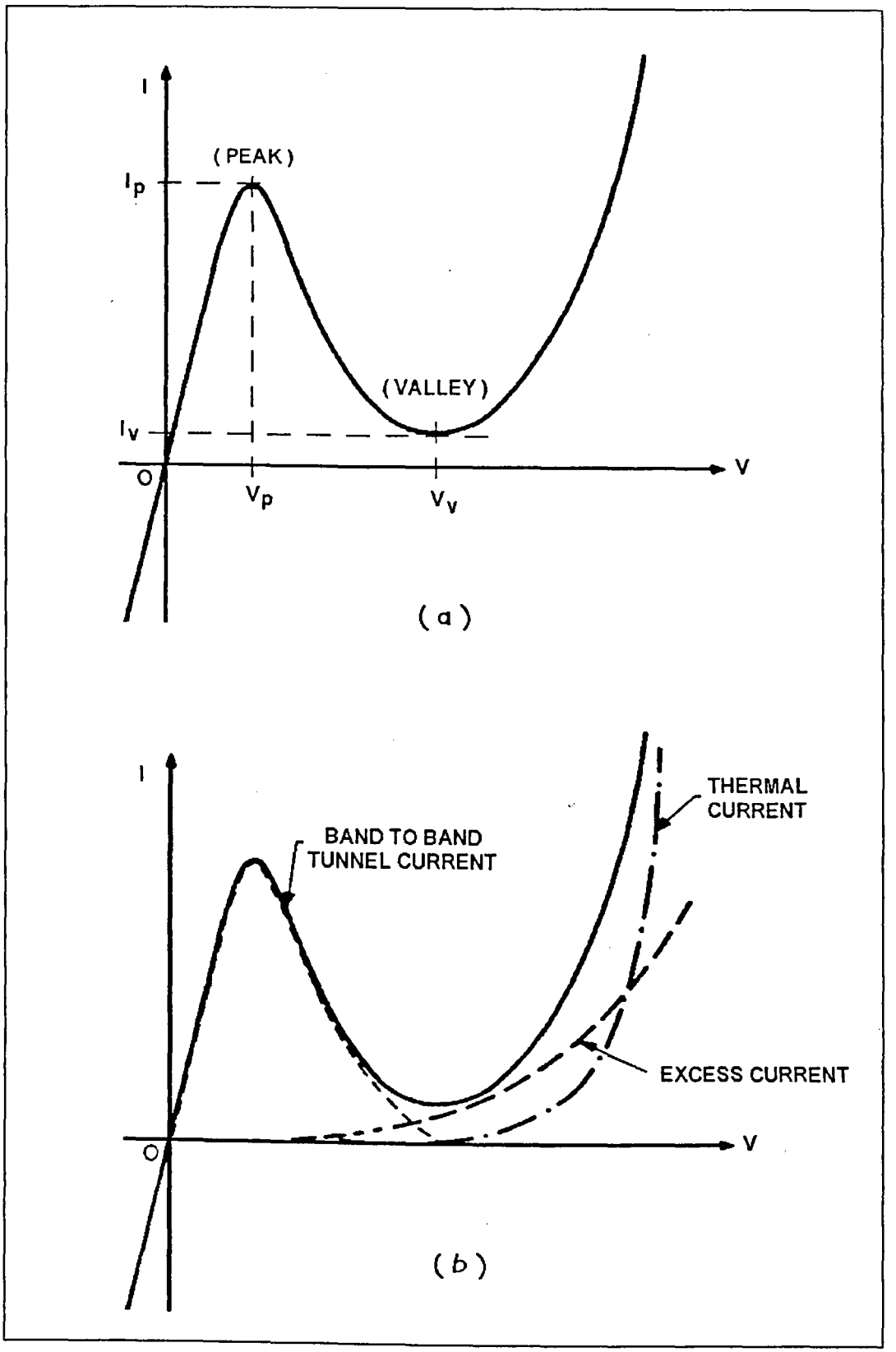


Figure 1.2 (a) Static voltage-current characteristics of a typical tunnel diode, (b) the static characteristics is analysed into three current components.

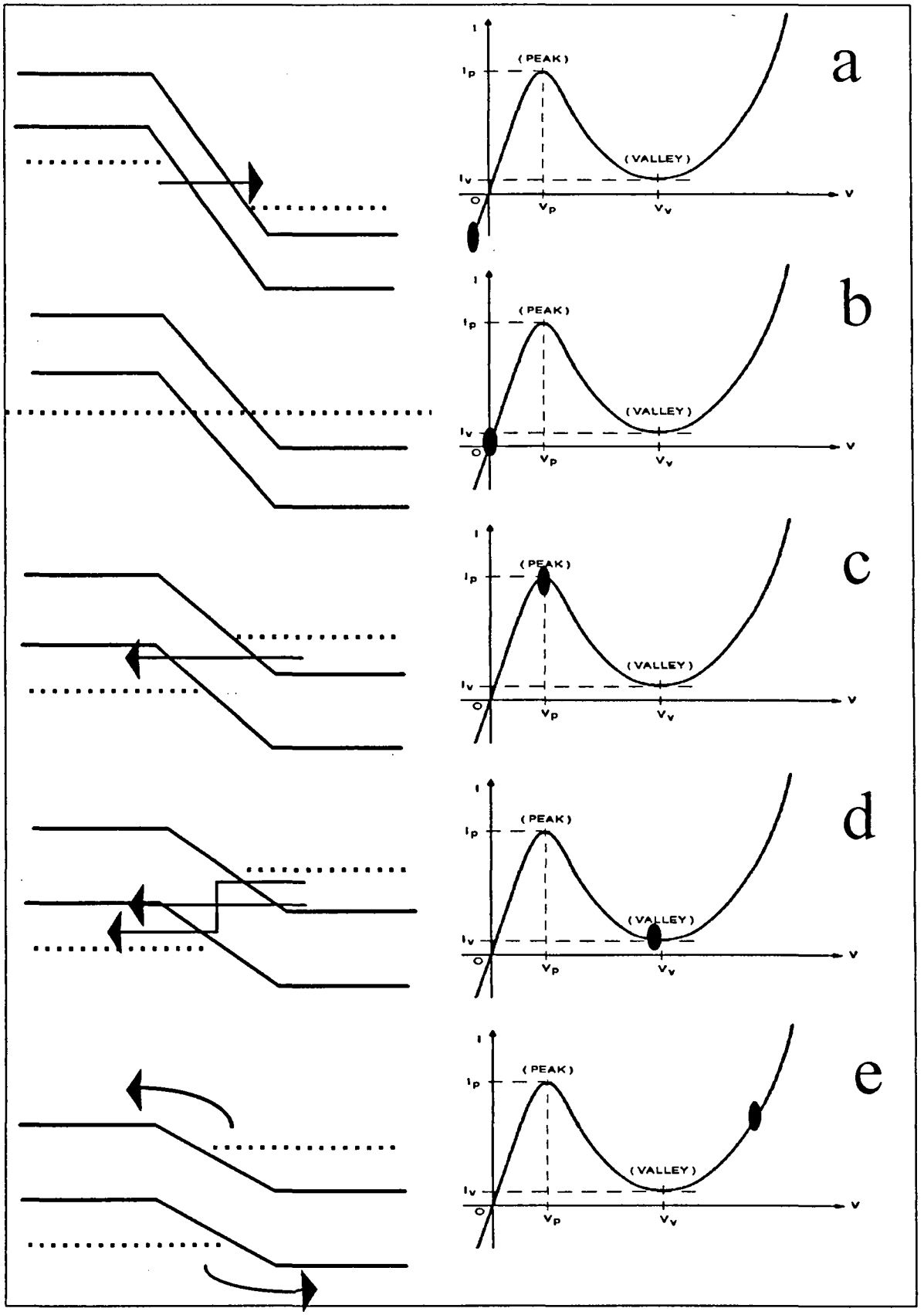


Figure 1.3 Simplified energy-band diagram of conventional tunnel diode under deferent biasing conditions.

We first discuss qualitatively the tunnelling processes at low temperature using the simplified band structures shown in

Figure 1.3. Note that the Fermi levels are within the conduction and valence band of the semiconductor. Under equilibrium condition (

Figure 1.3b) the Fermi level is constant across the junction. Above the Fermi level there are no filled states on either side of the junction, and below the Fermi level there are no empty states available on either side of the junction. Hence tunnelling currents cannot flow at zero applied voltage. When a biasing voltage is applied, the electrons may tunnel from the valence band to the conduction band, or vice versa. The necessary conditions for tunnelling are :

1. occupied energy states exist on the side from which the electron tunnels.
2. unoccupied energy states exist at the same energy levels on the side to which the electron can tunnel.
3. the tunneling potential barrier height is low and the barrier width is narrow enough so that there is a finite tunneling probability.
4. the momentum is conserved in the tunneling process.

Figure 1.3a shows electron tunneling from the valence band into the conduction band when a reversed bias is applied. The corresponding bias point is marked with a dot on the I-V curve. When a forward bias is applied (

Figure 1.3c) there exsits a band of energies for which there are filled states on the n side facing to unocupied states on the p side hence the electrons can tunnel from the n side to the p side. When the forward voltage is further increased, there are fewer available unoccupied states on the p side (

Figure 1.3d). If a voltage is applied in such a way that the bottom of the conduction band is exactly opposite to the top of the valence band, the tunneling conditions are no more satisfied. Thus at this point the tunneling current can no longer flow. With still a further increase in the voltage, the well-known thermal current flows (

Figure 1.3e), and drastically will increase with the applied voltage. Tunneling routes are shown in this figure by arrows.

In the present work we did not study the Esaki-diodes. Nevertheless in order to mack a comparison with resonant tunnelling devices which will be considered in the next section, Figure 1.4 showes the current-voltage characteristic measured at room temperature in our laboratory (IEMN) for a typical device.

In agreement with the qualitative approach ilustrated above, the device (comersially under the referebce 1N3857) exhibites a pronaunced NDC effect over a quite wide bias voltage range (in excess of 300 mV). The shoulder-type behaviour in the I-V curve is a result of the well known spurious selfe biasing effect. We will discuss of such parasitic effect in chapter 2.

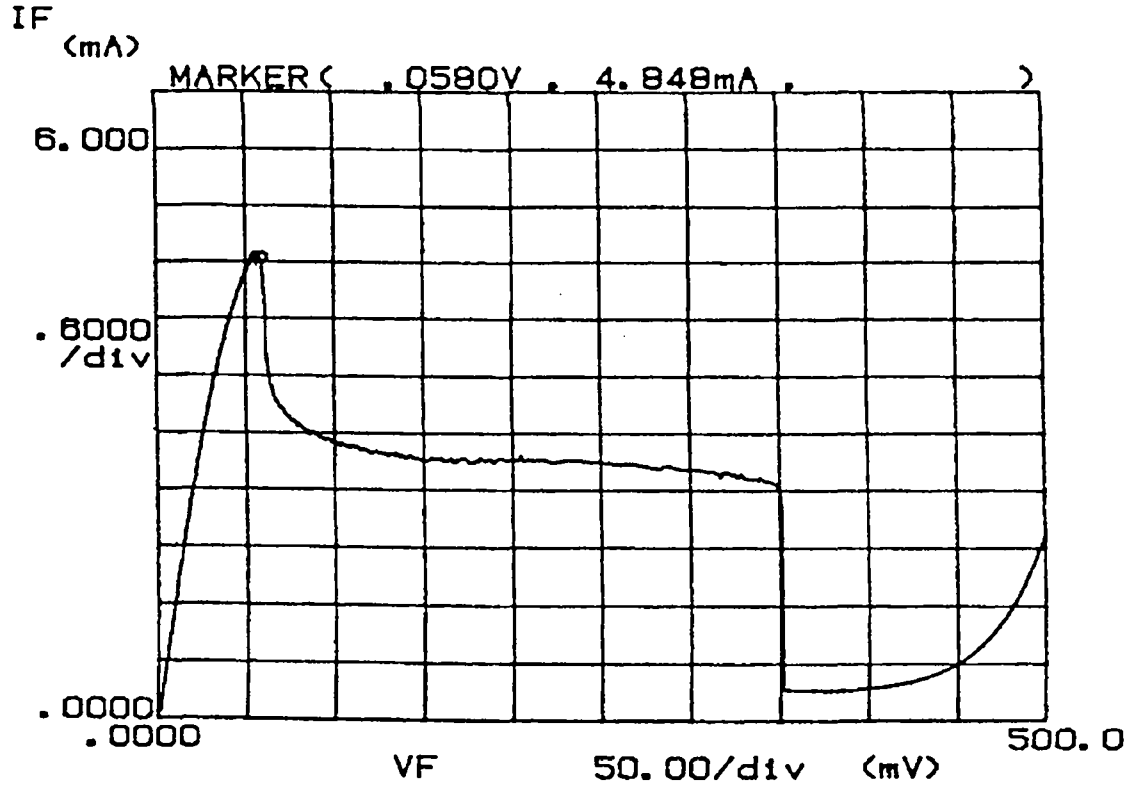


Figure 1.4 room temperature measurements of I-V characteristic of the Esaki-diode comercially referenced under the technical number 1N3857.

In contrary, the peak volyage is quit low (50 mV) for the device under test. The key figure of merit in this cas is clearlly the huge Peak-to Vally current ratio (16 :1) wich is a result of the interband tunneling mechanism. This is an important advantage and now there is an increas effort devoted to Antimonide based material systems involving such interband tunneling process.

In counter part, conventional Esaki-diod suffers from limited current capability alongwith very high capacitance level. This is explained by the very high doping leveles cloused to the tunnelling barrier which results in high RC time constant.

1.2.2 Resonant tunnelling structures :

The modern epitaxial techniques namely MOCVD and MBE, in the present time, make it available to build potential barriers (classically forbidden bands of energies)

using material modulation within the range of a few monatomic layers. The charge carriers (electrons or holes), near the corresponding band edge of the injection zone, have the ability to tunnel through the barrier. If for example, the structure contains only a single barrier, the quantum transmission probability is in the same order of magnitude of that of the conventional tunnel diode which is typically 1%.

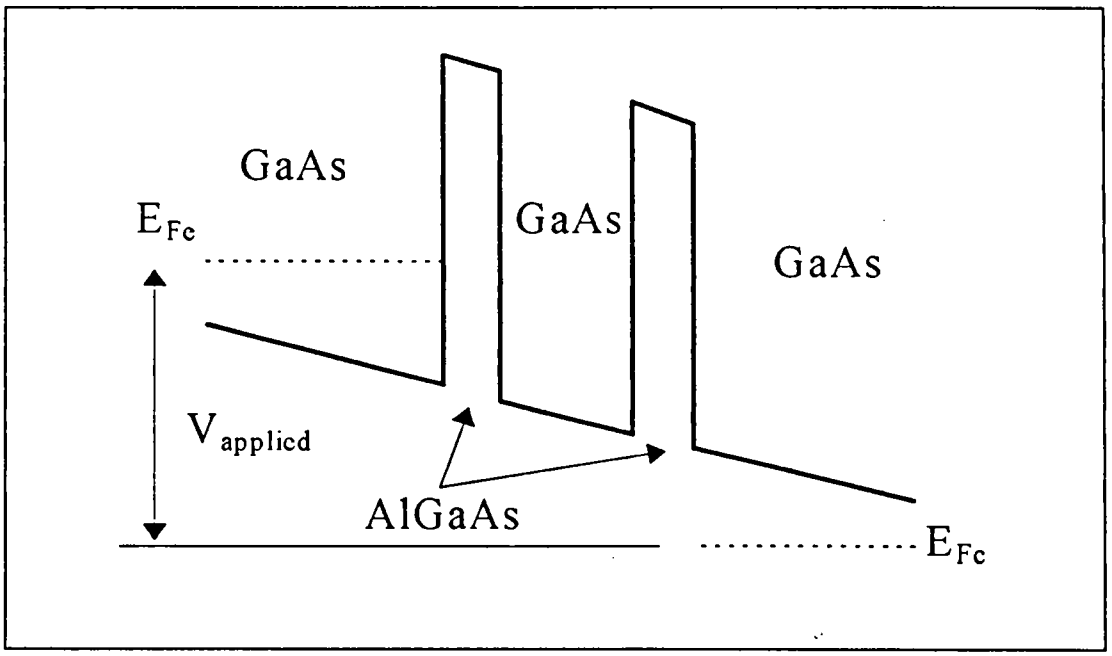


Figure 1.5 schematic diagram for simple resonant tunnelling structure.

If now one creates a potential well (classically allowed bands of energies confined between two potential barriers, see also Figure 1.5, one observes a resonant behaviour in the transmission probability. This is a direct result of the quantisation of electron energy in epitaxial growth direction (longitudinal direction). An example of the transmission probability calculated as a function of longitudinal energy is illustrated in Figure 1.6. One observes the selective unity transmission probability at the longitudinal resonant energy corresponding to the quantum well level. The transmission probability out of these resonant levels takes, approximately, the value of the transmission probability of a single barrier with a thickness equivalent to the sum of those of the two barriers which constitutes the resonant structure.

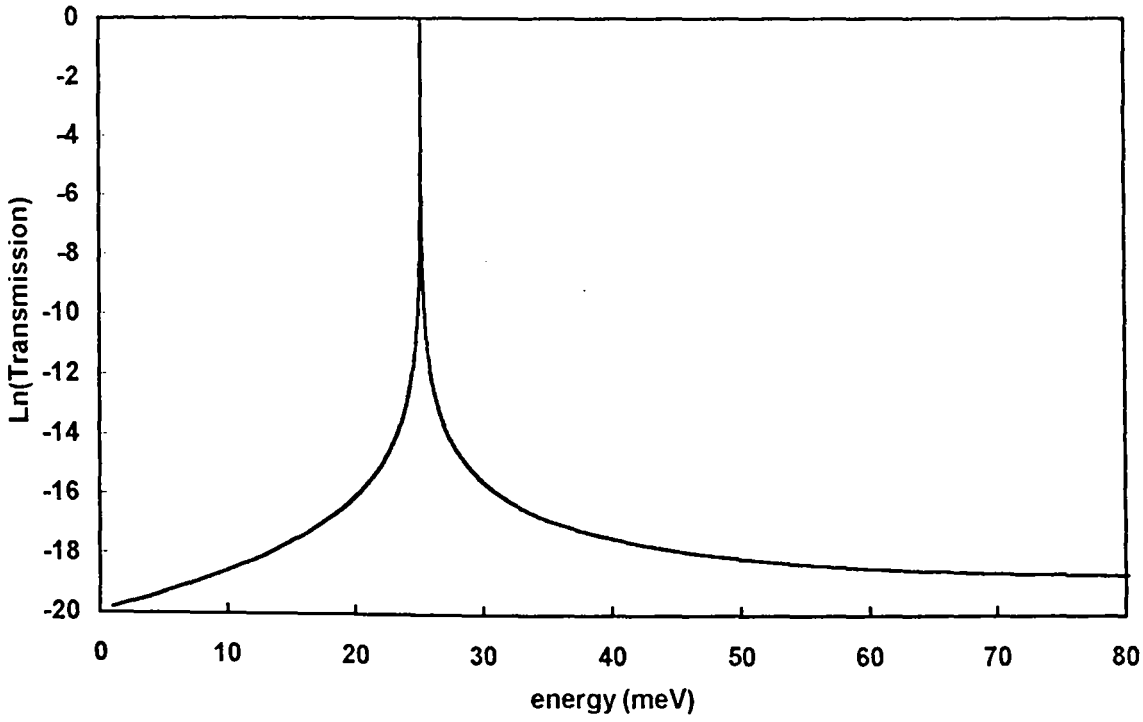


Figure 1.6 typical transmission coefficient.

The understanding of the resonant tunnelling phenomena might be achieved in the scope of the Fabry-Pérot optical interferometer model. It consists of two infinite parallel semi-reflecting mirrors with transmission coefficients denoted by t_l and t_r for the left and right mirrors respectively while they are separated by a distance L. Assuming a monochromatic wave of a Transverse Electric and Magnetic (TEM) mode is launched perpendicularly on the left mirror for instance, one can easily deduce the total power transmission coefficient to be as follows,

$$T = \frac{t_l t_r e^{-jkL}}{1 - e^{-2jkL} r_l r_r} \qquad \textit{Eq. 1-1}$$

In this equation r_l and r_r are the réflexion coefficient of the left and right-hand sides defined as :

$$r_l = \sqrt{1 - t_l^2}\ , \quad r_r = \sqrt{1 - t_r^2} \quad \text{and} \quad k = \frac{2\pi}{\lambda}\ . \qquad \textit{Eq. 1-2}$$

If the cavity length L is an integer multiples of $\lambda/2$, the overall transmission coefficient (T) in Eq. 1-1 will be unity. This corresponds to in-resonance condition. In contrast, for wave lengths for which the factor $e^{-2jkL} = -1$, the transmission $T \cong t_l t_r/2$, which is the transmission of an equivalent single barrier with two identical mirrors in series. In this case we say that the cavity is out-of-resonance. It is to be mentioned that the reflection coefficient of the semi-reflecting mirror equally with its resulting phase shift θ are independent of the incident light frequency. In the case of epitaxial potential barrier there is a phase matching condition which will be,

$$kL + \theta_l + \theta_r = (n-1)\pi \qquad \textit{Eq. 1-3}$$

where θ_l, θ_r (the phase shifts within the barriers) and k will be dependent of the electron energy equivalent to the electromagnetic wave frequency.

This intra-band resonant phenomena (in the sense that the conduction mechanism is unipolar) gives rise to NDC effect. Special attention will be paid to underlying mechanisms involved in a real device. In short, however, let us mention that there exists a strong difference between the conduction characteristics of Esaki diode and resonant tunnelling structures marked by the existence of high transmission level at resonance which leads to a high current capability. It will be shown that the voltage range over which the differential conductance is negative, is a consequence of complex mechanisms notably two-dimensional injection processes. On the other hand current contrast between the in and out of resonant conditions will be effected by scattering assisted tunnelling phenomena. As a result the Peak-to-Valley current ratio is drastically decreased.

1.3 Reciprocal space transfer effect :

For this very short and very brief report on charge transfer in K-space, we will distinguish between the bulk effect and the operation of a real device.

Reciprocal space transfer effect also named Gunn effect is an intrinsic property. The materials exhibiting this effect should have a conduction band characterised by a primary minimum located at the Γ point in the Brillouin zone (central valley) and a number (at least one) of satellite valleys at relatively higher energies with respect to the bottom of the central valley (Figure 1.7). Under the influence of high electric fields electrons gain enough energy compared to the difference between the central and satellite energy minima. Subject to different scattering mechanisms, some of these electrons could have an average of k value comparable with the k value difference between the two valleys. Under this condition these electrons could be transferred from the central valley to the satellite one. If the effective mass of the satellite valley is relatively higher than that of the central one (which is the case of III-IV semiconductor materials), the average electron velocity will be decreased and the material will exhibit a negative differential mobility region (Figure 1.8).

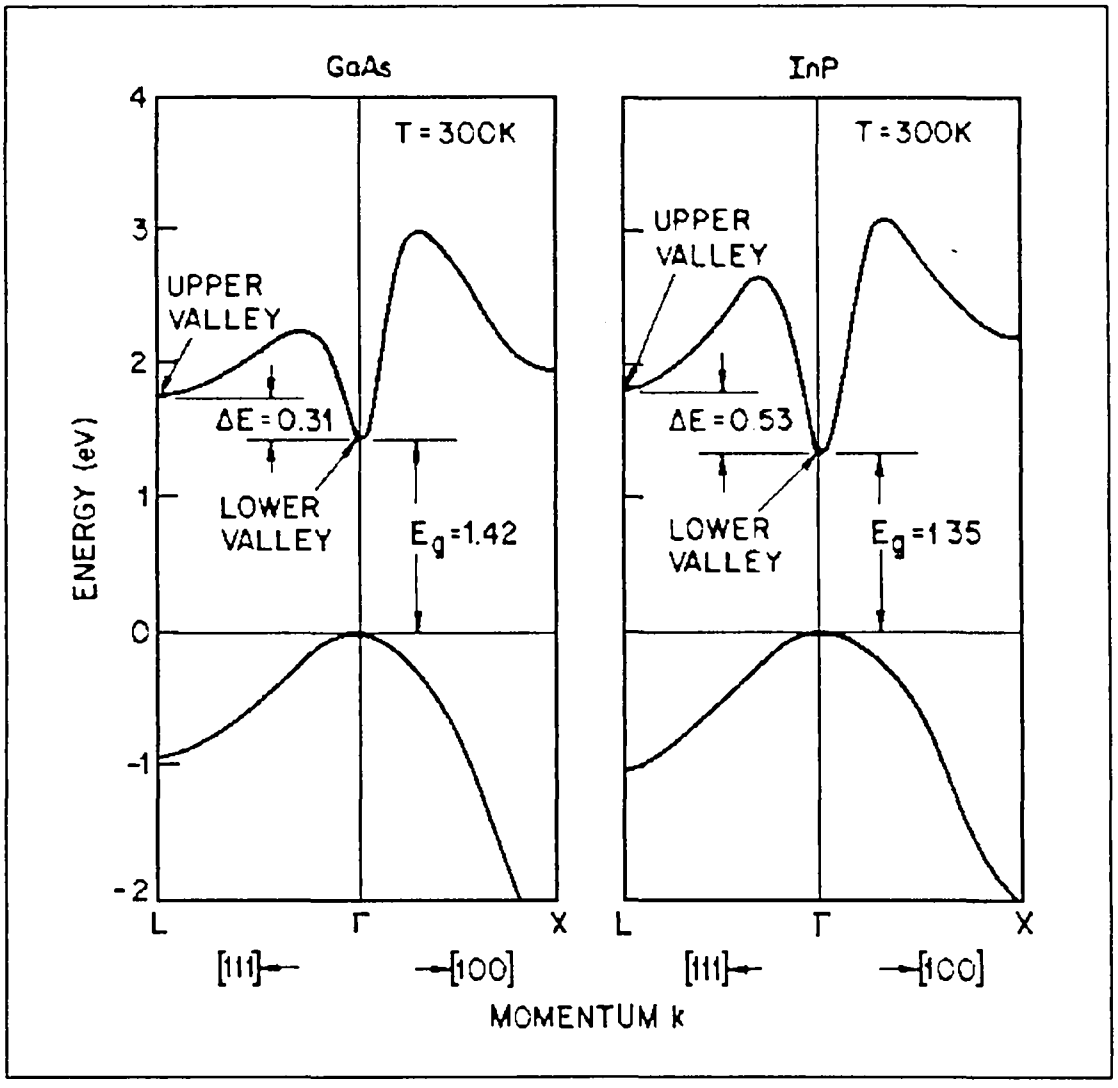


Figure 1.7 band structure diagram of GaAs

For the illustration purposes we plotted in Figure 1.8 the variation of average electron velocity versus the electric field for the turnery alloy $Ga_{0.7}Al_{0.3}As$ at room temperature, calculated by Mont-Carlo simulations technique (performed at IEMN). Also we assumed that the semi-conductor material is undoped.

In chapter(4) we will use such data as an input to the hydrodynamic model (SIMFET) originally implemented by [6] and extensively developed by [7][8][9], [7]this is after the necessary modifications which will be discussed later on in chapter(4).

From Figure 1.8 one can notice that the threshold critical electric field is around 5KV/cm. The peak velocity reaches 1.2 10^7 cm/s whereas the saturation velocity is close to 0.8 10^7 cm/s. On this bases one can expect that such negative differential mobility gives rise to an NDC effect since the current density depends directly on the velocity and the electric field is a direct result of the applied voltage.

However in real devices the condition of a uniform electric field is often not satisfied. In this case the space charge effect plays a key role that will be discussed in the following section.

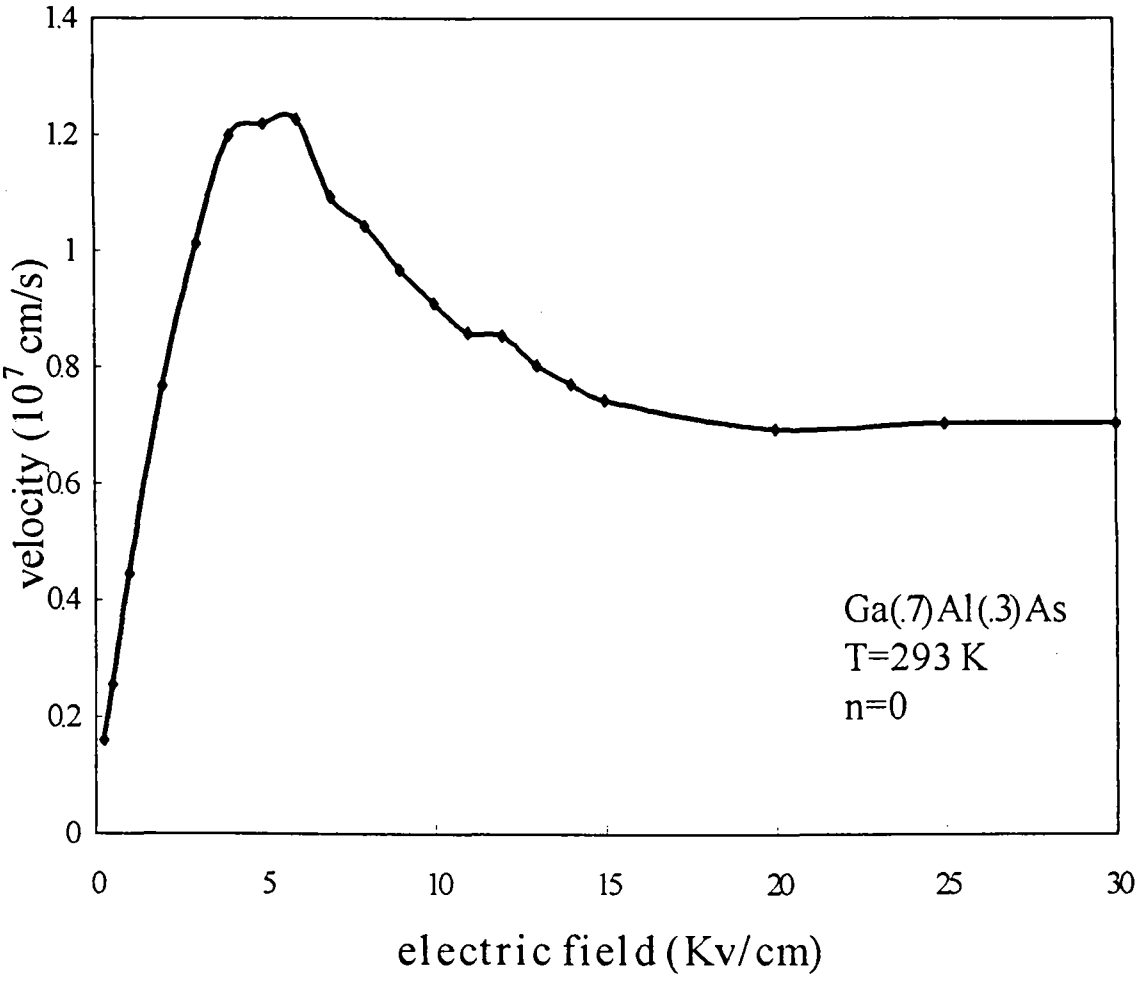


Figure 1.8 Mont-Carlo simulations of velocity-field characteristics of GaAlAs performed in IEMN

1.3.1 reciprocal space transfer device :

To get inside the operation of a real device let us consider Figure 1.9 taken from [10] where are schematically depicts the device active region, the electron density, the internal field as a function of distance along with the current density as a function of electric field.

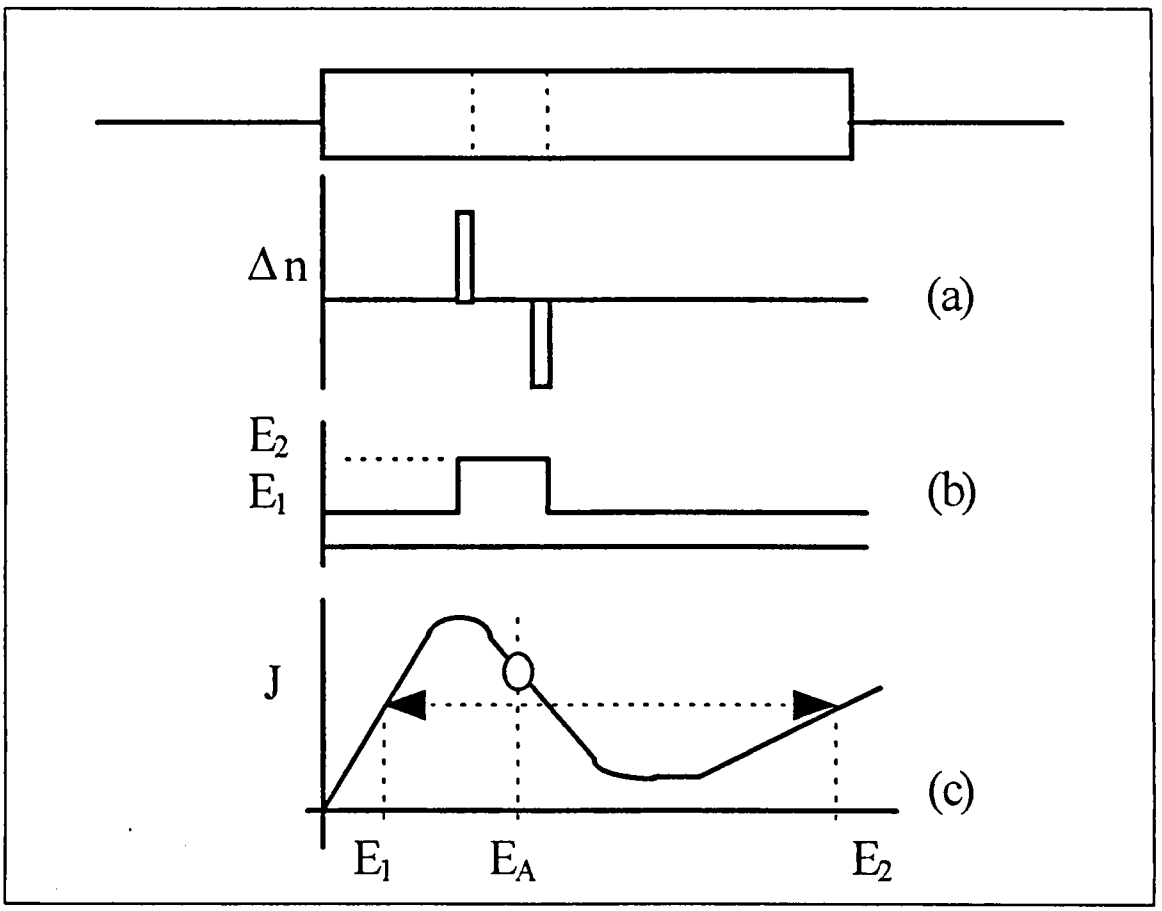


Figure 1.9 formation of an electron dipole layer in a perturbed medium of negative resistively, after Kroemer [10]

If the semiconductor material is biased such that the applied electric field (E_A) is in the negative differential mobility region, the material, initially electrically homogeneous, becomes heterogeneous. The origin of this heterogeneity might be any field perturbation. For example it might be at the neighbourhood of the

electrodes, where a region of higher resistance followed directly by a region of lower resistance thus formed, and hence a formation of accumulation-depletion region (high field domain) is started (Figure 1.9a) and would be developed in an attempt to reach current stability condition. The field inside the dipole domain (E_2) would be greater than the fields on each side of it (E_1) (Figure 1.9b). The two field values will tend toward equilibrium levels outside the NDR region where the high and low currents are equal (Figure 1.9c). The dipole layer moves through the active region and discharges at the anode. At this instant the field returns to the original uniform distribution, while a new dipole layer is starting to be formed, and the process repeats itself.

It is to be mentioned that the early-stage space-charge growth is given by $(n-n_o)=(n-n_o)_{t=0}\exp(t/\tau_R)$ and hence the maximum growth factor would be $\exp(T/\tau_R)$ where T is the sample transit time $T \approx L/v_s$, L is the sample length, v_s is the saturation velocity and τ_R is the dielectric relaxation time. For large space charge growth (domain mode of operation) this growth factor must be grater than unity thus making $n_o L \geq \epsilon_s v_s/q\mu^- \approx 10^{12} cm^{-2}$ for GaAs and InP.

There is another mode of operation, called limited space charge accumulation (LSA) mode, in which we drive the whole sample into the NDR region only for a very short period of time during which no significant charge accumulation can occur ; and hence the whole sample length would be in contribution to the device negative resistance.

In conclusion we learned that the Gun structures exhibit a transit time limitation analogous with any injection and transit time devices IMPTT diode for example. As a general rule the saturation velocity which is relevant for this transit process is relatively lower than the peak velocity. This imposes a limitation on the switching time of these devices. On the contrary the resonant tunnelling diodes exhibit a hot injection effect so that much higher velocity is expected in the adjacent transit layer.

1.4 Supperlattices :

In this section we will give a brief report on the mean conduction characteristics of semiconductor supperlattices.

Let us recall that the achievement of supper-lattice like electron transport requires the fabrication of tightly coupled quantum wells. Under this condition the overlap between the envelop wave function leads to the de-localisation of spatial electron probability distribution.

For infinite supper-lattices, energy mini-bands are created house energy positions and broadening are directly related, to the transparency of the barriers which depends on the geometrical parameters and the semiconductor materials.

The transport properties under bias are relatively complicated and several effects were proposed to explain the occurrence of Negative Differential Mobility. For instance it can be expected that the application of electric field leads to a transition between quasi-mini-band transport characteristics (wave function de-localisation) to (Wannier-Stark) ladder characteristics (wave function localisation). This leads in practice to NDC effect observable in the I-V relationship measured under D.C conditions.

However as for conventional Gunn-diodes, real devices shows a transit mode limitation with notably the formation of a dipole domain travelling through the active region. Therefore, from the device point of view the operation of supper-lattices used for instance as oscillators is quite comparable with that of Gun-diodes but with an effective drift velocity dependent on the min-band broadening. More over because the origin of the negative differential mobility is not the same, the frequency limitations are different. It is now well known that the relevant time in the k-space transfer devices is the inter-valley scattering time. In contrast, as seen above, the characteristic time of supper-lattices involves the transition probability between eigen states.

Concerning the shape of the I-V relation, the onset of the characteristic is typically abrupt where as the decrees in the current value in the NDC region is relatively gradual. On the other hand due to the spurious self oscillations one observes typically self-biasing effect which distorts the current variation versus voltage.

1.5 Real space transfer effects :

In the previous sections we have briefly discussed the underling mechanisms of transfer in k-space. Such transfer which leads to a negative deferential conductance effect is often used in the fabrication of transfer electron diode (TED). Recently it was proposed that NDC effect can also be achieved in three terminal heterostructures tacking the advantage of transfer in the geometrical space from a high mobility towards a low mobility region.

It is now well known that the mobility of a semiconductor is directly related to the band-gap of the semiconductor material. As an example, let us recall that the narrow gap InAs material exhibits one of the fastest drift velocities whereas AlSb material with wide band-gap shows a poor performance in terms of velocity. In addition most of the heterostructur transistors have doping modulation. Under this condition one can expect a significant decrees in the drift velocity at increasing voltages.

On these bases all conditions are met to realise a drift velocity modulation which is provided by the electron transfer from a narrow gap undoped material towards a wide gap highly doped material.

This real space transfer effect necessitates [11][12][13] that the electrons have to jump over the heterostructure barrier. The required excess energy is provided by the heating effect which is induced by the electric field. In this mechanism the charged carriers are conserved within the main current path in the active region. Another possibility consists in inducing current carriers modulation by collecting some of or approximately all the electrons from the main current path. There exists several

configurations of such devices, all are based on the collection of a current component using an extra electrode. For instance let us mention the work published by Luryi. In his work [14] he employs the substrate as an extra collector gate and call it NEgativ .Resistance Field Effect Transistor (NERFET). Figure 1.10 sketches the device proposed in [14].

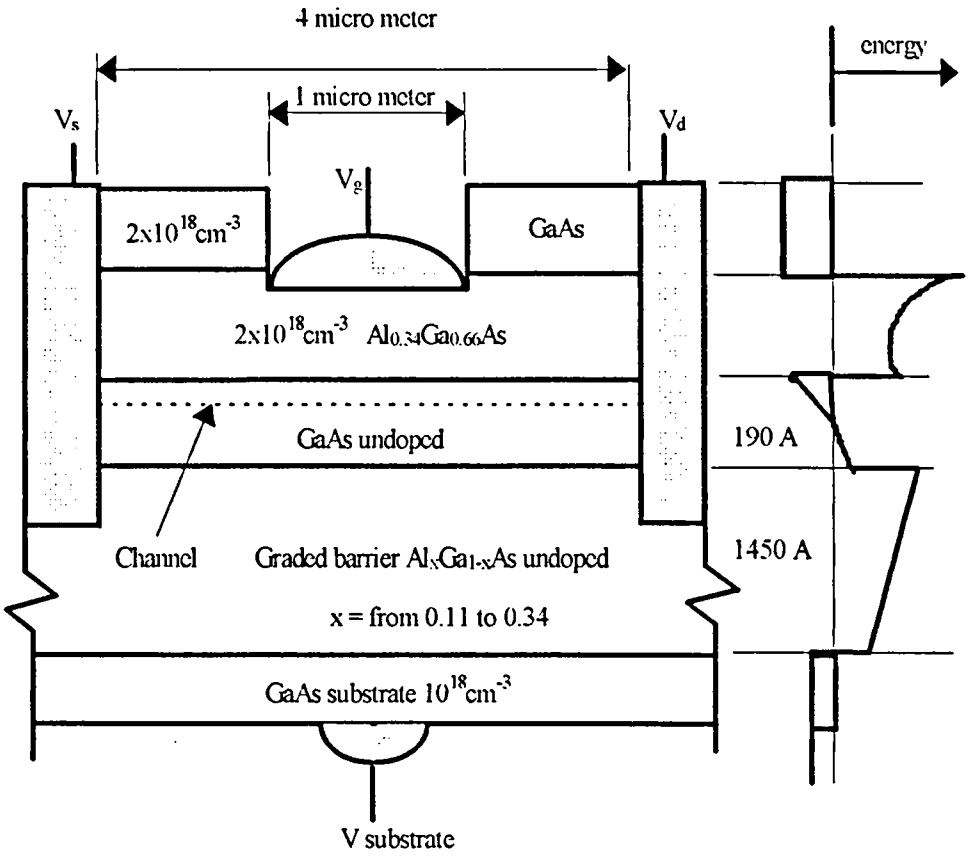


Figure 1.10 NERFET structure and energy band diagram after [14].

On the hand one can fined in the literature the same effect but using planer structures. Concerning this class of devices, one of the most elegant ways for achieving high performance is to use a Dual-Gate Field Effect Transistor (DGFET). The advantage of DGFET as compared with conventional FET have been investigated over the past [15][16] in connection with impedance level at the control terminals and the stability effect.

From the point of view of real space transfer effect such dual gate scheme is also favourable because varying the various voltages applied on the gates permits the optimisation of the heating effect necessarily required to observe electron transport in the high ban gap material. In addition it also permits the optimisation of carrier collection effect depending on the mode of operation.

In order to exemplify these degrees of freedom in device operation, let us consider the current flow lines and the electronic total kinetic average energy distribution depicted in Figure 1.11a and Figure 1.11b respectively. This figure clarify the formation of a high energy domain at the exit of each gate with the subsequent electron transfer into the wide band-gap highly doped $Al_{0.3}Ga_{0.7}As$ capping the GaAs undoped narrow band-gap material.

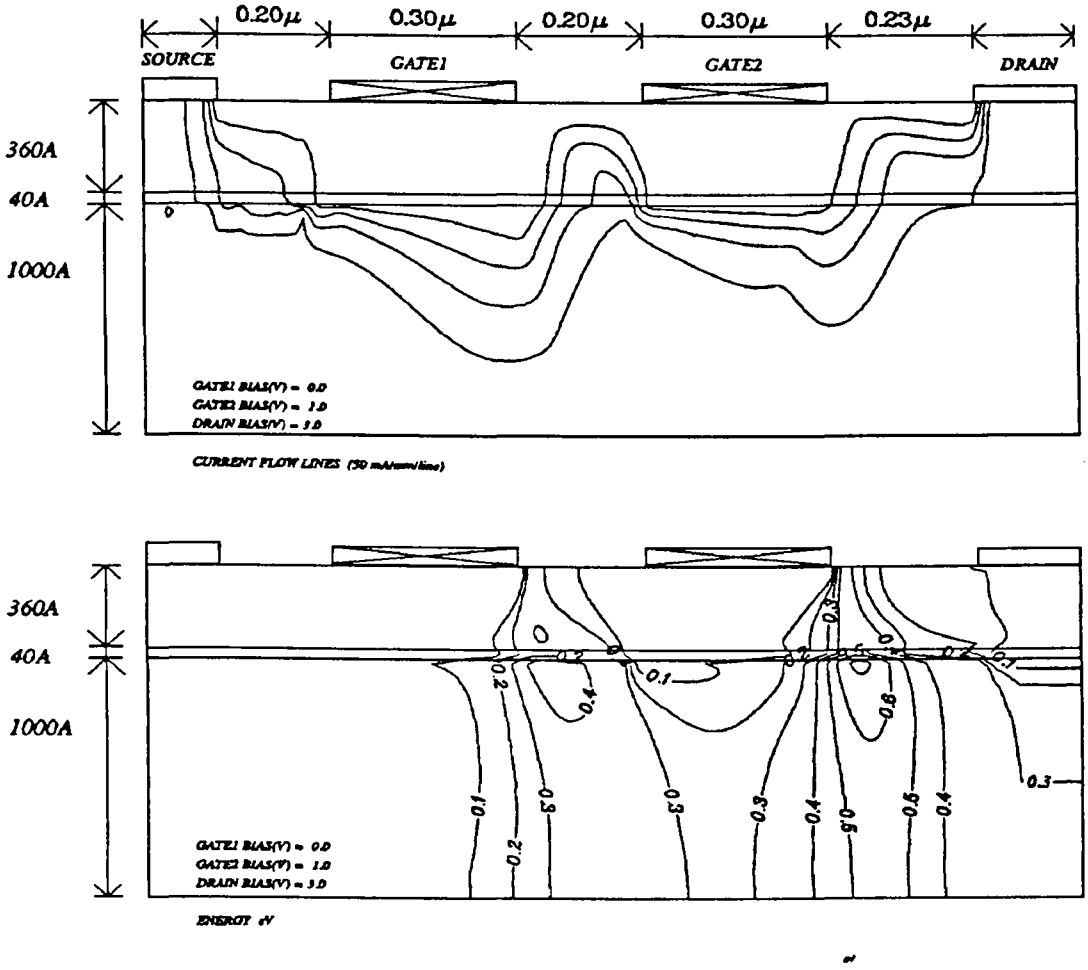


Figure 1.11 current lines (a) and energy distribution (b) [15],[16].

Also Thomas et all [17][18] have proposed a multiple-layer 2μm dual gate 2μm inter-gate separation field effect real space transfer transistor in which they use the second gate as a collector to modulate the drain current and hence creating a negative differential trans-conductance between the drain current and the second gate voltage. For the application of the real space transfer phenomena into two terminal devices See also [19].

1.6 Possible solutions for structures of transistor type :

Two terminal resonant devices with NDC characteristics have several valuable properties for circuit design (high speed, integrated-capability). Three terminal devices with controllable NDC characteristics are also attractive for circuit design. This is because they provide isolation between different circuit stages and allow a better control of device operation. In the same time NDC property in three terminal configuration increases the functionality of the device.

In order to fabricate a tunnelling device in a three-terminal configuration, there exists two main possible solutions. The first one involves creating a direct accessibility to control the resonant quantum well behaviour by either an Ohmic contact or a Schottky contact. The second one is the co-integration between a resonant tunnelling diode and a conventional transistor structure. In the beginning we shall discuss the first solution and then after the second one.

1.6.1 Direct accessibility to the quantum well :

1.6.1.1 Ohmic contact control :

In this technique [20][21] (Figure 1.12) a shallow Ohmic contact is directly deposited onto the quantum well (base layer) to enable a control of the value of potential in the quantum well (this technique is also known as potential control method). By this means the tunnelling current which may cross the base layer can be modulated [22][23]. The limitation of this kind of structure is the relatively high base access resistance resulting from the reduced thickness of the base active layer.

In fact, it is now well known that the pronounced quantum effects at room temperature implies the fabrication of narrow quantum well or/and the use of a very low effective mass well material, with high conduction band offsets. On the contrary the devices whose operation involves the deposition of a base Ohmic contact such as Heterostructure Bipolar Transistor should have a doping concentration level which has to be as high as possible along with a sizeable base.

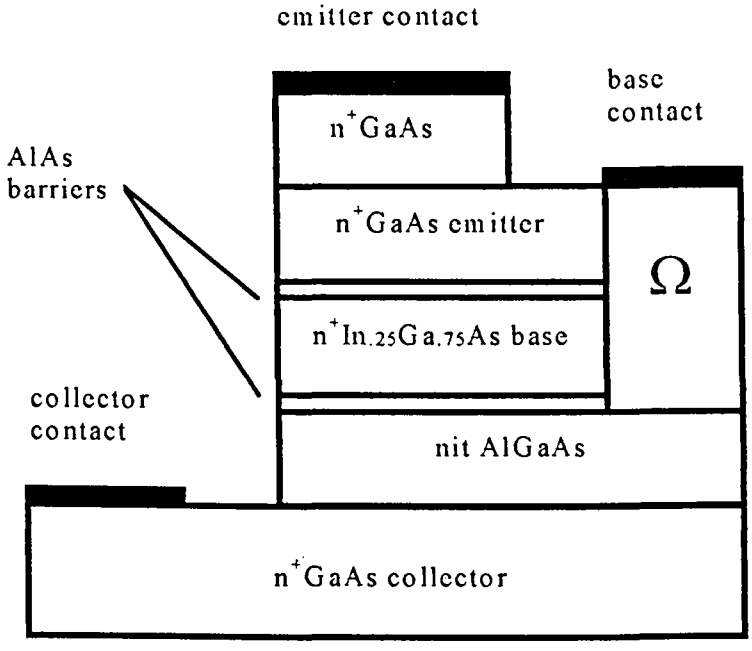


Figure 1.12 Schematic diagram showing a typical direct ohmic contact into the base layer (after Haddad et all. [20])

To overcome such trade-off between the observation of quantum effects and the achievement of a high accessibility to the base layer, one finds in the literature several solutions. From the material point of view, some authors proposed to take the advantage of the very low effective mass exhibited by InAs material and use it in the base layer in association with AlSb barriers. Although such a solution is a priori interesting, it suffers however many difficulties notably the fabrication of high quality epi-layers. On the other hand, it was proposed over the past to implement a buried base layer. Figure 1.13 depicts a schematic diagram illustrating the buried base layer band structure.

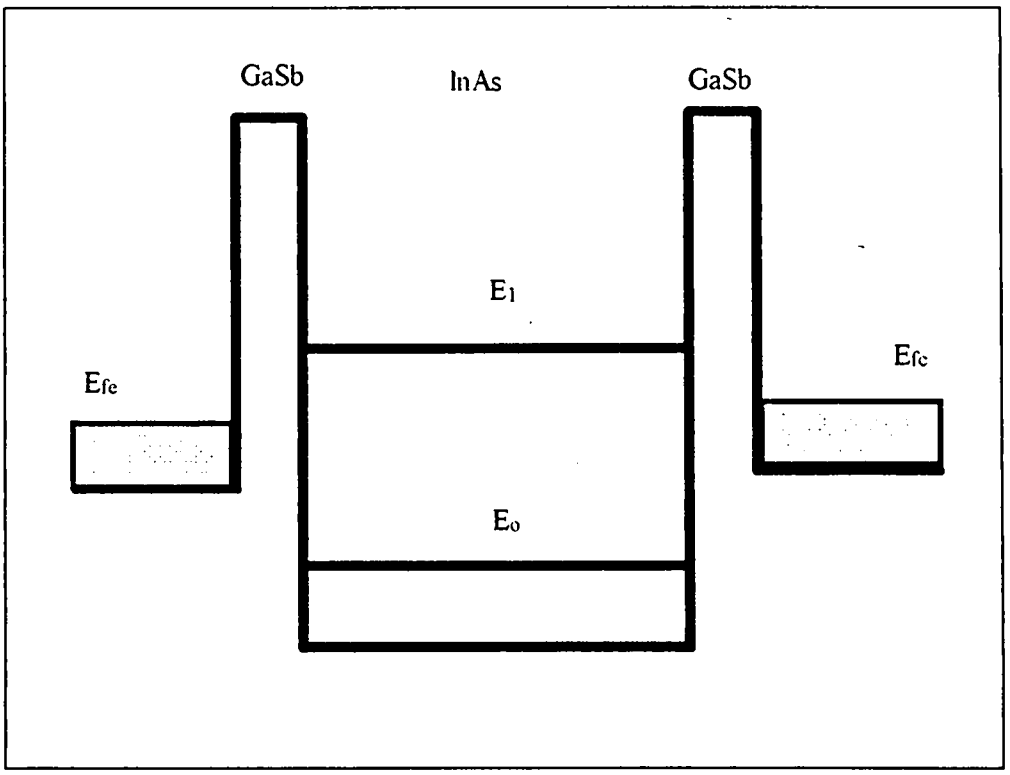


Figure 1.13 schematic diagram of buried base conduction-band structure.

By using a deep quantum well accommodating a ground level below the conduction band edge of the adjacent layer, it is possible to separate the electron population responsible of the emitter-collector current from that which plays the role of a reservoir close to the base contact. Unfortunately the tunnelling electrons which are injected through the first excited state experience many inter-band scatterings which gives rise to unreasonable base leakage current [24].

As a last example, let us mention the work of NTT group [25] which introduces the idea of coupled quantum wells as an attempt to reduce the base resistance. The key thought of their proposition is to increase the effective base thickness without decreasing the quantum resonant effect.

Practically in all the schemes based on unipolar structures and from the technological point of view we have to successfully fabricate a very shallow Ohmic contact. The main motivation is to avoid short circuiting between the base and the collector region. Such a requirement is often difficult to be realised using standard

Ohmic contact metalisation (AuGeNi). We will come back on this key issue later on in this work.

In the last paragraphs we emphasised the potential control method elaborated for a unipolar structures. It is also possible to realise a direct potential control by the deposition of an Ohmic-contact onto the base layer of Heterojunction Bipolar Transistor (HBT).Figure 1.14 depicts the Pseudomorphic BiQuaRTT energy band profile in GaAs/AlGaAs base materials, in equilibrium (a),and under bias to bring the emitter electron energy into resonance with the fourth allowed quantum well base state (after [30]).

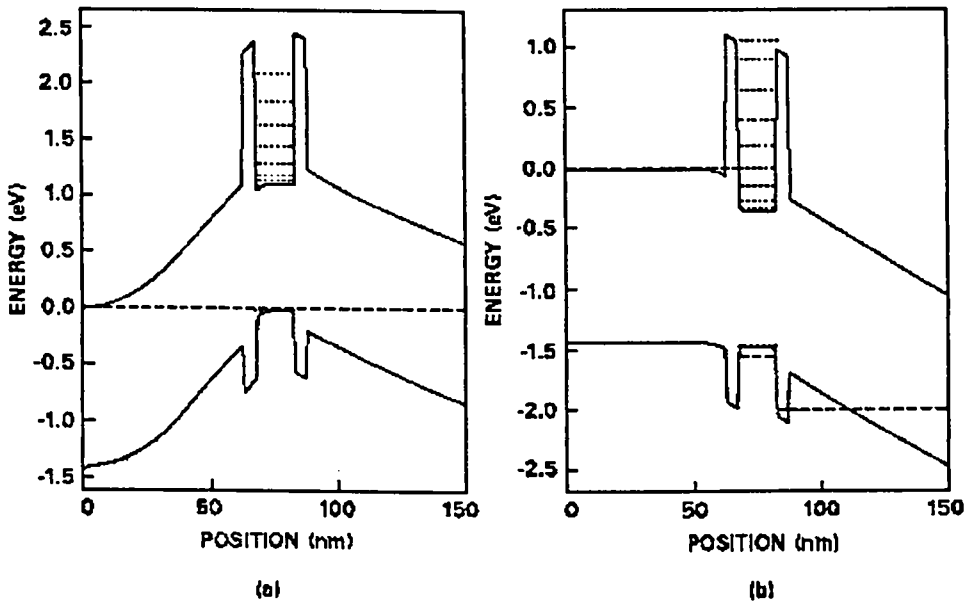


Figure 1.14 Pseudomorphic BiQuaRTT energy band profile, in equilibrium (a), under bias to bring the emitter electron energy into resonance with the fourth allowed quantum well base state. After [30]

It appears interesting to mention that the use of a lower band-gap material in the base region of BiQuaRTT is also favourable for a short bas transit time. The reasons are twofold : first of all we take benefit of high mobility, secondly hot electrons are injected into the base.

Another embodiment of the BiQuaRTT [26] which made use of superlattices in the emitter and collector sides. This provides an effective wider band-gap emitter than the GaAs quantum well. This supperlattice structure has also the advantage of hot electron injection. Besides, it increases the emitter injection efficiency by compressing the energy distribution of electrons in the emitter and collector sides which is a direct result of the formation of minibands.

1.6.1.2 Schottky contact control :

The second class of devices which enables the modulation of the tunnelling current through the double barrier by means of a third electrode concerns the Schottky based structures(Figure 1.15).

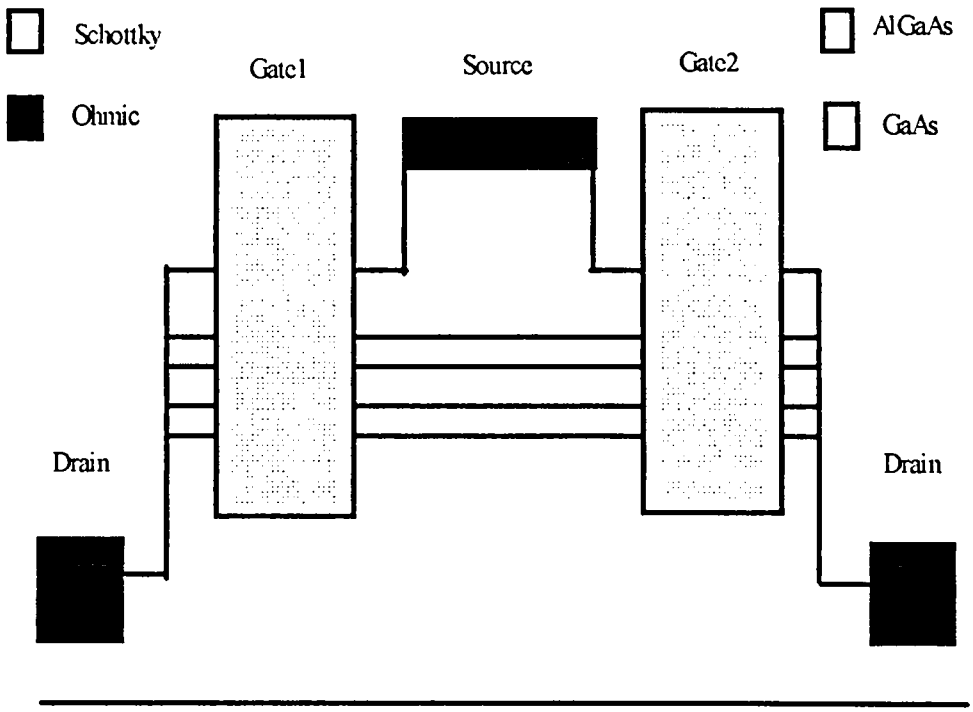


Figure 1.15Schematic cross-section of the heterodimensional Schottky-gated resonant tunnelling transistor (after Peatman et all.[27])

One of the first attempts published in the literature makes use of the deposition of side Schottky gates in close proximity to a buried double barrier heterostructure. In this case the fringing field induced by the gate electrodes allows the control of the

electron flux flowing from source to drain by field effect (this technique is also known as field effect control method). Some improvements in the implementation of the gates was published recently [27] involving a deep etching down to the collector drain region followed by the disposition of a Schottky metalization.

A very good example illustrating this technique is the structure proposed by Peatman et all. [27]. In this structure (Figure 1.15) the Schottky gate voltage modulates the drain current by modulating the effective cross-section area of the quasi-two dimensional electron accumulation layer (injection layer) which is created just above the source-side barrier under the influence of the drain-source bias.

1.6.2 Co-Integration of resonant tunnelling structure with active devices :

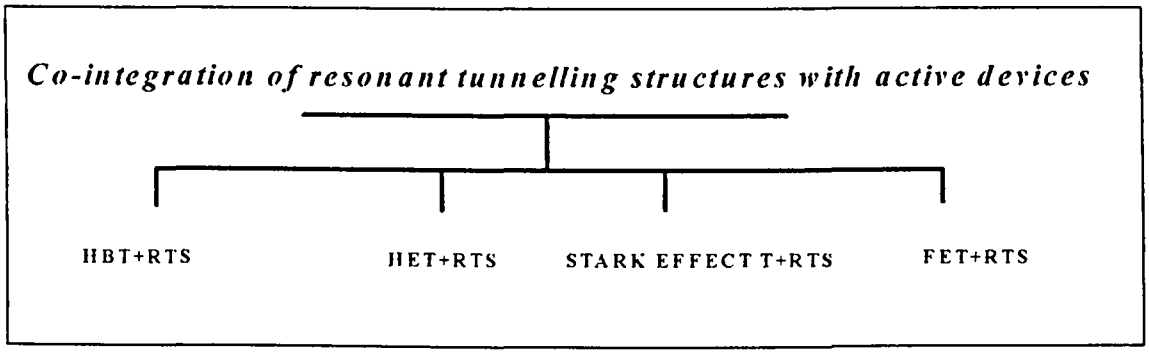


Figure 1.16 Family tree illustrating the main co-integration possibilities between the resonant tunnelling structure and active devices

In the previous section, we emphasised the fact that it is often difficult to satisfy the requirements of both a low resistance Ohmic contact and in the same time well-resolved quantum effects. To alleviate such a draw back one possible way is to separate the quantum structure, namely the double barrier heterostructure which is in a two terminal configuration, from the control cell which is in a three terminal scheme. By this means, one can preserve the main interesting feature of quantum devices notably the NDC effect while permitting the modulation of the overall current by an external voltage. However, it is worth-mentioning that the main disadvantage of this co-integration is that the frequency capabilities is governed by

the transistor part which has conventional transit time limitations. It is now well known that the search of a very high cut-off frequency depends on the gate length which has to be on sub-micron scale for pico-second time response. In this context, the fabrication steps of such high performance transistors requires an advanced level of technology in order to match the capabilities of both parts. Practically, the transistor part and the quantum structure have to be compatible in terms of intrinsic response time and current densities which governs the RC time constant.

1.6.2.1 Resonant Tunnelling Bipolar Transistor (HBRT) :

It is known that the key parameters to be optimised in the bipolar transistor are the base resistance and the base transit time. The former might be decreased by increasing the base doping level. Consequently undesirable increase in the base-emitter parasitic current results, which in turn might be avoided by introducing a larger band gap material as an emitter region. This creates in the valence band a higher barrier reducing the hole base parasitic current. Whereas in the conduction band the emitter-base hetero-junction exhibit on each side an accumulation and a depletion zones with a spike-shaped conduction edge. This decreases the macroscopic thickness of the base-emitter depletion width as compared with that of the conventional homogeneous p-n junction.

By introducing a wider band gap material for forming the resonant tunnelling structure in the emitter region, some of the above improvements could be further enhanced, but more importantly it gives the device the unique feature of negative differential conductance.

On the other hand, beside the room temperature operation possibility, the main application advantages of this co-integration can be summarised in the followings :

However, the fact that the flow of a high bas-emitter parasitic current degradates the device performance is not always true. Indeed, it can be shown [29] that a relatively high base-emitter voltage $\approx 1.0\ V$ is needed to drive the device into the in-resonance operating mode. Under this condition, the base current is relatively high,

hence reducing the current gain reduction beyond resonance. This causes an appreciable increase in the collector current PVR in the common-emitter transistor configuration under constant base current bias. A $|PVR|$ of 74 :1 is measured in the transistor operation mode whereas a $|PVR|$ of 17 :1 is measured in the double barrier operation mode for the same transistor of [29].

For having an overview of various solutions reported in the literature,

Figure 1.17 depicts a comparison of band structures [30] between the different RTBT's and the Bipolar Quantum Resonant-Tunnelling Transistor (BiQuaRTT) of Reed et al. [22].

In Figure 1.17 the top two and center structures are GaAs/AlGaAs materials whereas in the lower half of the figure, the devices were fabricated in the InAlAs/InGaAs system lattice-matched to InP.

Let us recall however that the BiQuaRTT depicted in the center of Figure 1.17 belongs to the direct access heterostructure family shortly discussed in section 1.6.1.

Finally one can conclude that the RTBT or BiQuaRTT transistor design needs the optimisation of a large number of parameters. Which offers a large degree of flexibility in the device design. In counterpart, such an optimisation using band gap engineering is a rather difficult task.

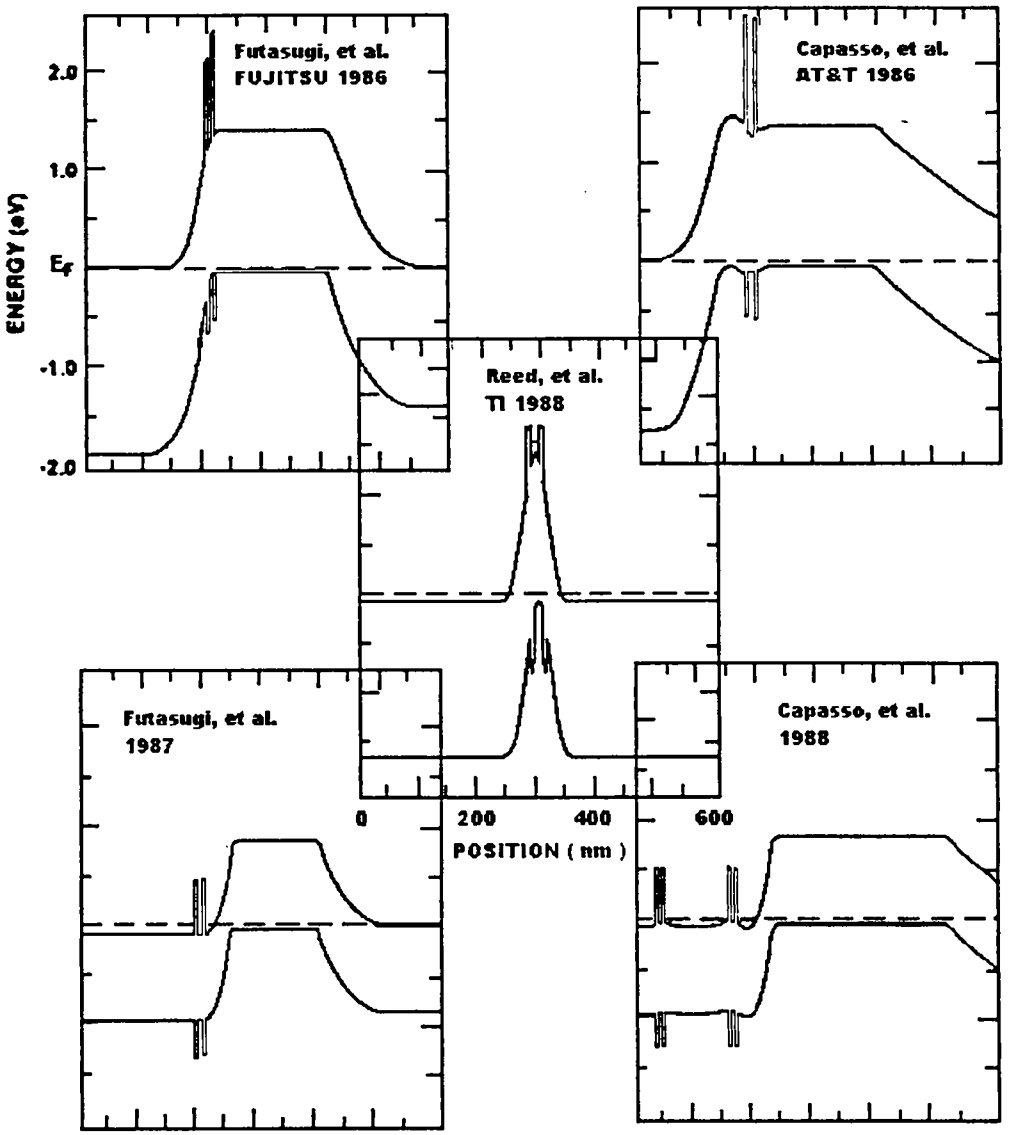


Figure 1.17 different band profiles of RTBTs illustrating a comparison between the different RTBTs and the Bipolar Quantum Resonant-Tunnelling Transistor (BiQuaRTT) of Reed et al. [22]. After [30]

1.6.2.2 Resonant Hot Electron Transistor :

A Resonant hot electron transistor (RHET) is an emitter-base-collector uni-polar vertical structure co-integrated in the emitter with a resonant tunnelling diode. The RHET uses a resonant tunnelling double barrier as a hot carrier injector as shown in Figure 1.18. It is to be mentioned that the conventional hot electron transistor uses a Fowler-Nordheim tunnelling barrier as a carrier injector [28].

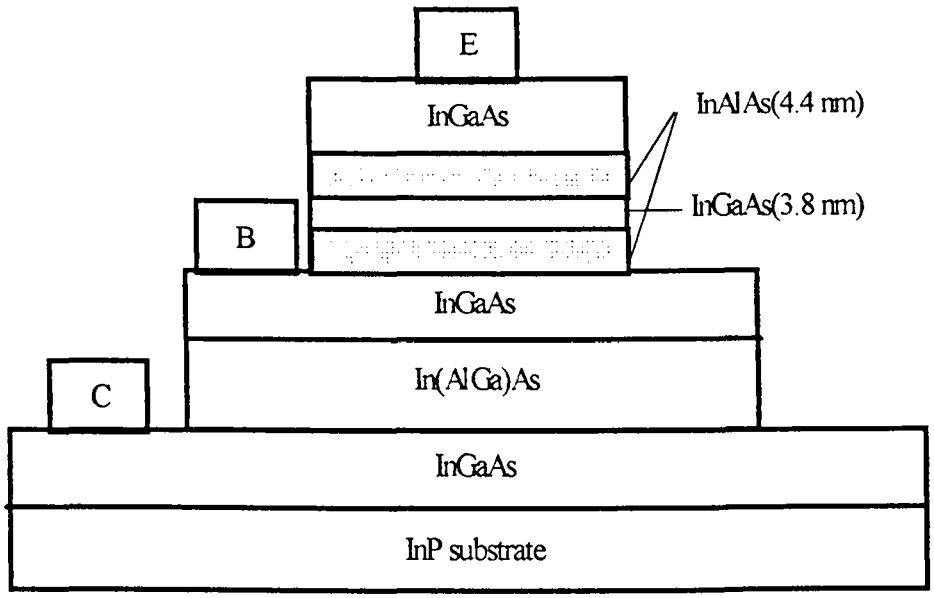


Figure 1.18 Schematic cross section of InGaAs based RHET

after Yokoyama et all.[31][32]

As the resonant tunnelling part injects a monoenergetic hot electron beam, the RHET is more suitable for investigating the physics of hot electron transport. The key design criteria in optimising the injection energy of hot electrons are : (i) the injected electrons energy has to be adjusted on one hand for avoiding any inter-valley scattering leading to drastic mobility degradation and on the other hand to surmount the collector barrier. (ii) to conserve the monenergetic feature of the hot injected electrons, the device has to be operated under low temperature conditions. For avoiding the phonon scatterings in the base layer.

In fact, it was demonstrated [29] that the energy difference between the Γ and L-points in the collector barrier is the allowed energy window through which the injected electrons could be collected into the collector side of the device without degradation of transport characteristics.

It is worth noting that the optimisation of this energy window is a tricky task. In fact there is a trade-off between the collector current capability which is enhanced by decreasing the collector barrier and the base-collector current which becomes higher. On the other hand, there exists a stringent requirement from the technology side indeed the collector epilayer. It should be low-doped thick layer to decrease the base-collector capacitance. Under this condition the growth of a mismatched epilayer is problematic.

One of the solutions to alleviate such drawbacks is the elaboration of structure fabricated from InP-based material system rather than from the conventional GaAs based system. Imamura and co-workers [31] have thus fabricated a RHET using InGaAs/In(AlGa)As as a new material based system. The Γ and L-points of the new material is linearly dependent on the Al percentage which makes the design task easier and more flexible (from the point of view of conduction band edge profile). In addition, the new material has a wider Γ-L-points separation of about 0.53 eV as compared with 0.31 eV for GaAs material (see also Figure 1.7).The wider Γ-L-points separation in the collector barrier permits to relatively increase the collector barrier and hence reducing the base-collector leakage current. Finally the current gain can be improved. In addition, the wide separation between Γ and L-points at the base-collector interface reduces the scattering probability into the upper satellite valley.

So fare we discussed the conditions of ballistic transport in the base region and the avoidance of inter-valley scattering. With respect to the double barrier transport properties, the use of InAlAs as the barrier material, which is a direct-gap semiconductor, reduces the indirect valley tunnelling effects thus decreasing the excess valley current (for further study on the excess current in double barrier see

also [33]). On the other hand, the InAlAs barrier has a smaller effective mass as compared with the effective mass of AlGaAs barrier. Consequently a higher tunnelling current densities can be achieved, with simultaneously a higher PTV ratio and a shorter switching time due to the decrease in the charging time of the tunnelling capacitance.

In summary, thanks to the ballistic transport over the active region, RHET's exhibit a high speed operation. In counterpart, the key limitation concerning the low temperature operation requirement is still a very important constraint.

1.6.2.3 Inverted base-collector tunnel transistor (Stark effect transistor) :

In the previous sections we show how it is difficult to satisfy the trade-off between the base leakage current which causes a degradation in the device performance and the realisation of a good accessibility to the base layer. One of the elegant solutions which was proposed in reference [34], is to interchange the relative positions of the base and collector from the conventional emitter-base-collector sequence (Figure 1.19).

Worth noting that the large collector-base barrier shown in Figure 1.19 necessitates the application of a relatively high base voltages ($\approx 4V$), to permit an effective change in the potential conditions within the inverted collector quantum well.

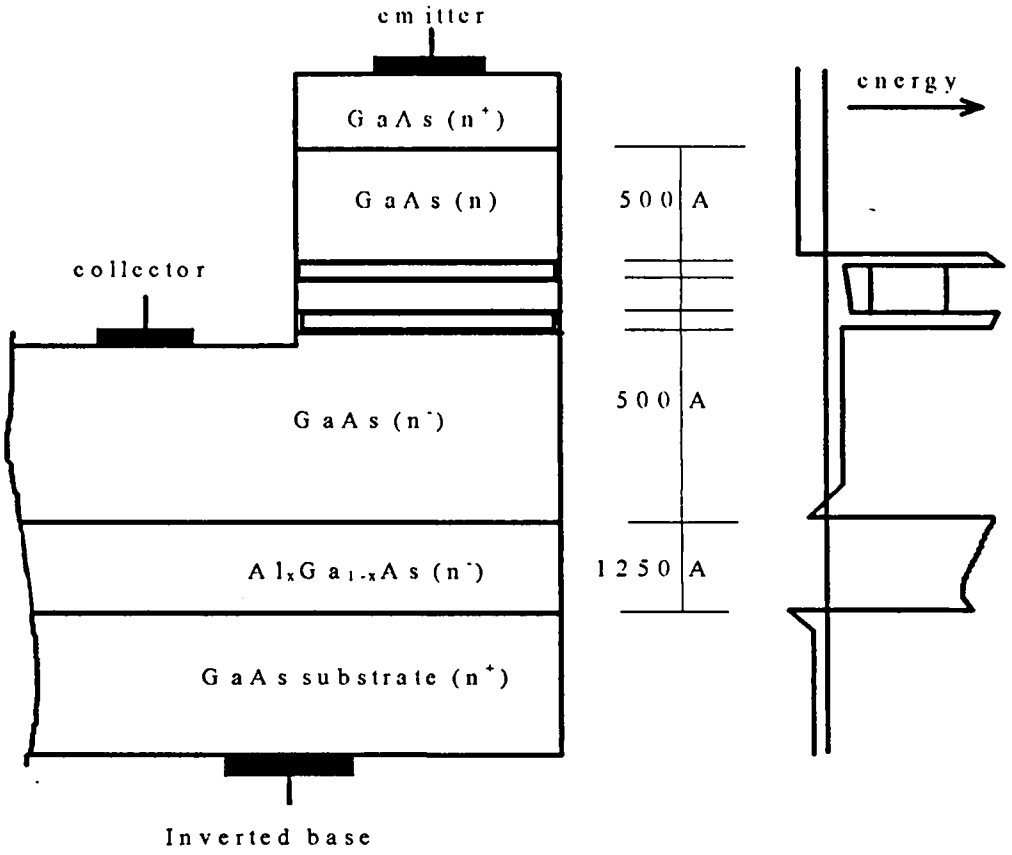


Figure 1.19 Schematic diagram of a cross section of the proposed negative resistance Stark effect transistor, the conduction band profile is also illustrated.

The advantage of this type of structure is the relatively large collector-base barrier which suppresses the base current to a very large extend. Not to mention that the relatively high inverted-base voltages implies the condition of low temperature operation. This is to prevent the base thermal current leakage from the collector to the base.

The design parameters are summarised as follows [34] :

- the barriers and the quantum-well are lightly doped.
- the $Al_xGa_{1-x}As$ barrier is not too thick to drop all the base-emitter voltage.
- the device geometry is such that the collector contact does not completely shield the emitter from the inverted-base.

From the fabrication view point however, the technological difficulty of realising a very shallow Ohmic contact is no more existing. This is because the wide low doping collector-base barrier avoids the short circuit between the base and the collector electrodes.

Subject to the above design conditions, the electric fields produced by applying a voltage to the base can modulate the positions of the resonant levels in the quantum well and thus the base voltage can control the emitter-collector tunnelling current. Although these device were analysed in the context of $GaAs/Al_xGa_{1-x}As$ heterojunction technology, other semiconductors could be used.

1.6.2.4 Resonant Tunnelling Field Effect Transistor (RTFET) :

Another example is the co-integration of resonant structure with FET transistor [35]. This co-integration can be made by the series connection of two discrete devices namely a two terminal resonant tunnel diode and a commercial FET transistor. Obviously the monolithic integration permits to reduce the geometric dimensions, the parasitic elements and increases the device performance in terms of AC and noise characteristics. These are essential needs for high-speed applications. In the structure of [35], the resonant tunnelling carrier injector is directly integrated with the source electrode see also Figure 1.20.

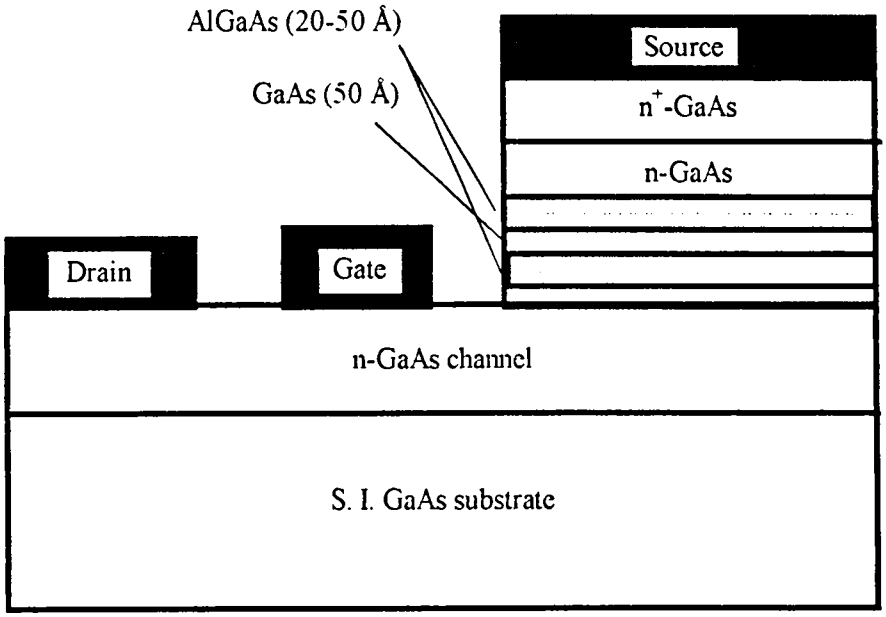


Figure 1.20 Schematic diagram of cross section of resonant tunnelling MESFET

One of the key issue in the RTFET operation is related to the energy relaxation mechanisms of electrons before experiencing a tunnelling effect. we have to stress on the fact that hot electron effects occurs within the transistor channel. Under this condition it is important to cool the electron population in order to avoid any high temperature tunnelling effects which are known to degradate the overall device performance. To this aim, various solutions exist including the use of a large drain-

gate interdistance, the growth of a high doped epilayer, or the implementation of the DBH in the source side of the device. Clearly, an optimisation effort is needed with here again, some trade-off between the overall device response time and the complexity in the fabrication.

Worth noting that there is no restriction on the room operation of the device structure illustrated in Figure 1.20.

To achieve an efficient voltage swing on the resonant tunnelling part, the channel resistance under the gate should be smaller than the effective resistance over the double barrier. This means that the FET part should be operated in the high conductance linear regime. Gradually, as the gate bias is increased, the required drain voltage which gives the resonant condition is increased. At the earlier reverse gate voltages the channel resistance is almost perfectly linear and hence the overall I-V characteristic is identically shifted to a higher drain potential due to the increase in the voltage drop across the channel. On the other hand in the later reverse bias voltages where the effect of FET-channel non-linearity is highly pronounced, the I-V shift is associated with a decrease in both the peak current and the PVR. Finally at large reverse gate voltages near the pinch-off the channel resistance absorbs practically the whole drain voltage and the resonant part is in out-of-resonance condition.

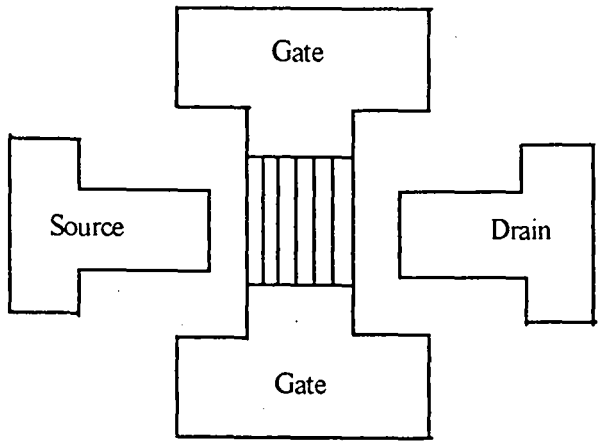
Although this type of devices is introduced in the context of GaAs/AlGaAs in Figure 1.20, it could be implemented in other materials. Therefor, it may be advantageous to use InGaAs material which has better transport properties notably a higher electron mobility.

1.6.3 Field induced tunnelling structures :

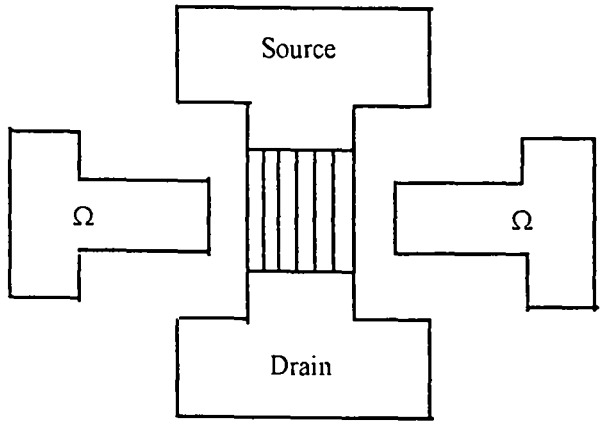
The field induced resonant tunnelling structures depends on the idea of creating field effect induced quantum wells using lateral Schottky nano-contacts. Khalid Ismail, Dimitri A. Antoniadis and Henry I. Smith had fabricated many types of field effect tunnelling structures [36], the lateral-surface superlattice devices (LSSL) [36], the multiple parallel quantum wires (MPQW) [36] and the planar resonant tunnelling field effect transistor (PRESTFET) [36]. Figure 1.21 shows a schematic design of each type of devices. It is to be mentioned that the idea of a PRESTFET was firstly suggested and implemented by Chou et al. [37].

The planar version of quantum structures can avoid most of the shortcomings of the vertical devices discussed above notably ; the achievement of a good contact to the nano-epi-layers while avoiding the short-circuiting between them. Not to mention that the planar structures are more promising to circuit integration. The key limitation at the present time is the requirement of a very low temperature operation. As the inter-gate separation is relatively large (technology limited to ≈ 50 nm) then the associated induced quantum potentials are wide. This leads to relatively narrow energy separations between the corresponding resonant levels.

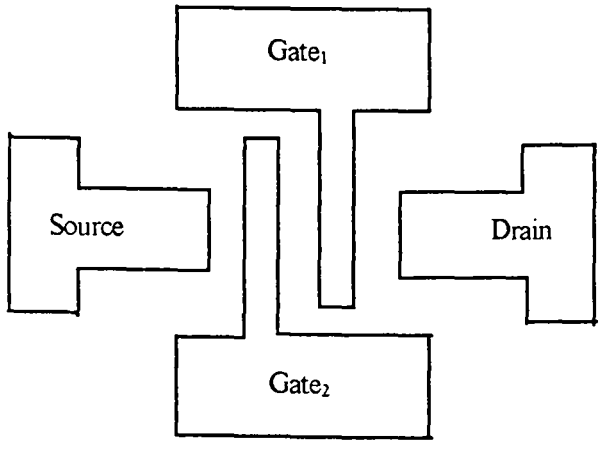
It is clear that thermal effect will play an important role concerning this planner devices. First of all, any increase in the operating temperature will tends to smear out the quantum level ladder. Second, non coherent transmission mechanisms will strongly influence the conduction properties. At last, thermally activated processes will occur degrading the resonant features.



(a) top view of the grating gate FET (LSSL)



(b) top view of the MPQW.



(c) top view of dual-gate PRESTFET

Figure 1.21 (14) shows schematic design of each type fabricated by Kkaled Ismail et all after Kkaled Ismail et all [36]

If one now tries to alleviate these drawbacks by employing ultra narrow gate separations, one faces the problem of the fringing effects. Indeed the overlap of the potential induced by each gate finger tends to decrease the difference between the maxima and the minima in the overall potential profile. In addition the resonant states are pushed up to higher energy and become non accessible. Basically they may exist different solutions using either an optimised doping profile and/or a buried gate technology but nevertheless it appears that room temperature operation is a very challenging issue and its research field is steel opened.

1.6.4 Solutions combining optical control gate :

To increase the functional capability of the resonant tunnelling diode, another option rather than adding an electrode is the optical control. This is mainly achieved through the use of inter-band light absorption in the neighbourhood of the double barrier. When electron-hole pairs are generated for a device under illumination, the electrons are instantaneously swept out from the depletion region. The latter is designed to absorb most of the incident optical power. In contrast, the holes are accumulated near the hetero-interface barrier [38][39]. It is noticed that a steady state condition will be reached for an equilibrium balance between the electron-hole generation rate and the hole escape current. The key role of the accumulated holes is to modify the internal field. As a consequence, the conduction band profile along with the resonant states are modified by the external light.

Figure 1.22 shows the advantage of resonant tunnelling photo-detector over the conventional one as regards the input/output conversion ratio. This figure illustrates the relative shift between the two operating points (P1, P2) on the load-line arising from the optically controlled shift of the V-I characteristics.

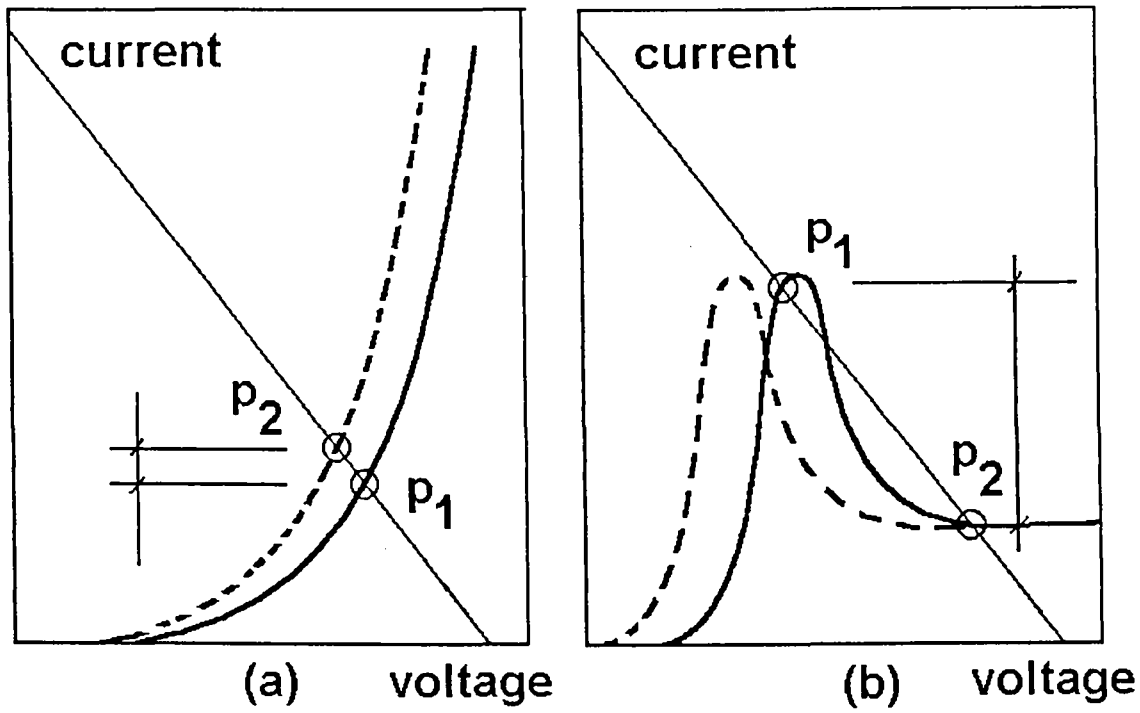


Figure 1.22 schematic diagram illustrating a comparison between the conventional photo-detector (a) and the NDR one (b).

1.7 Conclusion :

We have discussed in this first chapter the various physical mechanisms responsible for negative differential conductance in the current voltage characteristics of a two terminal device. This concerns mainly the inter-band tunnelling and the intra-band resonant tunnelling structures, the k-space and real space devices and some generalisations to supperlatices.

Som fundamental concepts have been introduced concerning the device structure, the inherent physics, the current voltage characteristics and in some cases the intrinsic response time. In the following chapter we will tray to improve our understanding about NDC devices by focusing our attention on resonant tunnelling structures. This will concern the two dimensional injection effects (chapter 2) and incoherent transmission effects (chapter 3).

As regards the three terminal devices, we have reviewed some of the most promising options by discussing, on one hand, the possibility of direct access into the quantum well by an Omic or Shottky contact control and on the other hand, the co-integration of resonant tunnelling structure with active devices. On this area, more specific studies will be performed with special attention paid to the field induced resonant tunnelling structures (chapter 4).

1.8 references :

[1] Capasso et al., Quantum functional devices : resonant tunnelling transistors, circuits with reduced complexity, and multiple-valued logic, IEEE Transactions on Electron Devices, Vol.36, No.10,2065 (1989).

[2] A. Sellai, M.S.Raven, D.P.Steenson, J.M.Chamberlain, M.Henini, and O.H.Hughes, Double-barrier resonant tunnelling diode three-state logic, Electronics Letters, January 1990, Vol.26 No.1.

[3] Yokoyama N., Resonant Tunnelling Hot Electron Transistor (RHET) :potential and application. Bstract of the 18th Conference on solid state devices and materials, Tokyo, p. 347-350, 1986.

[4] L.Esaki, New Phenomenon in Narrow Germanium p-n Junctions, Phys.Rev.,109, 603 (1958).

[5] L.Esaki, Discovery of the Tunnel Diode, IEEE Trans. Electron Devices, ED-23, 644 (1976)

[6] Ebrahim Saadoun, M.Sc. theses , Cairo University 1983.

[7] T. Shawki, G Salmer and O.L.El-Sayed, MODFET 2D hydrodynamic energy modeling : optimization of subquarter-micron-gate structures, IEEE Trans. Electron Devices, Vol.ED-37, p.21-30,1990.

[8] Tarek Ahmad Shawki, thèse de doctorat, Université des Sciences et Technologies de Lille, 1990.

[9] Khaled Mahmoud Sherif thèse de doctorat, Université des Sciences et Technologies de Lille, 1994.

[10] H. Kroemer, Negative Conductance in Semiconductors, IEEE Spectrum, 5, 47 (1968).

[11] K. Hess, Real space transfer : Generalised approach to transport in confined geometry, Solid-State Electron., vol. 31, p319, 1988.

[12] R.Sakamoto, K. Akai, and M. Inoue, Real-space transfer and hot-electron transport properties in III-V semiconductor heterostructures, IEEE Trans. Electron Devices, vol. 36, p2344, Oct. 1989.

[13] J. Lasker, A. A. Ketterson, J. N. Baillageon, T. Brock, I. Adesida, K. Y. Cheng, and J. Kolodzey, Gate-controlled negative differential resistance in drain current characteristics of AlGaAs/InGaAs/GaAs pseudomorphic MODFET's, Electron Device Lett. Vol. 10, p528, Dec.1989.

[14] A. Kastalsky, S. Luryi, A.C. Gossard and R. Hendel, A field-effect transistor with z negative differential resistance., IEEE electron device letters, VOL. EDL.5, No.2, February 1984.

[15] Adel Refky Mikhail, Performance analyses of dual gate modulation doped field effect transistor, Master degree thesis, Cairo University, 1993.

[16] Khaled Sherif, Adel Refky, Tarek Shawki, Osman El-Sayed and Georges Salmer., two-dimensional Hydrodynamic Simulation of Sub-micrometer Dual Gate MODFETs. Solid-State Electronics Vol.38 No4,pp. 917-929, 1995.

[17] Thomas E., Koscica and Jian H. Zhao, Field effect real space transfer transistor, IEEE Electron Device Lett. Vol. 16, May 1995.

[18] Thomas E., Koscica and Jian H. Zhao, Frequency doubling in GaAs/AlGaAs field effect transitor using real space transfer, IEEE Electron Device Lett. Vol. 16, Dec. 1995.

[19] Jan-Shing Su, Wei-Chou Hsu, Yu-Shyan Lin, Wei Lin, Chang-Luen Wu, A novel InAlAs/InGaAs two-terminal real-space transfer diode, IEEE Electron Device Lett. Vol. 17, Feb. 1996.

[20] S. Mohan, P. Mazumder, R. K. Mains, J. P. Sun, and J. I. Haddad, Ultrafast Piplined Adders Using RTTs, Electronics Letters vol.27 No.10 May 1991.

[21] G. I. Haddad, R. K. Mains, U. K. Reddy, and J. R. East, A proposed narrow band-gap base transitor structure, Superlattices and Microstructures, 5,(3),p. 437, 1989.

[22] Reed M. A., Frensley W. R., Matyi R. J., Randall J. N., Seabaugh A.C., Realisation of a three terminal resonant tunnelling devices : the Bipolar Quantum Resonant Tunnelling Transistor. Appl. Phys. Letters, 54 (11), p.1034-1989.

[23] Shulman J. N., Waldner M., Analysis of second level resonant tunnelling diodes and transistors. Journ. of Appl. Phys., 63(8),p.2859-1988.

[24] D. Lippens, P. Mounaix, H. Leroux, O. Vanbésien, V. Sadaune and E. Barbier, Buried-quantum well resonant tunnelling transistor in the $Al_xGa_{1-x}/In_yGa_{1-y}As$ material system. Int. Workshop on Quantum Effect Physics, Luxor, 1992.

[25] T. Waho, K. Maezawa and T. Mizutani, Resonant Tunneling in a Novel Coupled-Quantum-Well Base Transistor : Jpn. J. Appl. Phys. 30 (12A) p. L2018-1991.

[26] M. A. Reed, W. R. Frensley, R. J. Matyi, J. N. Randall, and A. C. Seabaugh, Appl. Phys. Lett., 54, 1034 1989.

[27] W. C. B. Peatman, E. R. Brown, M. J. Rooks, P. Maki, W. J. Grimm and J. Shur, Novel resonant tunneling transistor with high transconductance at room temperature, IEEE Electron Device Letters vol. 15, (7) July 1994.

[28] Heiblum, M., Solid State Electron. 24, 343-366, 1981.

[29] Alan C. Seabaugh, Edward A. Beam, Albert H. Taddiken, John N.Randall and Yung-Chung Kao, Co-integration of resonant tunneling and double heterojunction bipolar transistors on InP, IEEE Electron Device Letters, Vol.4 No.10, October 1993.

[30] A. C. Seabaugh, W. R. Frensley, Y. C. Kao, J. N. Randall, and M. A. Reed ,Quantum-Well Resonant-Tunnelling Transistors, (invited), IEEE/Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits, Cornell University Ithaca, New York, August 7-9, 1989.

[31] Imamura, K., S. Muto, H. Ohnishi, T. Fujii, and N. Yokoyama, Program of 45th Annual Device Research Conference, Santa Barbara, 1987.

[32] Yokoyama N., IMAMURA K. , MUTO S., HIYAMIZU S. and NISHI H., A new functional Resonant Tunnelling Hot Electron Transistor RHET. Jap. Jour of Appl. Phys., Vol.24, No.1, p. L853,1985.

[33] Hiroyuki Fukuyama, Takao Waho and Takashi Mizutani, Current-voltage characteristics of GaAs/AlAs double-barrier resonant tunnelling diodes with a Si-planner-doped barrier, J.Appl.Phys.79(3), p1801, February 1996.

[34] A.R. Bonnefoi, D.H. Chow, and T.C. McGill, inverted base-collector tunnel transistor, Appl. Phys. Lett. 47(8), October 1985.

[35] A. R. Bonnefoi, T. C. McGill, and R. D. Burnham, Resonant Tunneling Transistors with Controllable Negative Differential Resistances, IEEE Electron Device Letters, Vol. EDL-6, No. 12, December 1985.

[36] Khalid Ismail Ph.D these, Massachusetts Institute of Technology (MIT), May 1989.

[37] S. Y. Chou, J. S. Harris, Jr., and R. F. W. Pease, Appl. Phys. Lett. 52, 1982 (1988).

[38] H.S.Li,Y.W.Chen, K.L.Wang, D.S.Pan, L.P.Chen, and J.M.Liu, J.Vac.Sci.Technol.B 12, 1269 (1994).

[39] C.R.H.White, M.S.Skolnick, L.Eaves, M.L.Leadbeater, M.Henini, O.H.Hughes, G.Hill, and M.A.Pate, Phys.Rev.B 45,6721 (1992).

CHAPTER 2

2. EFFECT OF TWO DIMENSIONAL INJECTION ON PURELY COHERENT RESONANT TUNNELLING :

2.1 Introduction :

In a real double barrier heterostructure (DBH) device the injection region exhibits a modulation of doping concentration. This is to satisfy the trade-off between a low resistance access region and a high quality heterostructure. As a consequence, the cladding layer of most of the earlier devices have relatively low doping ($N_d \approx 10^{16} cm^{-3}$) along with undoped spacers [1][2]. In this condition, when the device is under bias, an accumulation zone (Figure 2.1a) is formed in front of the DBH. This accumulation layer strongly modifies the injection condition. This conclusion is widely recognised in the literature [3]. Notably, it was established that the picture of a three dimensional injection is no-longer valid, and that the resonant condition stems from the anti-crossing between the quantum levels attached, from one side, to the accumulation layer and, from the other side, to the quantum well itself [4][5][6][7].

In addition, the two dimensional injection character is enhanced by means of a heterostructure quantum well which is buried prior to the DBH [8][9]. Such an idea is illustrated in Figure 2.1b with buried wells cladding the DBH.

We have to emphasise that this effect is intrinsic and does not concern the loss of coherence resulting from the different scattering mechanisms.

From the point of view of the I-V properties, and in the scope of the 2D injection effect, there would be some expected modifications in resonant voltage conditions. This is concerning a steep decrease in the tunnelling current would be occurred when the resonant state of the quantum well is shifted just below the quantum state of the accumulation layer.

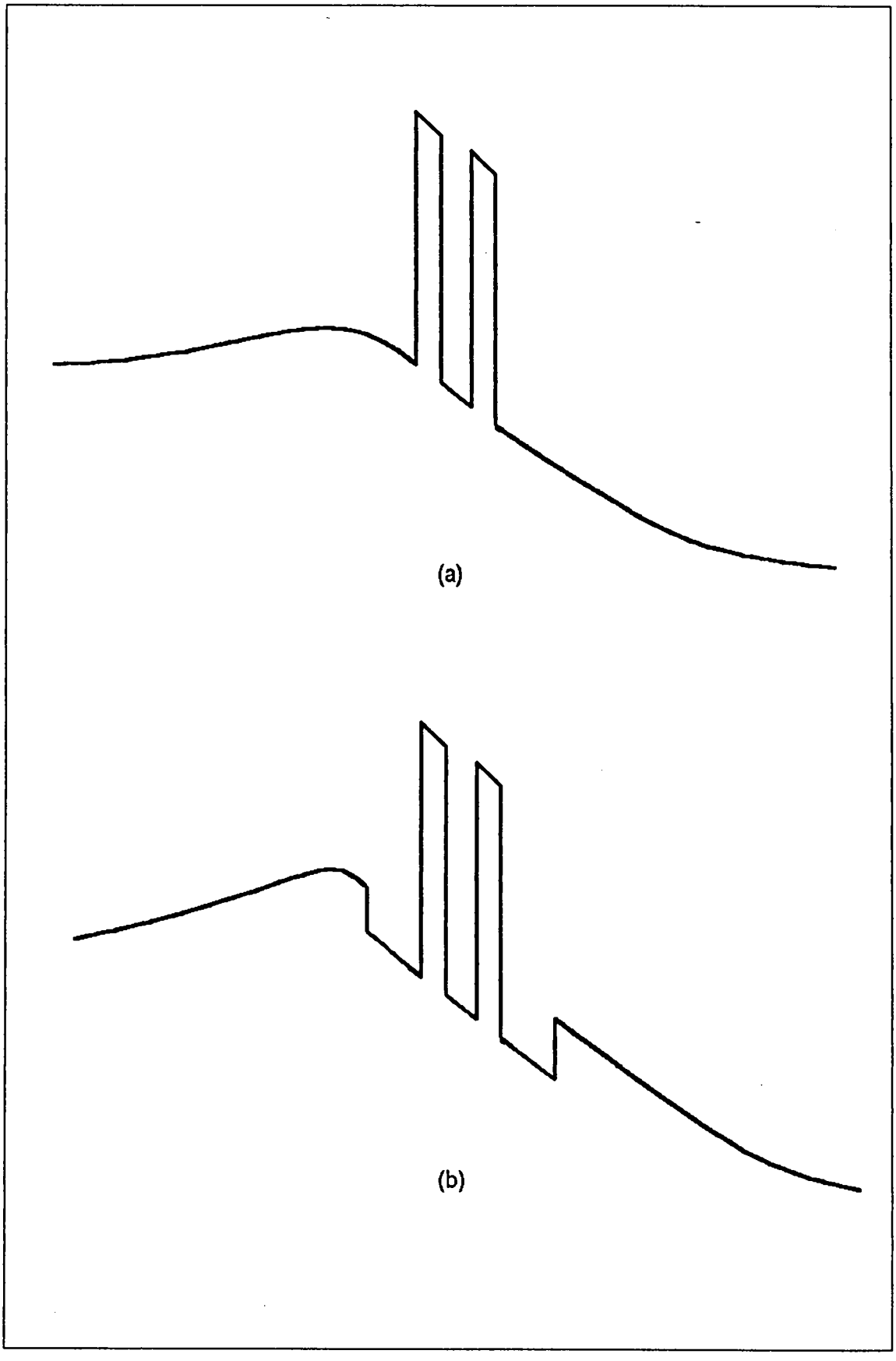


Figure 2.1 schematic band profile of DBH under bias with the formation of accumulation layer (a), and with the buried well cladding layers (b).

However experimentally, the transition between the on-resonant and the off-resonant states, is not found so sharp. In the earlier studies, this was often explained as an effect of the relatively low doped cladding layers which contribute to the overall voltage drop across the device.

Recently, due to the progress in the modern epitaxial growth technology, along with the motivation to increase the peak current density, it was proposed to sandwich the DBH in-between highly doped regions , let us say ($N_d \approx 10^{18} cm^{-3}$).

Surprisingly, although the effect of the low doped cladding layer is suppressed, the device fabricated using this idea, exhibited a relatively wide voltage range over which the NDC effect can be seen.

In this chapter, we shall try to address this issue which is of a key importance in the understanding of resonant tunnelling devices. As an example, this voltage range determines the impedance level and the output power when the device is used as an oscillator. Also for the digital applications, this voltage range is of prime importance when a steep transition is often preferable.

To deal with this issue, we will distinguish between two different aspects : First, the effect of scattering assisted phenomena which tend to sustain a finite transmission probability in the out-of-resonance condition, which is studied in the next chapter. Second, the coupling between the accumulation quantum well and the DBH quantum well, which is investigated in details in this chapter.

To support this separation of the two effects, we shall assume a very low temperature operation (4.2 K).

The first step in the analysis will be to show that a purely coherent 3D/3D approach is not appropriate to describe the injection properties in a real device. On this basis we shall discuss the development of a new model which defines a two dimensional supply function. One of the main problems encountered, is related to the openness of the quantum system. Concerning this issue, we will start from the notion of local density

of states previously introduced by Lassning under flat band condition [10] which will be generalised in the present work.

Special attention will be paid to the wave function normalisation and some preliminary results will be also discussed by introducing the effect of non-parabolicity which influences the effective mass in the accumulation zone.

Finally, to support the validity of the new formalism we will report a comparison between the simulation results carried out in the present study and the experiments conducted by the NTT group [11].

2.2 3D/3D purely coherent approach :

2.2.1 Current equation :

It is well known that the probability tunnelling current can be calculated by the following equation [12]:

$$J = q\int_0^{\infty}\int_{P_x}^{\infty}\frac{P_x}{m^*}g(\varepsilon,P_x)f_1T(\varepsilon_x)(1-f_2)dP_xd\varepsilon \qquad \textit{Eq. 2-1}$$

where :

- J is the probability tunnelling current density from the emitter (side 1) to the collector (side 2) and q is the electron charge
- p_x/m^* is the x-component velocity defined as the x-component momentum divided by the effective mass in the x-direction which is taken as the propagation direction.
- $g(\varepsilon, p_x)$ is the density of states in the momentum range p_x to $(p_x+\partial p_x)$ and energy range $d\varepsilon$ to $(\varepsilon+d\varepsilon)$, whereas f_1 is the Fermi-Dirac probability density function of occupancy at side 1. As a consequence the product $g(\varepsilon, p_x)f_1dP_xd\varepsilon$ is the number of electrons in the momentum and energy infinitesimal-windows defined above.

- $T(\varepsilon_x)$ is the tunnelling probability of transmission from side (1) to side (2).

It is to be mentioned that the factor $(1- f_2)$ is introduced to the probability current equation to account for the probability of finding an empty state in the other side. The introduction of this probability of vacancy can be interpreted as that, a complete reflection would be occurred if electron tunnels directly to an occupied state. The validity of the last statement will be discussed later on in this chapter.

After the change of variable ($d\varepsilon_x = P_x / m^* dP_x$) into Eq. 2-1 we get :

$$J = q\int_0^{\infty}\int_{\varepsilon_x}^{\infty} g(\varepsilon, P_x) f_1 T(\varepsilon_x)(1 - f_2) d\varepsilon_x d\varepsilon \qquad \textit{Eq. 2-2}$$

2.2.2 Density of states :

The density of states necessarily required to calculate the probability current density is defined as the subset of states having total energy from $d\varepsilon$ to $(\varepsilon + d\varepsilon)$ and x-momentum from p_x to $(p_x + \partial p_x)$ per unit spatial volume per unit total energy per unit x-momentum (local density of states) [12]

Assuming that we take the well known three dimensional (3D) density of states in momentum space (Eq. 2-3) and multiply it by the total volume of momentum space ($4\pi p^3 / 3$), we get the total number of states enclosed by a sphere of radius p (Eq. 2-4). The constant-momentum sphere is illustrated in Figure 2.2. By taking the first derivative with respect to ε we get the well known 3D aerial density of states in the unbounded system (Eq. 2-5).

$$G = \frac{1}{(2\pi\hbar)^3} \qquad \textit{Eq. 2-3}$$

$$n_t = \frac{1}{(2\pi\hbar)^3}\left[\frac{4}{3}\pi(2m\varepsilon)^{3/2}\right] \qquad \textit{Eq. 2-4}$$

$$\frac{\partial n_t}{\partial \varepsilon} = \frac{1}{(2\pi\hbar)^3} 2\pi (2m)^{3/2} \sqrt{\varepsilon} \qquad \textit{Eq. 2-5}$$

Our problem is different, since we want only a subset of those states corresponding to a momentum from p_x to $(p_x + \partial p_x)$. Thus instead of multiplying by the spherical momentum volume of constant energy sphere we multiply by the volume of spherical slice which lies between two constant planes, p_x and $(p_x + \partial p_x)$ in the momentum space as illustrated in Figure 2.2.

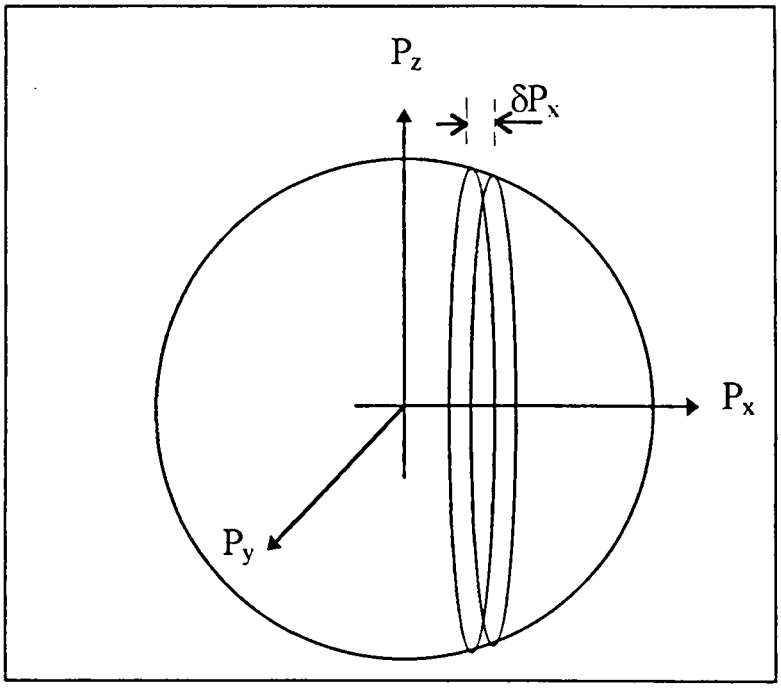


Figure 2.2 constant energy sphere in the momentum space

Thus the required number of states (Δn) is calculated as follows :

$$\Delta n(\varepsilon, p_x) = (2\pi\hbar)^{-3} \pi (2m\varepsilon - p_x^2)\, \delta p_x \qquad \textit{Eq. 2-6}$$

derivating with respect to the total energy gives,

$$\frac{\partial\, \Delta n(\varepsilon, p_x)}{\partial \varepsilon} = \frac{2\pi m}{(2\pi\hbar)^3}\, \delta p_x \qquad \text{Eq. 2-7}$$

which is the total-energy density of states having momentum from p_x to $(p_x + \partial p_x)$. Thus further dividing by ∂p_X we get the total-energy and x-momentum density of states (Eq. 2-8) which is constant independent of neither the energy nor the momentum.

$$g_{3D} = \frac{2\pi m}{(2\pi\hbar)^3} \qquad \textit{Eq. 2-8}$$

We find out that the same result can be obtained by dirivating the total number of states (Eq. 2-4) with respect to the energy (Eq. 2-9), and then with respect to the momentum (Eq. 2-10), not to mention to divide by two to take into account the momentum either in the positive or in the negative direction.

$$\frac{dn_t}{d\varepsilon} = \frac{4\pi m}{(2\pi\hbar)^3}\sqrt{2m\varepsilon} = \frac{4\pi m}{(2\pi\hbar)^3}p \qquad \textit{Eq. 2-9}$$

$$\frac{1}{2}\frac{d}{dp}\left(\frac{dn_t}{d\varepsilon}\right) = \frac{2\pi m}{(2\pi\hbar)^3} \qquad \textit{Eq. 2-10}$$

from Eq. 2-8 and Eq. 2-10 it is noticed that the local density of states in a 3D zone is independent of energy and of position.

2.2.3 3D/3D supply function :

In the literature, a supply function is usually defined in such a way that the tunnelling current density can be evaluated from the integration of the product of this supply function times the quantum tunnelling probability with respect to the energy.

Referring to Eq. 2-2 and taking into account that we must subtract the reverse tunnelling current from side (2) to side (1) we get the net current density :

$$J = qg\int_0^\infty T(\varepsilon_x)\,d\varepsilon_x \int_{\varepsilon_x}^\infty f_1(1-f_2) - f_2(1-f_1)\,d\varepsilon \qquad \textit{Eq. 2-11}$$

consequently the supply function (SF) is defined by the following integral :

$$SF = K_B T \int_{\varepsilon_x}^{\infty} f_1(1 - f_2) - f_2(1 - f_1)\ d\varepsilon \qquad \textit{Eq. 2-12}$$

using simple algebra, it can be shown that SF is written as,

$$SF = \ln\left\{\frac{1 + \exp[1 + (\varepsilon_{f1} - \varepsilon_x)/kT]}{1 + \exp[1 + (\varepsilon_{f2} - \varepsilon_x)/kT]}\right\} \qquad \textit{Eq. 2-13}$$

This equation referred as the supply function, vanishes at equilibrium condition ($\varepsilon_{f1} = \varepsilon_{f2}$).

2.3 comparison between the purely coherent 3D/3D and 2D/3D approaches :

We will discuss of a comparison between 3D injection and 2D injection taking Figure 2.3 as an illustration. In this figure, and by focusing our attention on the emitter zone, we can define three specific regions : first a highly degenerate 3D-layer, then a bump resulting from the doping modulation and at last, the accumulation layer. The bump results from the diffusion mechanism due to the doping gradient and in the same time, from the blocking process of the DBH. Starting from this conduction edge profile, several energy domains and tunnelling paths can be defined with respect to the quantum level in the accumulation layer (ε_{acc}). Bellow ε_{acc} and assuming a purely coherent tunnelling is sustained, the feeding of the accumulation layer is provided by path (a), consequently, through the bump. It is clear that such a probability of supply is extremely weak. On the other hand, when the injection energy matches the quantum level ε_{acc} , a finite probability can be found (path b). Path c describes the case in which the injection energy is just above the bump height. For the later, the electrons have the possibility to experience some relaxation mechanisms permitting extra feeding via the extended states. At this stage, we have the choice between two different approaches. The first one corresponds to a current calculations from emitter 3D region on the left hand side to the collector 3D region on the right hand side of the device. Under this approach, the device is taking as a whole, including both the bump and the DBH. We use Eq. 2-11 with the supply function described by Eq. 2-13 for a

transmission directly through and over the bump. For the calculation of the peak current, despite the crud assumption of this approach, we often take advantage of a relatively efficient tunnelling supply. This stems from the fact that ε_{acc} , at the peak voltage, is usually located close up to the edge height of the bump. In contrast, such an analysis fails to correctly describe the off-resonance condition, with notably a peak-to-valley ratio highly over estimated.

The second approach consists to assume an injection mechanism starting from the accumulation layer. In this region, we have previously noticed the existence of quasi-bound states and also extended states. We will assume a local thermal equilibrium in this zone. This assumption permits us to include the different relaxation mechanisms without considering them in details in this zone.

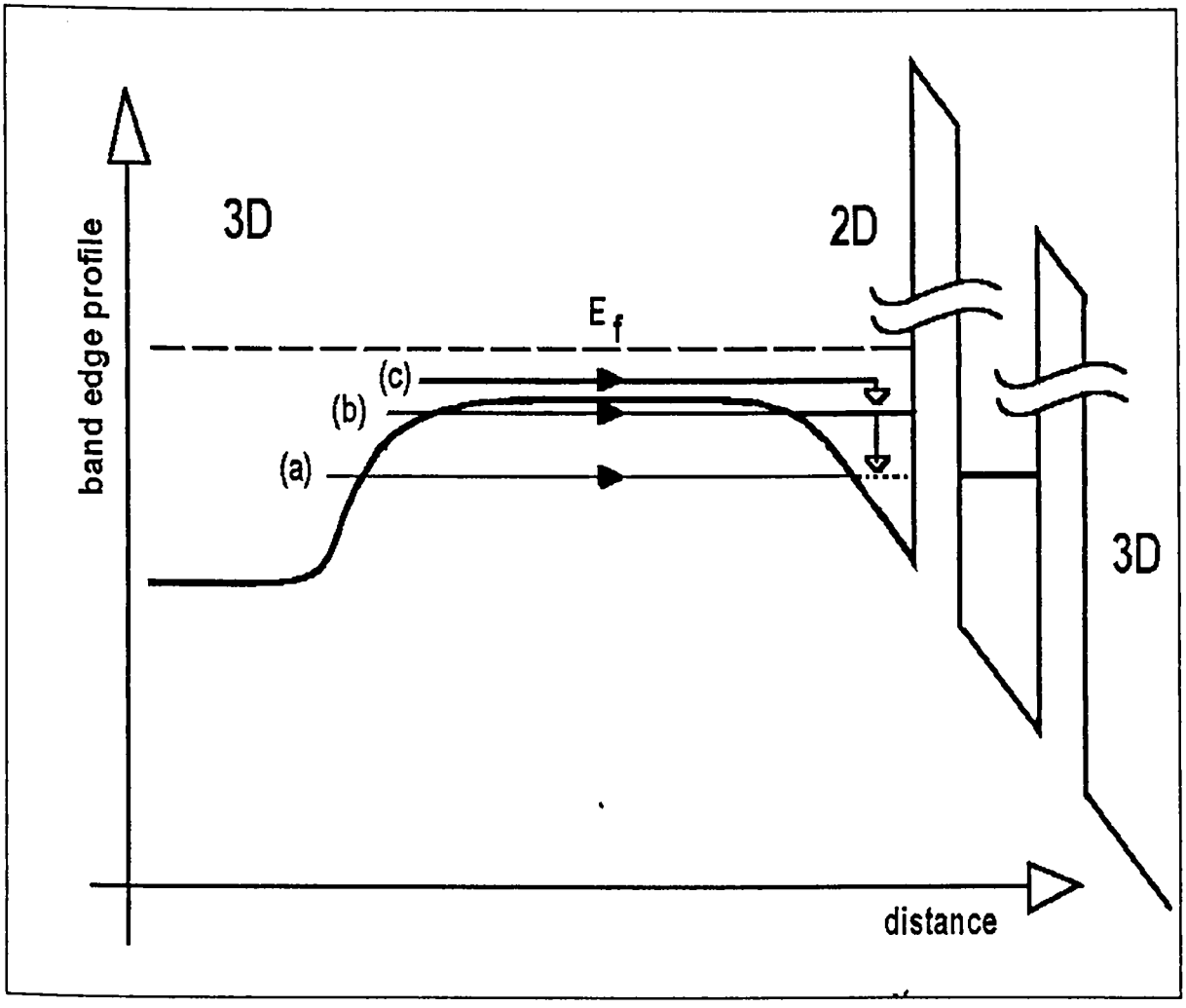


Figure 2.3 the purely coherent 3D/3D approach compared with 2D/3D approach.

In addition, the quantum states in the accumulation layer corresponds to a confinement situation which is intermediate between true bounded states and extended states, under this condition, it is expected that the dispersion relation in the growth direction (noted x in the following) is deviated from the parabolic behaviour. This influences the effective mass within this region. With respect to this issue, we will try to investigate this effect by a renormalisation of the effective mass which is found to be smaller than the 3D-bulk value. However, let us mention that, it is often problematic to define a local effective mass everywhere within the structure in particular when the electron energy is below the conduction band edge and hence within the forbidden gap. In the previous 3D/3D analyses, it is often assumed that, the conduction band effective mass governs any charge transfer. With such assumption, the value of the renormalised effective mass has a negligible influence on the current value and notably the valley current. This is not the case if the 2D/3D approach is applied.

2.4 Quasi-2D local density of states :

In the introduction of this chapter, we have showed the limitations of a purely coherent 3D/3D approach. This concerns the voltage range which is found broader in the experiment for NDC effect and the valley current which is under estimated. This motivates the need of a more realistic physically based refined model for which less discrepancy could be found between the predicted results and the experimental data. To achieve this goal, it is believed that, an almost exact value of the local density of states in the accumulation layer has to be defined. On the other hand the 3D/3D supply function model (Eq. 2-13) and its physical interpretation has to be re-examined (later on in this chapter) and replaced by a 2D/3D supply function.

First of all we shall start by the model proposed by Lassning et. Al [10]. In this model, the local density of states was assumed directly proportional to the quantum probability of existence.

2.4.1 Lassning model [10] :

In an attempt to study the influence of structural parameters on the purely coherent electron life time in a quasi-bound state, R.Lassning and W.Boxleitner had used the semi-opened structure depicted in Figure 2.4.

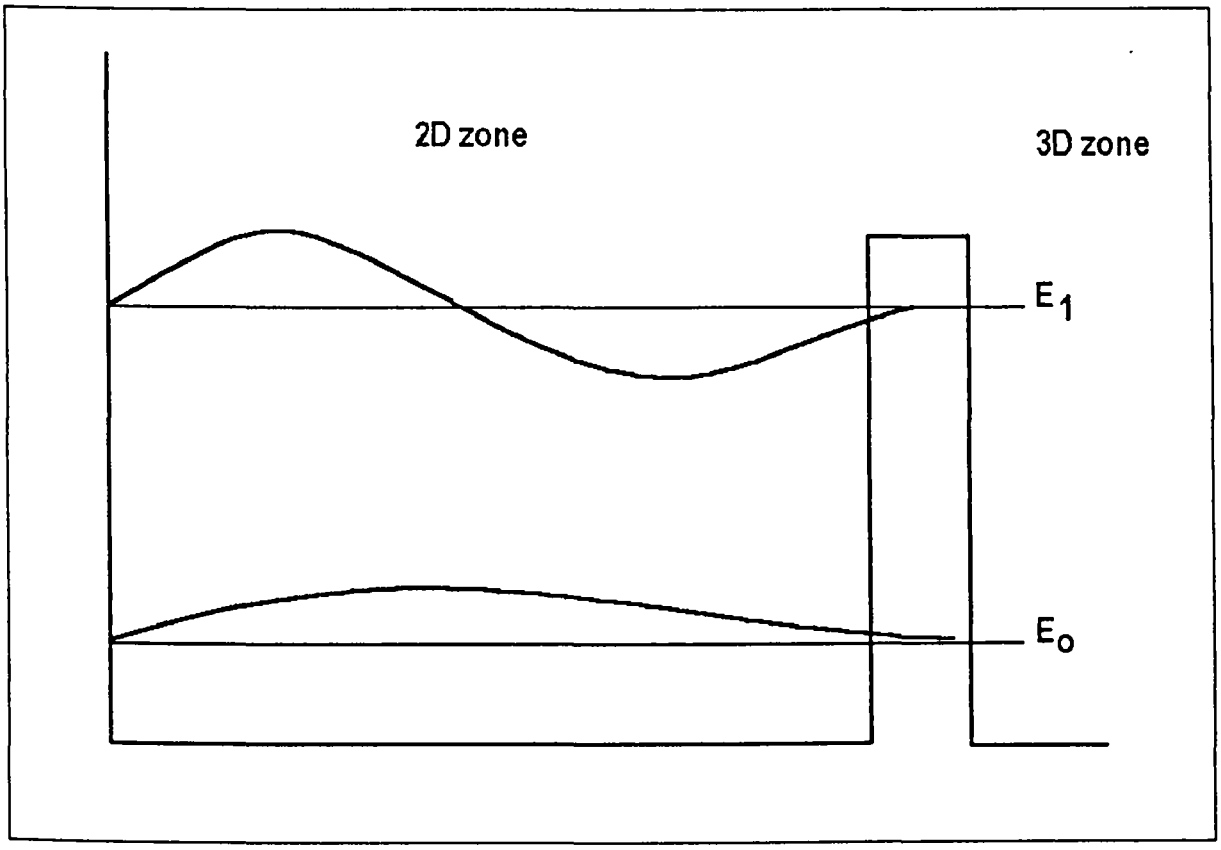


Figure 2.4 schematic for the model structure used for the tunnelling study of Lassning et. All [10].

The main assumption of the analysis published in reference [10] is the use of a flat potential profile. As a consequence, a sinusoidal wave function is the solution of the Schrodinger wave equation in the propagation zones. Also the attenuation of the wave function within the barrier can be described by means of a very simple exponential dependence. On the other hand, the amplitude of the wave function in the 3D zone was chosen to be constant independent neither of energy nor of position. At last, let us

mention that only the relative values were considered, which were sufficient to derive the quasi-bound state life time, but failed to address the current voltage issue. The local density of states in their model were then defined as usual to be directly proportional to the probability of existence spatially averaged in the quasi-2D region.

2.4.2 Analytical calculation of local density of states using zero field approximation :

Worth noting that based on the WKB approximation, the wave function could be analytically written in the regions of classically allowed energy bands as,

$$\psi(x)=\frac{A}{\sqrt{k(x)}}\ \sin\left\{\left[\int k(x)\,dx\right]+\delta\right\} \qquad \textit{Eq. 2-14}$$

whereas in the regions of classically forbidden energy bands as,

$$\psi(x)=\frac{1}{\sqrt{\alpha(x)}}\left\{A_1\exp\left[-\int\alpha(x)\,dx\right]+A_2\exp\left[\int\alpha(x)\,dx\right]\right\} \qquad \textit{Eq. 2-15}$$

where :

- A_1 & A_2 could be assumed for a first order approximation as constants in each region.
- $k(x)=\sqrt{2m^*\varepsilon_x(x)/\hbar^2}$
- $\alpha(x)=\sqrt{2m^*[v_b-\varepsilon_x(x)]/\hbar^2}$
- m^* is the electron effective mass, ε_x is the energy in the propagation direction and $\hbar$ is the Blanck's constant.

In a practical device, we face the problem of accounting for the variation of the potential profile relatively far from the equilibrium at the resonance condition. For the sake of simplicity we well assume that only the propagation k(x) and attenuation $\alpha(x)$

exhibit a spatial dependence through ε_x, the energy in the growth direction, in the sinusoidal and exponential terms. It is believed that, keeping a constant amplitude does not modify too much the resonance condition. Such an assumption simplifies drastically the analytical procedure.

LAYER	MATERIAL	DOPING (cm^{-3}) FABRICATED, SIMULATED	THICKNESS (NM) FABRICATED, SIMULATED
cap	GaAs	$1*10^{19}$, $1*10^{18}$	50, 50
upper electrode	GaAs	$5*10^{17}$, $5*10^{17}$	270, 50
upper spacer	GaAs	n.i.d., n.i.d.	1.4, 1.4
upper barrier	AlAs	n.i.d., n.i.d.	4.0, 4.0
well	GaAs	n.i.d., n.i.d.	6.0, 6.0
lower barrier	AlAs	n.i.d., n.i.d.	4.0, 4.0
lower spacer	GaAs	n.i.d., n.i.d.	5.0, 5.0
lower electrode	GaAs	$5*10^{17}$, $5*10^{17}$	300, 50
substrate	GaAs	n^+, $1*10^{18}$	--, 50

Table 2-1 table(1) double barrier growth structure [11]

of the resonant tunnelling device used as a testing vehicle in this chapter.

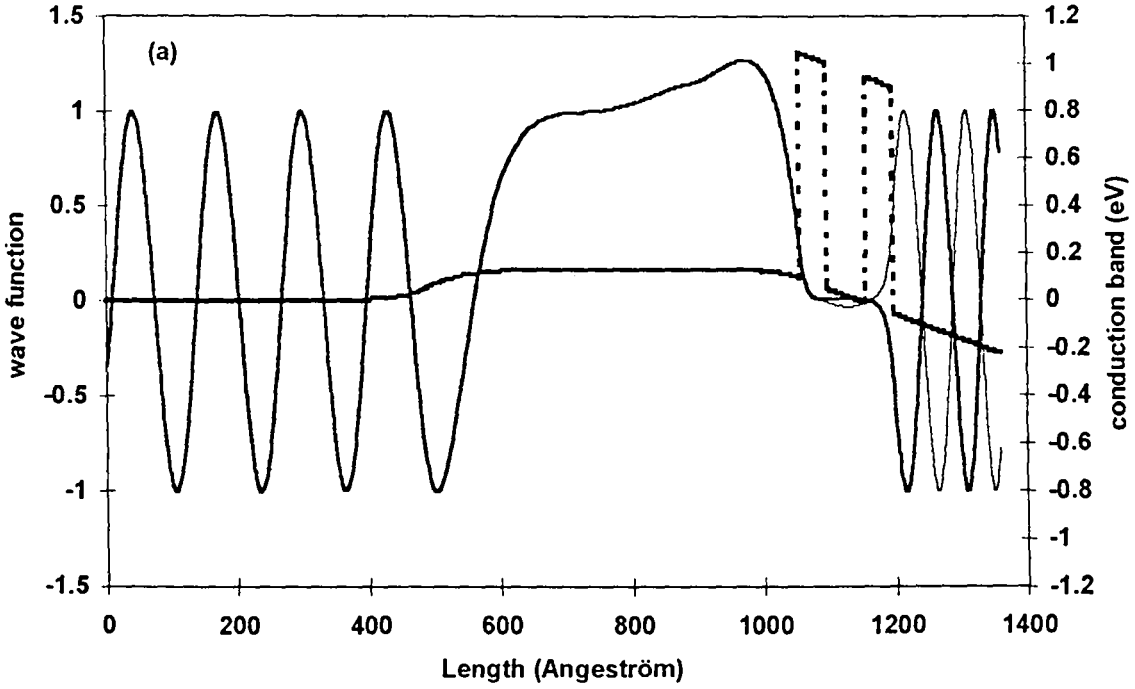
Table 2-1 illustrates the structural parameters of the resonant tunnelling diode which was fabricated by T. Waho et al [11] and which is used as the test vehicle throughout the study in this chapter. On the other hand, Figure 2.5 depicts the conduction band profile of this structure, which is calculated using the Thomas-Fermi approximation [13]. From the potential profile and for energy values corresponding to path a (see Figure 2.3) six changes between the allowed and forbidden regions have to be defined concerning the continuity of the wave function and its first derivative. This

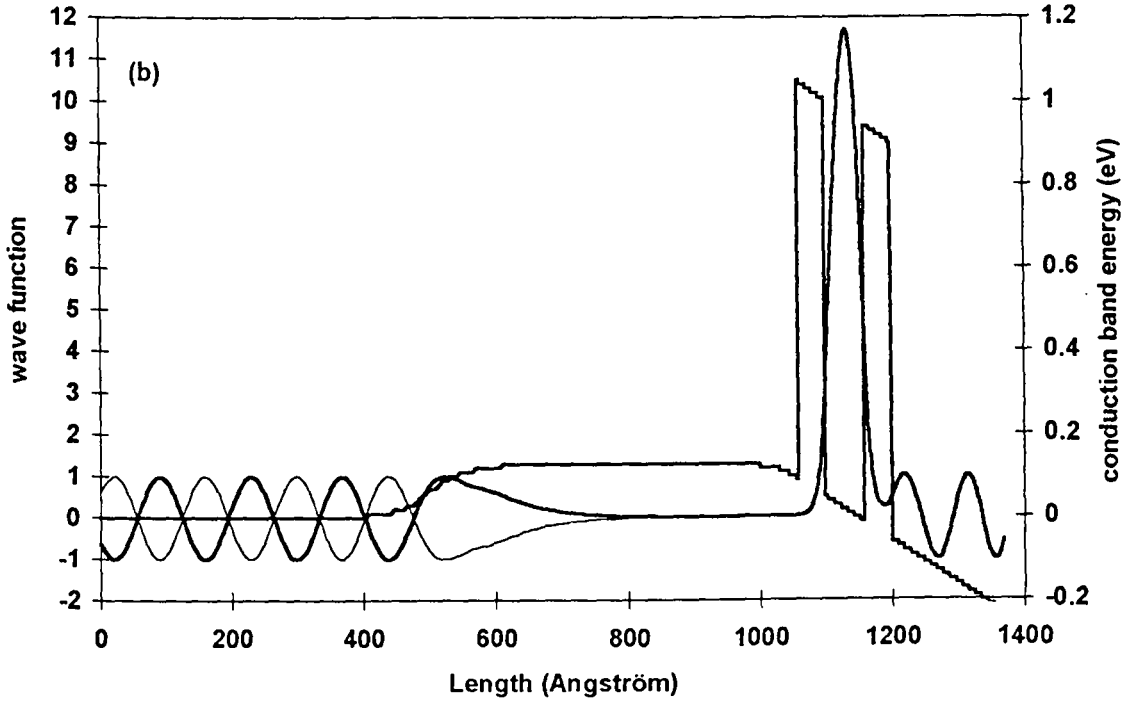
results in twelve analytical equations with twelve unknowns. While elaborating this analytical approach, the unity amplitude of the sinusoidal wave function approximation is chosen to normalise the wave function in both the emitter and the collector 3D-zones. In addition, the phase is left as a floating unknown variable to be determined by the analytical solution at each energy level.

Using this analytical procedure, the variation of the wave function versus distance can be calculated for each value of ε_x, resulting in a wave function spectroscopy of the microstructure.

It is important to state that we find out two phase values satisfying the unity normalisation condition in both the emitter and collector 3D-regions.

Figure 2.5 shows a plot of these two solutions for an applied voltage of 0.3 Volt. When the energy is close to resonance condition within the accumulation layer (a). In (b) the energy is close to the resonance within the quantum well of the DBH. In (c) we plotted the solution at the DBH resonant energy.


(a)
wave function
conduction band (eV)
Length (Angeström)
1.5
1
0.5
0
-0.5
-1
-1.5
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
0
200
400
600
800
1000
1200
1400


(b)
wave function
conduction band energy (eV)
Length (Angström)
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
200
400
600
800
1000
1200
1400

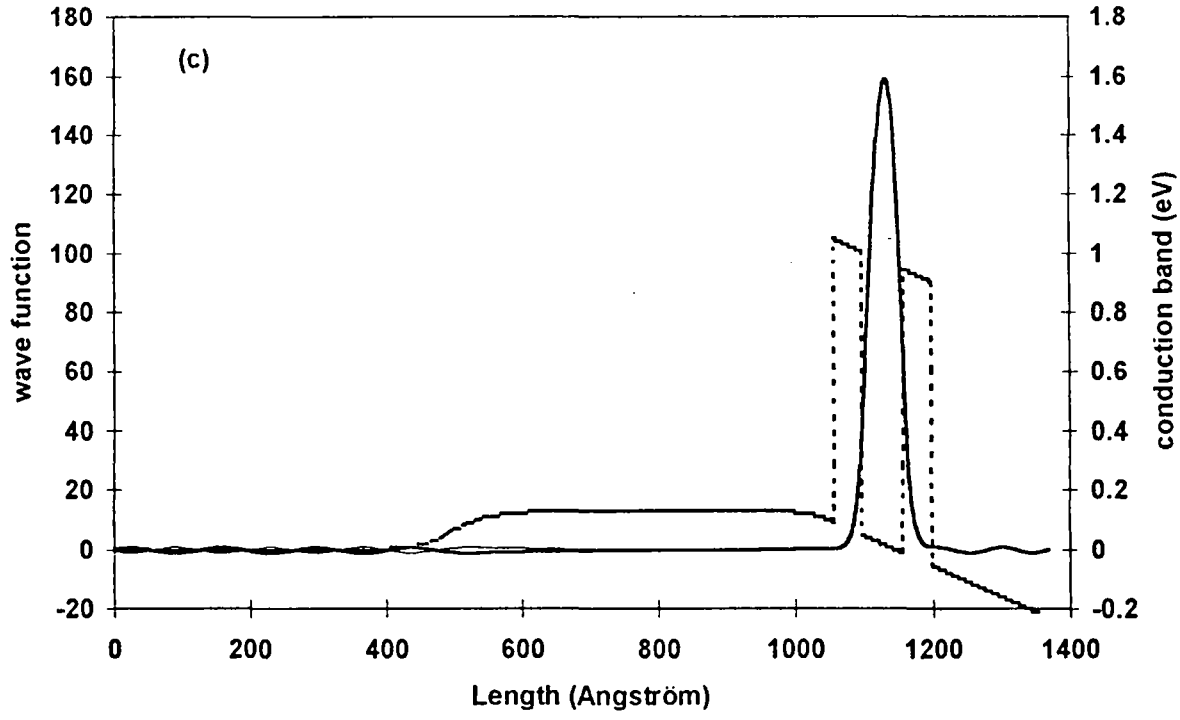


Figure 2.5 analytical wave function for zero-field approximation and unity amplitude in both emitter and collector (a) at quasi-resonance energy in the accumulation zone, (b). at the neighbourhood of the quasi-resonance energy in the quantum-well, © at quasi-resonance energy in the quantum-well

For illustrating the calculations of density of states we pay a special attention to the bias condition corresponding to ε_{acc} lying at an higher energy than ε_o. Let us recall that in a 3D/3D approach it is normally believed that the current is turned off because of the anti-crossing between ε_{acc} and ε_o.

The resulting local density of states as a function of energy in the quantum-well is depicted in Figure 2.6a.

We use for plotting this local density of states a normalisation procedure which will be explained here after. The key feature in this figure, is plotted in logarithmic scale, is the high degree of selectivity observed at the resonance of the density of states which is found here around 25 meV. This is a characteristic of a high degree of quantum

resonant confinement. On the other hand, Figure 2.6b depicts the analytical solution of the local density of states as a function of energy in the accumulation zone. In this figure, corresponding to the same applied voltage, two features are now seen at the previous energy value of 25 meV and at 40 meV with a very weak resonant structure.

This is an original result which is to our knowledge not mentioned previously in the literature. Indeed, although we investigate the eigenfunctions in the accumulation layer, we find a resonant structure characteristic of the quantum well level ε_o. In the following we will speak about an effect of signature due to the coupling between the accumulation and the central well of the DBH. This peak is sharp and well defined. From the point of view of current properties it is important to note that the existence of that local density of states, not predicted in a simple analysis, permits one to consider an extra energy level of injection. The second weak confined peak reveals the own resonance of the accumulation zone. Quantitatively, the former resonant structure (quantum well signature) exhibits peak values quite comparable t the level reached in the extended states corresponding to a 3D-local density of states.

The analytical approach used above relies on quite crude assumption. Nevertheless, some novel trends have been found in the understanding of resonant tunnelling through DBH. At this stage, it seems interesting to re-examine the validity of this highly approximated analytical approach and also to establish in what extent the notion of local density of states can be generalised.

To achieve this goal, one has to study the equivalence between 2D & 3D density of states which could be derived from the above calculated local density of states.

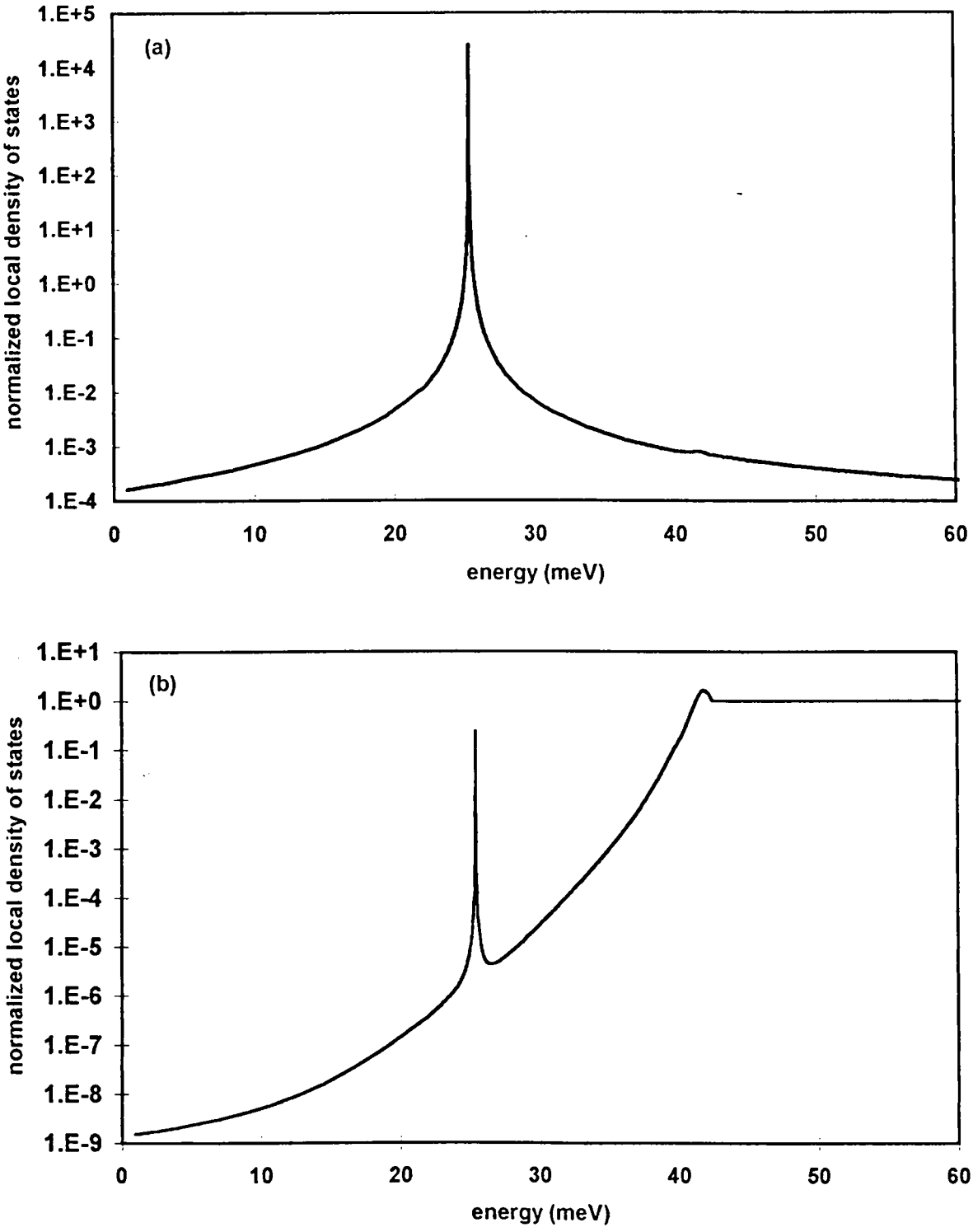


Figure 2.6 analytical solution of the local density of states as function of energy (a) *in the quantum-well, (b) in the accumulation zone.*

2.4.3 Equivalence between 2D & 3D density of states :

It is now well known [14][15] that the 2D volume density of states calculated for a layer of equal thickness to that of an infinite depth quantum well compared to the 3D volume density of states of an unbounded system, are equal for all energy eigen-values (ε_n) (Figure 2.7). In particular, this result was pointed out by Dingle [14][15] and in the following we will speak about the Dingle equality.

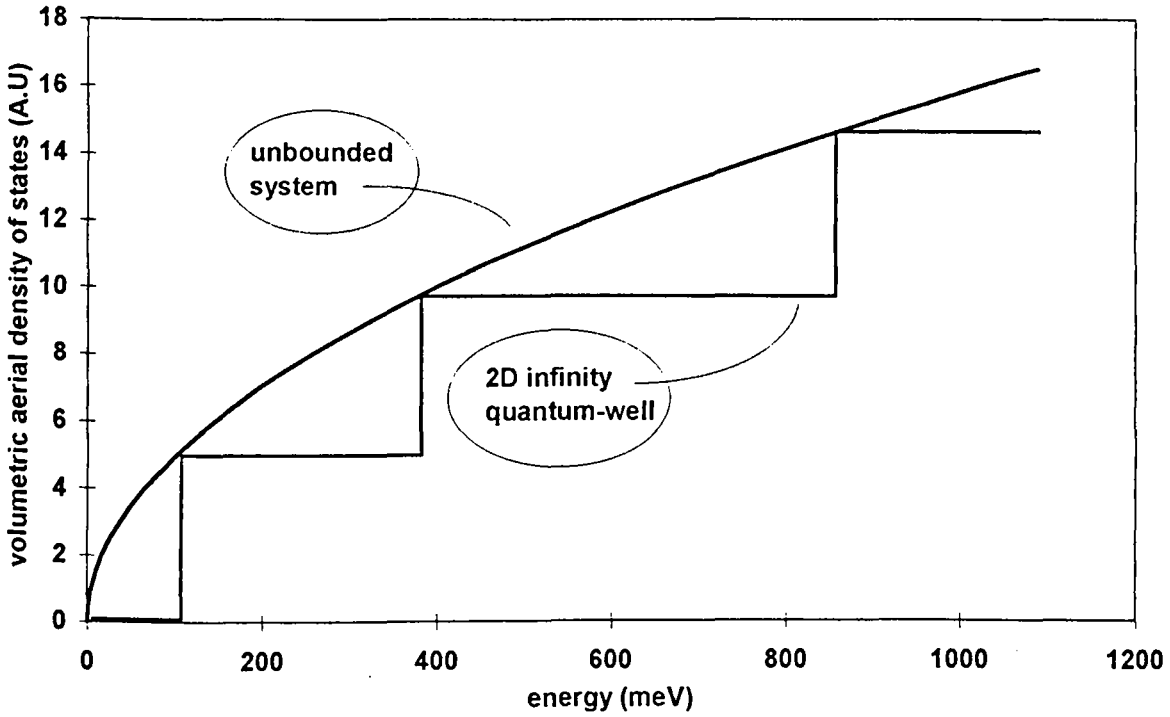


Figure 2.7 comparison between volume density of states in both unbounded system and 2D infinity quantum well after Dingle [14]

Let us rewrite Eq. 2-10 which summarises the mathematical definition of the local density of states with respect to the momentum-free conventional density of states denoted by the capital letter G.

$$g = \frac{1}{2}\frac{d}{dp}G \qquad \text{Eq. 2-16}$$

Integrating Eq. 2-16 and multiplying the result by a factor of two to take into account the momentum in both negative and positive directions we get,

$$G = 2\int g \; dp = \sqrt{2m}\int \frac{g}{\sqrt{\varepsilon}} \; d\varepsilon \qquad \textit{Eq. 2-17}$$

where

- g is the local density of states.
- $p = \sqrt{2\varepsilon / m^*}$ is the quantity of motion.

Substituting by the local density of states in the 3D-unbounded system Eq. 2-8 into Eq. 2-17 we get the well known expression for G_{3D} :

$$G_{3D} = 2\int g_{3D} \; dp = \frac{2\pi m^*}{(2\pi\hbar)^3}\sqrt{2m^*}.\int \frac{d\varepsilon}{\sqrt{\varepsilon}} = 2.\frac{2\pi m^*}{(2\pi\hbar)^3}.\sqrt{2m^*}.\sqrt{\varepsilon} \qquad \textit{Eq. 2-18}$$

Starting from this definition of the local density of states, it is possible to apply it within the DBH quantum well and within the accumulation layer respectively. Figure 2.8 shows the variation of density of states in the quantum well noted in capital G versus energy.

First one can note the practically perfect step like behaviour of the 2D density of states calculated for the present structure with respect to the complete bound situation. This is a direct consequence of the high degree of confinement previously noted in the above section (see Figure 2.6a). under this condition, the Dingle equality is verified, which is a good check for the model.

Figure 2.8 shows the density of states calculated in the accumulation layer. The normalised local density of states (g/g_{3D}) plotted here, exhibits a peak around 40 meV and a signature around 25 meV. Also shown, is the variation of (G_{acc}/G_{2D}) calculated by integrating g with respect to the momentum averaged over the accumulation zone. At last, for comparison we plotted the equivalent (G_{3D}/G_{2D}).

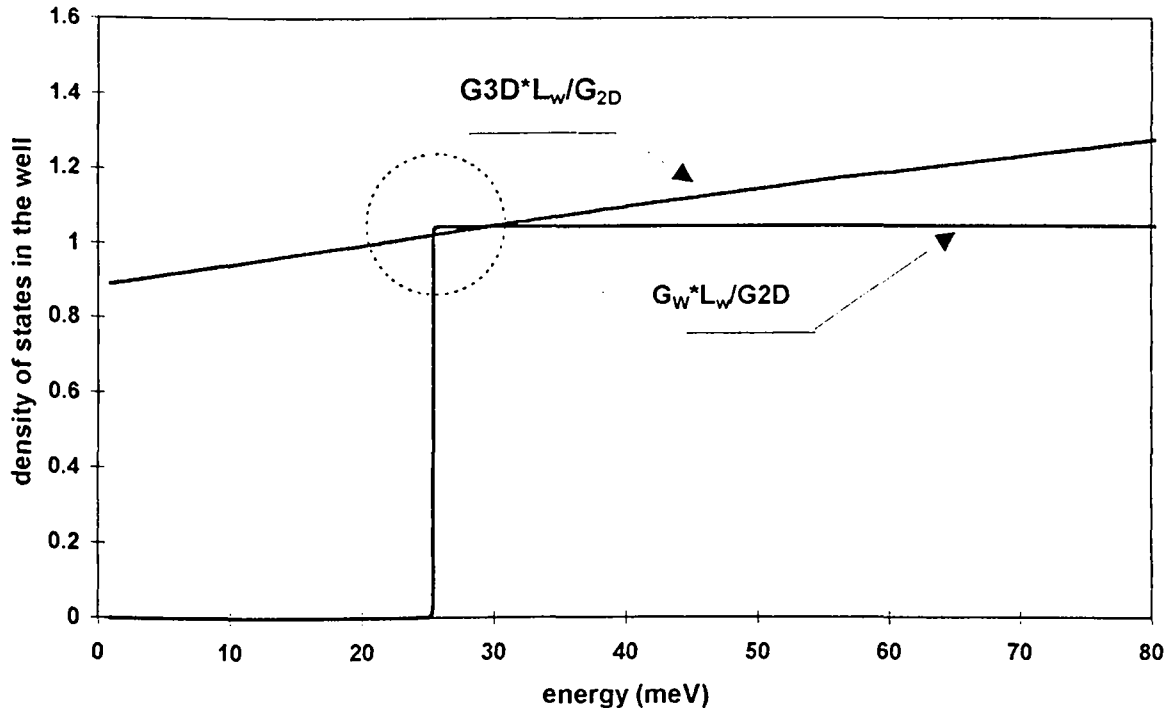


Figure 2.8 the equivalence between the quantum-well density of states and the corresponding 3D density of states (Dingle [14][15] equality is verified at resonance).

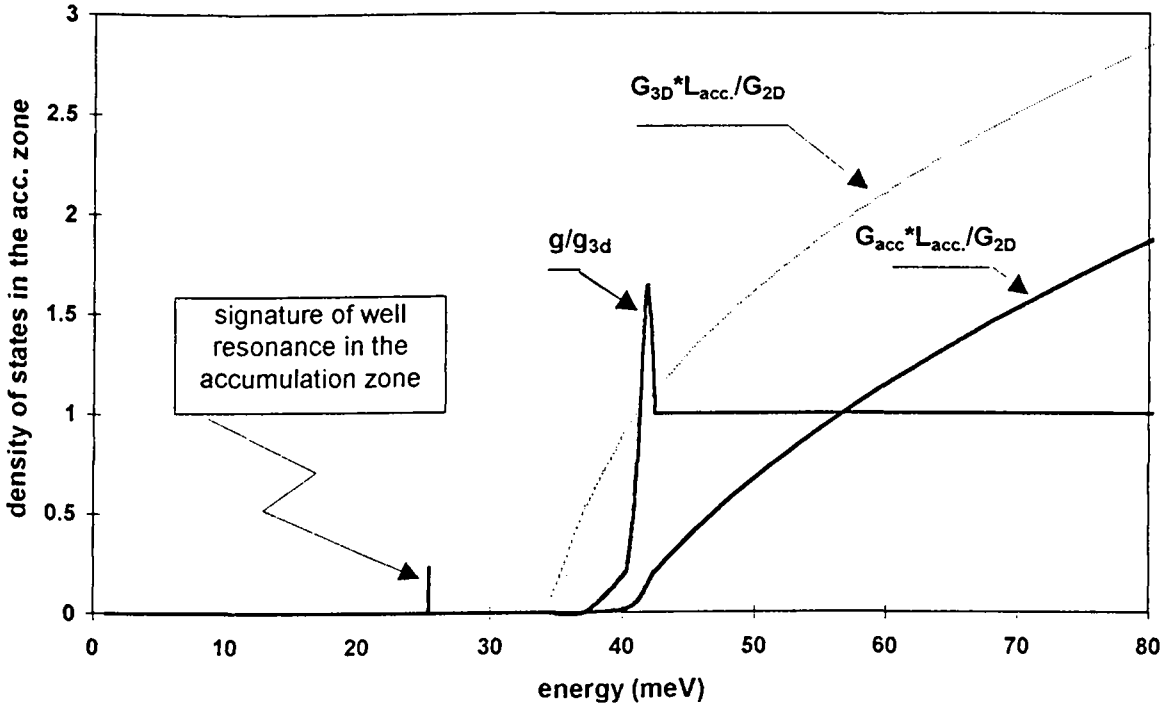


Figure 2.9 the equivalence between the accumulation zone density of states and the corresponding 3-D density of states (Dingle [14][15] equality is not verified at resonance).

For both later terms, we define L_{acc} which can be interpreted as the width of the accumulation layer calculated at the energy $\varepsilon=\varepsilon_o$. The salient feature in this figure, is the loss of the step like characteristic on one hand, and the large discrepancy between the normalised values of G_{3D} and G_{acc}. This is a direct consequence of the openness of the system which in the present case, is intermediate between a highly confined system and extended states.

At this stage of a very simple analytical approach, it is interesting to further investigate the discrepancy between the density of states in the accumulation layer and its corresponding 3D-density of states. We thus developed a more realistic model avoiding in particular the zero field assumption.

2.5 Exact model of the local density of states :

Let us recall that physically the local density of states is not only the measure of finding an electron in a given energy state and in a given position, but also we have to select from the whole k-distribution the states having the current direction. This means that it is not correct to solely consider the local density of states as a function of the probability of existence. Obviously this probability of existence reveals potential carrier accumulation, the best example is the case of a strict bounded state, but does not describe how are distributed the k-vectors in specific location. Therefore, a first conclusion is that, it is imperative to go further beyond the definition of Lassning. This is to account for both features related to the carrier density and to the k-vector. Clearly we have thus to start from the current definition which leads to the basic idea of this work.

The idea of defining a local density of states g is to assume that an electron stream is able to activate and probe the local density of states. Also in other words, this means that g depends not only of the probability of existence which can be related to a carrier density, but also of the momentum $\hbar\, k(x)$ and hence of the velocity.

Now, we try to prove that the local density of states (g) is defined to be linearly proportional to the maximally allowed current density. The later, locally could cross

the material in the energy and momentum ranges considered (dJ_{max}). The above new definition can be directly deduced as follows : The differential current density (dJ) is statistically described by the following equation [12],

$$dJ = 2(p_x/m^*)g(p_x)f(\varepsilon)dp_x d\varepsilon \qquad \textit{Eq. 2-19}$$

using the substitution of variables, $(p_x/m^*)\cdot dp_x = d\varepsilon_x$, we get,

$$dJ = 2g(\varepsilon_x)f(\varepsilon)d\varepsilon_x d\varepsilon \qquad \textit{Eq. 2-20}$$

To obtain the maximally allowed differential current density, the Fermi-Dirac distribution function $f(\varepsilon)$ is replaced by unity and hence one gets,

$$dJ_{max} = 2g(\varepsilon_x)d\varepsilon_x d\varepsilon \qquad \textit{Eq. 2-21}$$

The question is how to quantum mechanically activate all the allowed local density of states in a real open system as considered in this study. The probability current (dJ_{max}) is quantum mechanically defined by the following equation [12].

$$dJ_{max} = \frac{C_1}{2j}\left|\psi(x,\varepsilon_x)\frac{\partial\psi(x,\varepsilon_x)^*}{\partial x} - \psi(x,\varepsilon_x)^*\frac{\partial\psi(x,\varepsilon_x)}{\partial x}\right|$$

$$= C_1 k(x,\varepsilon_x)\left|\psi(x,\varepsilon_x)\right|^2 \qquad \textit{Eq. 2-22}$$

where C_1 is constant, Ψ^* stands for the conjugate complex of the wave function Ψ, and $k(x,\varepsilon_x)$ is the propagation coefficient along the growth direction (x) and at the energy (ε_x).

From Eq. 2-21 and Eq. 2-22 the local density of states can be written as,

$$g(x,\varepsilon_x) = C_2 k(x,\varepsilon_x)\ |\psi(x,\varepsilon_x)|^2 \qquad \textit{Eq. 2-23}$$

Worth-noting that, because the quantum system is open, all the conduction states are equally-likely activated by injecting a forward and an equal backward quantum mechanical probability current. Under this condition the local density of states is calculated using the following equation :

$$g(x,\varepsilon_x) = C_2 k(x,\varepsilon_x)\left(|\overrightarrow{\Psi(x,\varepsilon_x)}|^2_{FA} + |\overrightarrow{\Psi(x,\varepsilon_x)}|^2_{BA}\right) \qquad \textit{Eq. 2-24}$$

which is equivalent to,

$$g(x,\varepsilon_x) = C_2 k(x,\varepsilon_x)\left(|\overleftarrow{\Psi(x,\varepsilon_x)}|^2_{FA} + |\overleftarrow{\Psi(x,\varepsilon_x)}|^2_{BA}\right) \qquad \textit{Eq. 2-25}$$

where (FA) stands for a forward activation and (BA) for a backward activation.

Recalling the 3D local density of states Eq. 2-8 and applying Eq. 2-24 or Eq. 2-25 in the 3D zones (collector or emitter region),.the normalisation constant (C_2) can be directly calculated

In order to illustrate the difference afforded by the new definition, we compare in Figure 2.10 the space variation of the local density of states first defined as proportional to the quantum probability of existence, which is the wide spread conventional definition and the results calculated using Eq. 2-24 or Eq. 2-25. For the former the plot of $|\Psi^2|/30$ exhibits a well defined maxima at location away from the first heterointerface. This is explained by the bias condition as seen later close to the peak current voltage which is quite low (0.2 V). Under this condition ε_{acc} is very close to the bump height leading to a probability of existence which is not maximised in the central region of the accumulation zone. One can speak about a strong repelling effect at the heterostructure barrier for this specific case of very low one side bounded state. Within the DBH quantum well, we identify a peak whereas in the emitter and

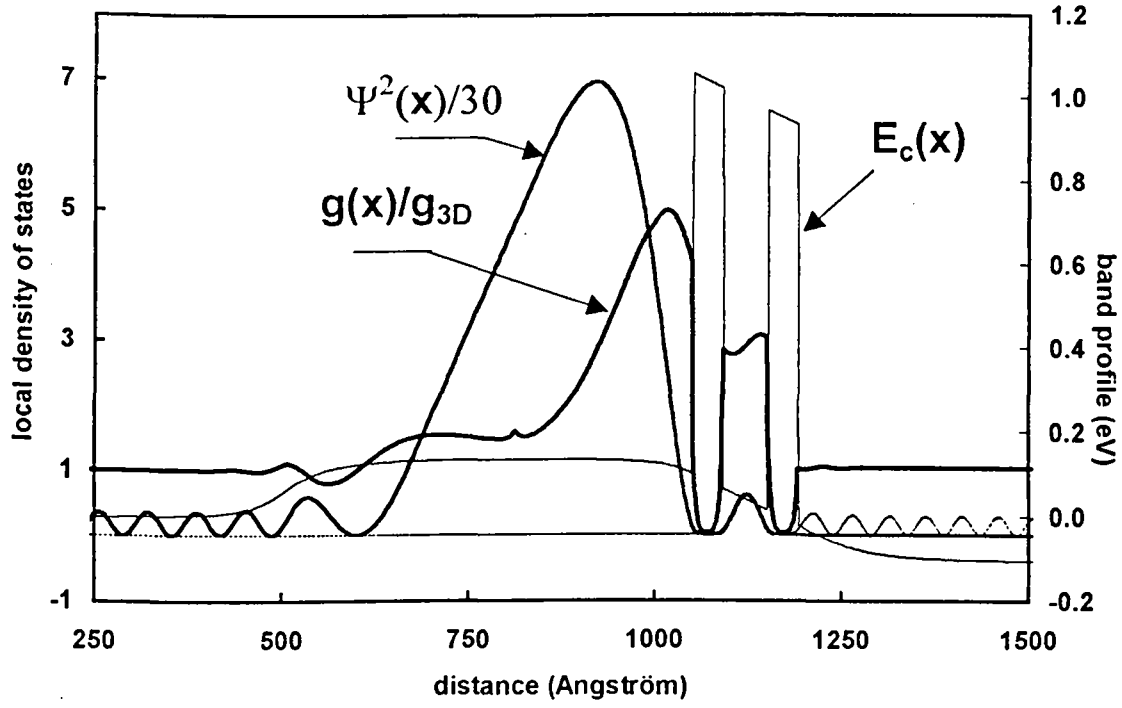
collector regions small ripples can be seen as a consequence of the built-in of a standing wave pattern. In addition a close view of the amplitudes in these adjacent regions reveals amplitude difference.

Turning now to the variation versus distance of g(x), by noticing the calculated values using the new definition and then normalised with respect to g_{3D}, one can find that most of the drawbacks pointed out above can be alleviated. First of all, one can see that the local density of states is now pushed in close proximity to the emitter heterointerface along with rather constant value within the quantum well and also constant value equal to one in the emitter and the collector regions.

The first feature is a direct consequence of the k-contribution. In fact, in the accumulation layer the electrons experience very efficient round-trips which induce a strong local or more exactly regional currents. However, in this bouncing back and forth motion, the net current in the direction of bias is relatively small. In terms of the life time in the quantum state, the electrons spend a relatively long time as compared with the escape tunnelling time. Hence, there is an accumulation of local current which reflects the confinement of the carrier on one hand by the emitter barrier and on the other hand by the electrostatic bump. The same situation is encountered within the DBH quantum well. Again an accumulation in the local density of states is noted. However in that case the local density of states g(x) is found quite uniform over space. This result, a priori in contrast with the general view of the local density of states which should exhibit a peak within the quantum well. This points out the fact that a lack of carriers is compensated by high magnitude of k(x) vector. This compensation effect is quite equivalent to the one observed in the simulation of devices where the current conservation permits to explain an increase in the average velocity in the regions of low carrier concentration density. This conservation principle is not satisfied over the hole structure because, as mentioned before, we have to distinguish between the regional bouncing current and the net current. The later is measurable in the 3D unbounded regions as expected with a constant value. The last remark with respect to the normalisation issue is of major concern. Figure 2.10b is a zooming view of the DBH region at the resonant energy in the quantum well. Concerning the

existence probability, we have considered the forward (solid line) and the backward (dashed line) excitation respectively. Also shown is the normalised local density of states (g/g_{3D}). In summary, the use of the new definition has the following consequences : first, the density of accumulation in the front of the DBH is now centred within the accumulation layer with non negligible value at the heterointerface. Second, a quasi uniform density of states is obtained in the central quantum well and at last a normalisation procedure can be performed from the surrounding regions. Concerning the later issue, it is often thought that the unity probability normalisation condition of $|\Psi^2|$ defined for a close system can be used for an open system. This restricts the application of such a rule to the resonant energy for which most of the wave function is well localised. In contrast, under non resonant condition the openness of the system plays a key role and the reference to the out side regions is mandatory. Generally speaking, we are facing the normalisation issue whenever we have an open system and the example of the accumulation layer is particularly representative of that question. It is worth mentioning that, even in the zero field analytical calculations out lined in section 2.4.2 the current probability is intrinsically used. Indeed, assuming only constant regional amplitude is equivalent to multiplying the amplitude of $|\Psi^2| \alpha^{al}$ $1/k(x)$ by the wave vector $k(x)$.

Last, let us mention that Lassning's model is a very special case of the general definition proposed here, in the sense that assuming the $k(x)$ vector is constant all over the structure suppresses the k-dependence.



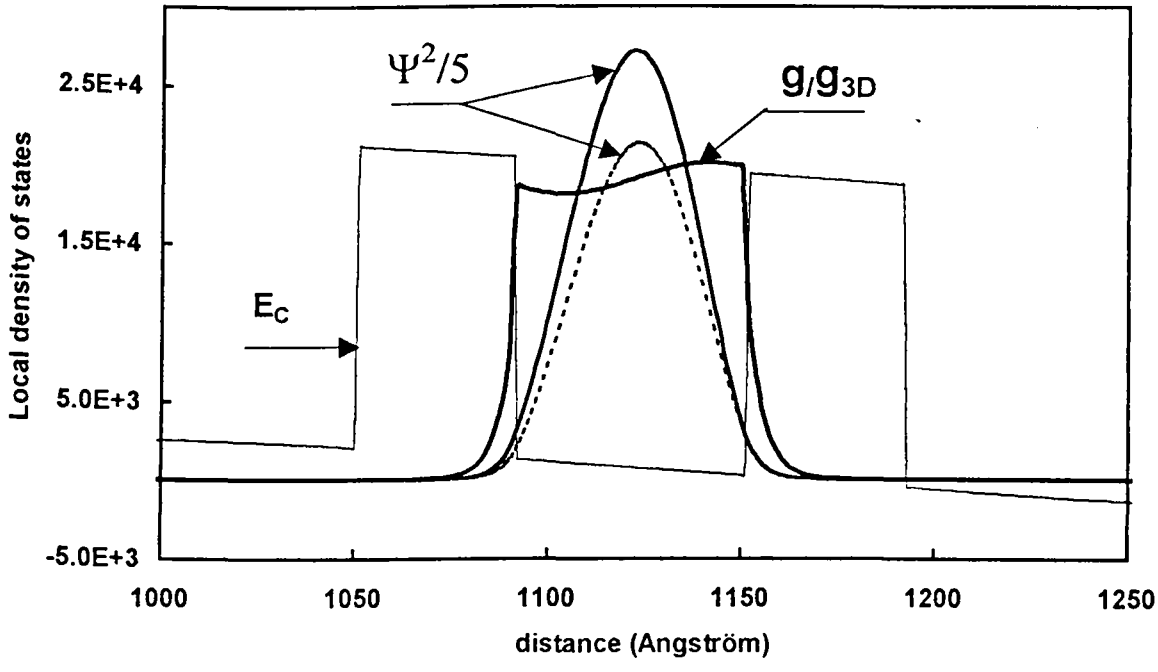


Figure 2.10 difference between the local density of states defined as proportional to quantum probability of existence $|\Psi^2|$ which is the well known definition [10] and our original definition as described by Eq. 2-24 or Eq. 2-25 (a) in the accumulation zone (b) zooming on the quantum well.

Let us now consider the energy dependence of the local density of states which are described for the quantum well region and for the accumulation zone in Figure 2.11 (a) and (b) respectively. For the former, the difference between the previous analytical results show in Figure 2.6a are negligible. In contrast, in the accumulation layer we find out the same trends but with a much more refined evaluation especially close to the resonance related to the quantum well. It can be surprising that at first glance the resonance attached to ε_{acc} is much broadened than that of $\varepsilon_{r\text{-}well}$, this is a direct consequence of energy positioning of ε_{acc} at the frontier between tightly bound system and highly extended state. In terms of life time inversely proportional to the broadening, we obtain a very short dwelling time for ε_{acc} whereas the residence time is long for the signature and for the DBH quantum well itself. Such a very short life time can be interpreted or explained by very efficient escaping processes over the bump. It remains now to reconstruct the conventional momentum-less density of states versus energy and to compare it with the equivalent G_{3D} multiplied by L_{acc}. This comparison is done in Figure 2.12. The main conclusions drawn from the analytical analysis are still verified for the DBH quantum well and the accumulation layer but with a real improvement in terms of quantitative behaviour specially for the later. Now, we are sure that the discrepancy between the G_{3D} and G_{2D} is not a result in the accuracy of the theoretical approach but is a direct consequence of the openness of the system.

In more details regarding the local density of states depicted in Figure 2.11 in the quantum-well and in the accumulation zone, one can notice the well-defined signature of $\varepsilon_{e\text{-}well}$ detected in the accumulation zone.

On the other hand, we also find out that the Dingle [14][15] equality can be verified by integrating the numerically calculated local density of states in the quantum well with respect to the regional average momentum, not to mention to multiply by a factor of two to account for positive and negative momentum. Although, this equality is well verified in the quantum well, a disagreement is seen in the accumulation zone. See also. Figure 2.12.

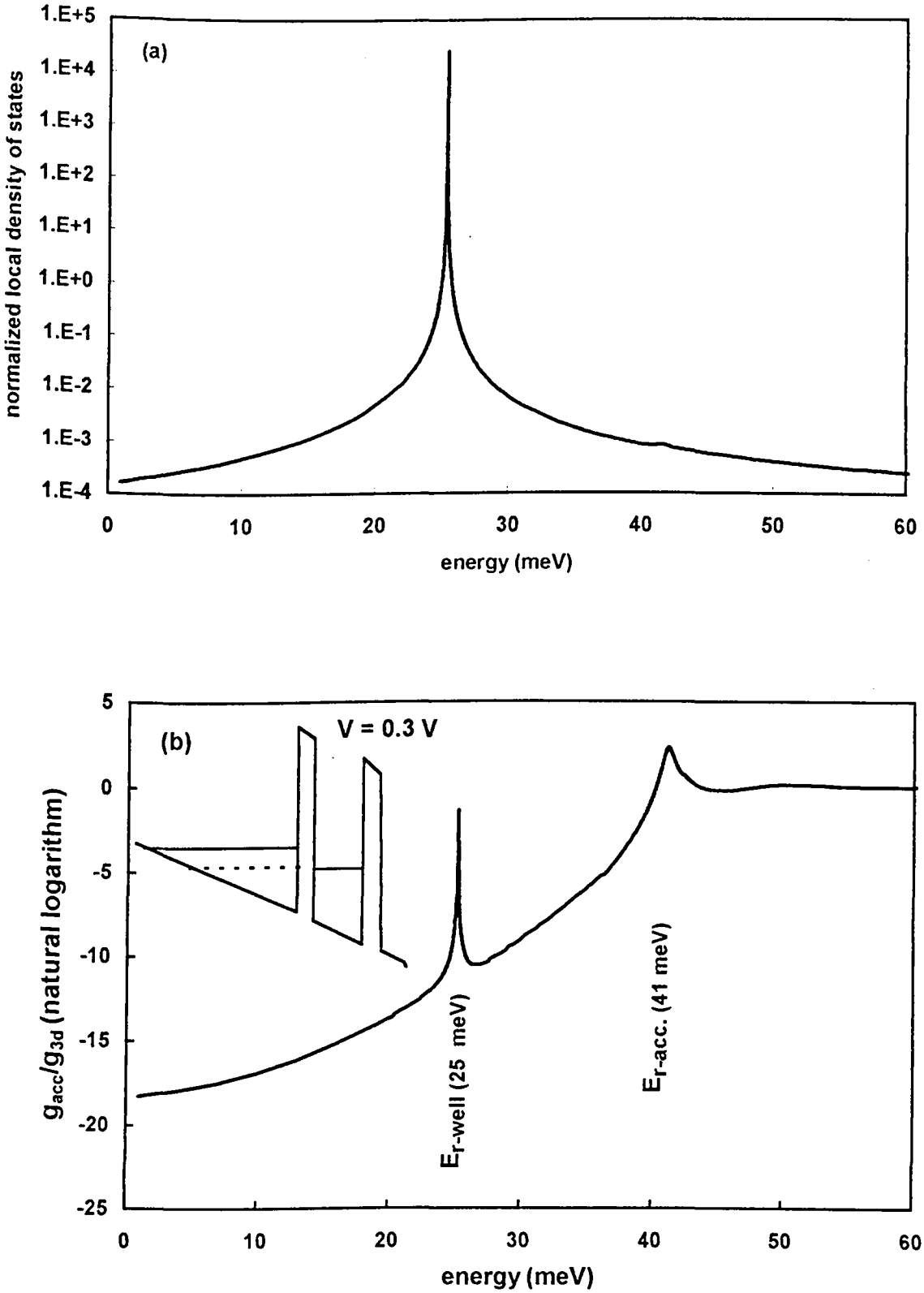


Figure 2.11 depicts the local density of states numerically calculated using the exact model described by Eq. 2-24 or Eq. 2-25 (a) in the quantum-well and (b) in the accumulation zone.

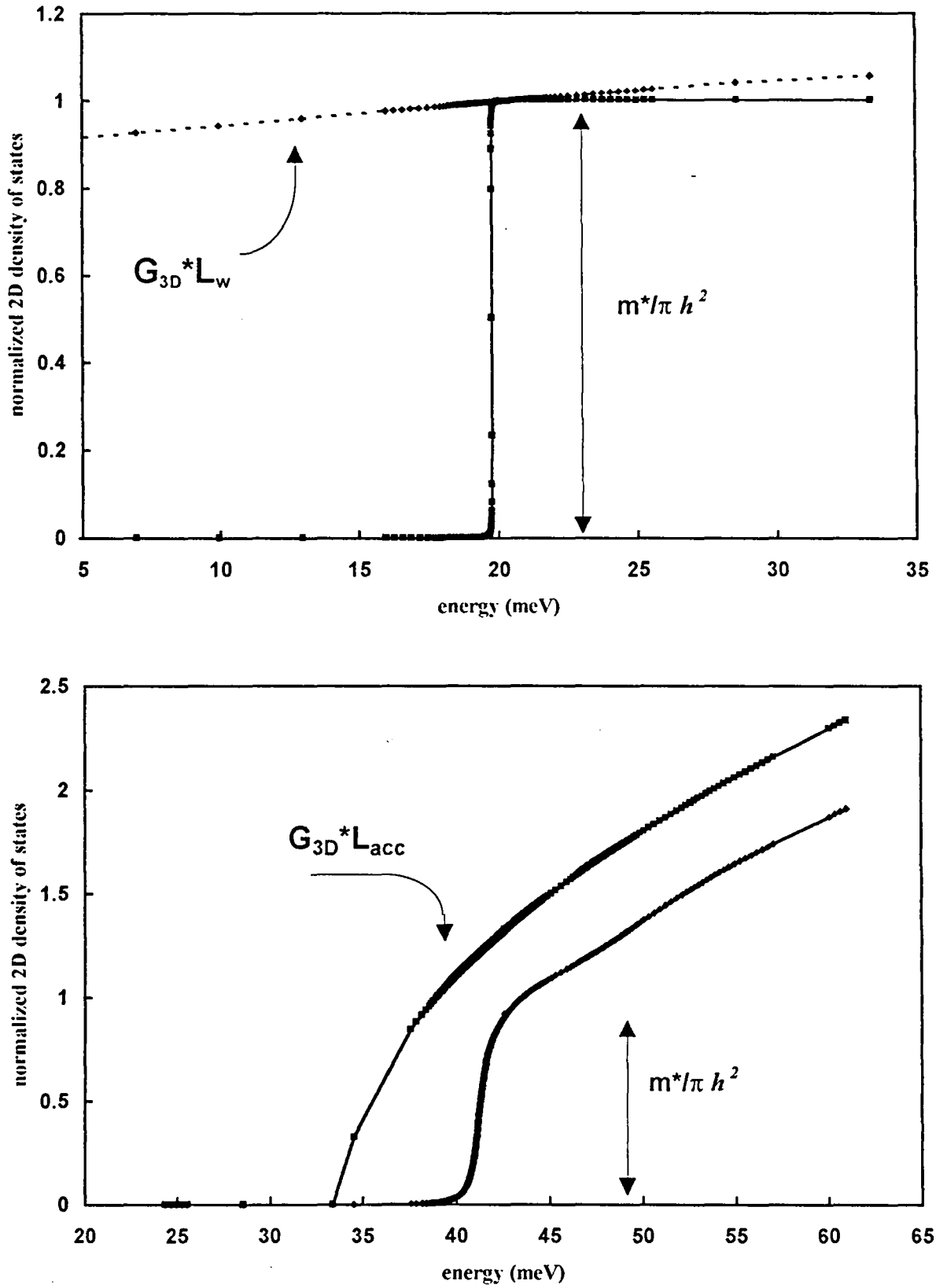


Figure 2.12 the equivalence between the quantum-well density of states and the corresponding 3D density of states

(a) in the quantum well (b) in the accumulation zone.

Comparing Figure 2.8 and Figure 2.12a we find that the zero-field analytical approximation might be a good approximation in the quantum-well region. On the contrary, as regards the local density of states in the accumulation zone, one notice a remarkable big difference between Figure 2.9 and Figure 2.12b. This means that it could not be accepted to approximate the local density of states in the accumulation zone by the zero-field analytical approximation.

Although the discrepancy between the accumulation zone density of states and its equivalent 3D one is dramatically decreased when the exact definition is elaborated, the Dingle equality [14][15] is still not satisfied in this region. In fact the quasi-bound states in the accumulation zone are not sufficiently confined as those in the quantum well, which causes the (k-ε) relationship to be considerably non-parabolic in the accumulation zone. This necessitates the re-normalisation of the effective mass in the accumulation zone. It is to be reported that in our model, the effective mass is re-normalised to achieve the Dingle equality [14][15] in the accumulation zone. The resulting re-normalised effective mass is found to be in the order of 64 % of the 3D effective mass.

Before considering the supply function we would like to comment on the starting energy point of G_{3D} whose variation could alleviate the discrepancy aforementioned. Indeed, basically to calculate the momentum-less density of states requires to integrate g with respect to the average regional momentum with the reference to the same starting energy ε_o for both g_{3D} and g. confidence in this procedure can be found subsequently when $G_{acc}*L_{acc}$ is matching the conventional $G_{2D} = m^* / \pi \hbar^2$.

2.6 Supply function model :

As regards the supply function (SF), let us recall the well known definition proposed in Eq. 2-12, and systematically applied here in the case of 2D/3D injection.

$$SF = K_B T \int_{\varepsilon_x}^{\infty} g_{er} f_e (1 - f_c) - g_{cr} f_c (1 - f_e) \; d\varepsilon \qquad \textit{Eq. 2-26}$$

Denoting g_{er} and g_{cr} to be the relative local density of states on both sides of the double barrier resonant structure.

After expansion of the integrand one gets,

$$\begin{aligned} SF = {} & g_{cr}(\varepsilon_x) \ln\left\{ 1 + \exp[\, 1 + (\varepsilon_{fe} - \varepsilon_x)/kT \,]\right\} \\ & - g_{cr}(\varepsilon_x) \ln\left\{ 1 + \exp[\, 1 + (\varepsilon_{fc} - \varepsilon_x)/kT \,]\right\} \\ & - K_B T \int_{\varepsilon_x}^{\infty} f_c f_e (g_{cr} - g_{er}) d\varepsilon \end{aligned} \qquad \textit{Eq. 2-27}$$

Out of resonance, it is important to note that the local density of states in the injection region (g_{er}) is small as compared to that in the collector region ($g_{cr} = 1$). This gives a negative supply function resulting a negative current value which is not physical. On the contrary, in the energy ranges where the cross resonance occurs an overestimated highly current values might be calculated, depending on the degree of confinement of the injection zone.

In our opinion the unphysical character of the negative and over estimated currents stems from the fact that we considered an electron stream between two subsystems of different density of states (different dimensionality). This not will be the case when g_{er} and g_{cr} are equal and in that case the product of $f_e(1\text{-}f_c)$ and vice versa are justified. In the present work investigating different dimensionality system, this is not the case and we have to re-examine the definition of the supply function. The underlying idea is

to assume that the magnitude of current flow will be limited either by the number of occupied states on one side or the unoccupied states on the other side, depending on the comparison between both. Physically, this limitation concept is based on the saturation limit of the degree of occupation. Therefore, when the number of injected electrons is greater than the available states, one can thought that the electrons undergoes a compete reflection. In that case, the number of occupied states will govern the current. In contrast, the number of injected electrons will be the key figure for controlling the magnitude of current when the available states on the other side is not the limiting factor. Mathematically, we can summarised the above arguments by the following equations.

$$SF(\varepsilon_x) = \int_{\varepsilon_x}^{\infty} \delta SF_{ec}(\varepsilon) - \delta SF_{cc}(\varepsilon) \qquad \textit{Eq. 2-28}$$

where :

$$\delta SF_{ec} = \begin{cases} g_e f_e \delta\varepsilon & \Leftrightarrow\ g_e f_e \le g_c(1-f_c) \\ g_c(1-f_c)\delta\varepsilon & \Leftrightarrow\ g_e f_e \ge g_c(1-f_c) \end{cases} \qquad \textit{Eq. 2-29}$$

Figure 2.13(a) shows the supply function plotted in logarithmic scale using the analytical solution of local density of states reported in Figure 2.5 for a bias voltage of 300 mV. As seen latter, this corresponds to a bias point above the peak voltage in the I-V characteristics. Figure 2.13(b) depicts the results obtained from the numerically calculated density of states shown in Figure 2.10. For the latter, the re-normalised effective mass was introduced. Both curves exhibit the same trends with a well peaked feature reflecting the signature effect as discussed previously. At this stage, the derivation of current density-voltage characteristics is straightforward and is discussed in a next section along with a comparison with experimental data.

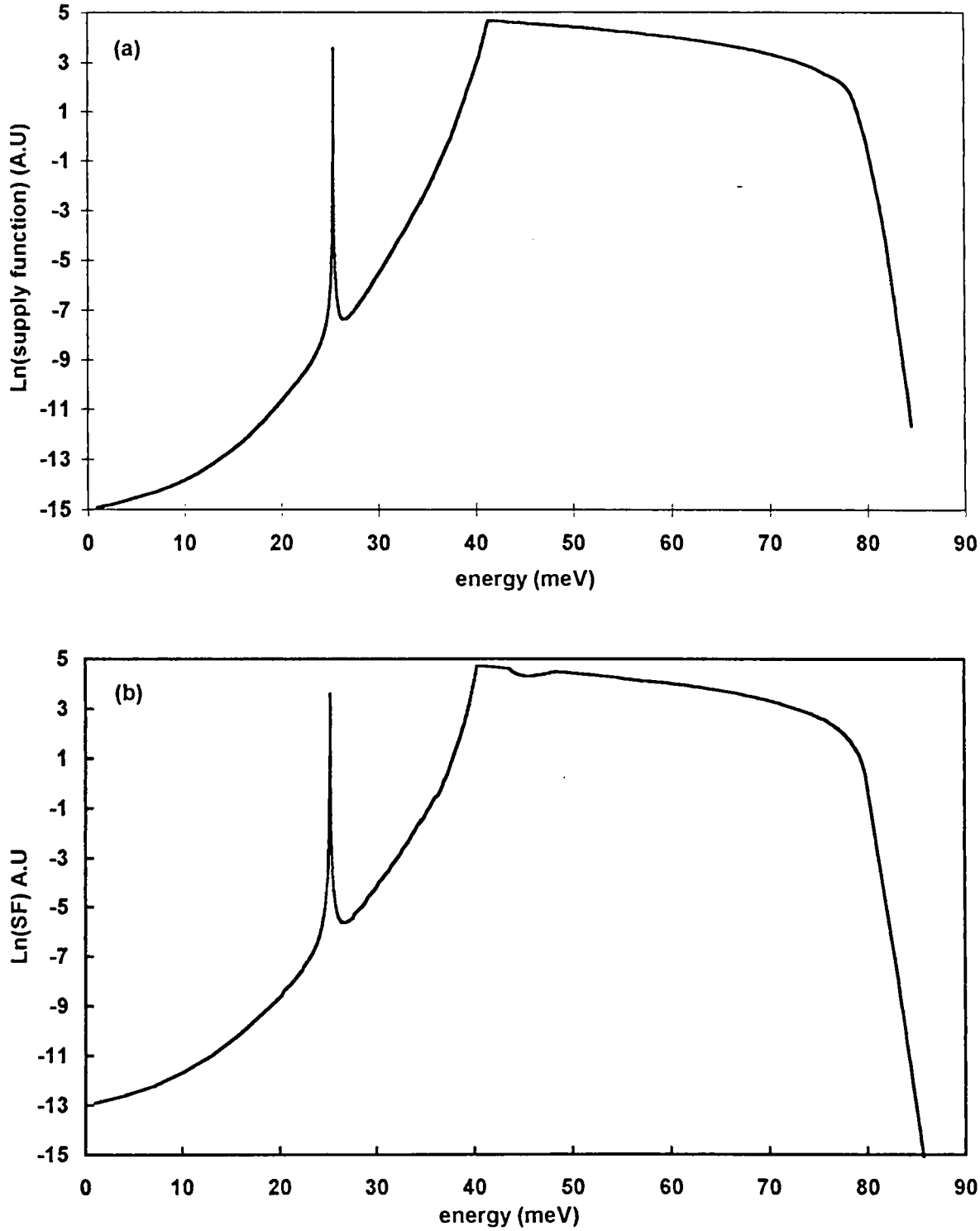


Figure 2.13 supply function calculated using Eq. 2-28 and Eq. 2-29 where it is calculated (a) using the analytical solution of local density of states depicted in Figure 2.6 and (b) the numerically calculated local density of states depicted in .Figure 2.11. It is to be mentioned that the numerically local density of states is calculated using the re-normalised effective mass.

2.7 Comparison to experiment :

In this section we compare the simulation data (solid line) of the current-voltage characteristics calculated at 4K under various assumptions with the experimental results (dashed line) recently published in [11]. Overall, the agreement is good (Figure 2.14), notably concerning the shoulder type behaviour in the negative differential resistance region. Also, the peak current density along with the voltage range over which NDC effect is observed fit quite well. As expected, the peak to valley ratio cannot be described in magnitude according to our model which is based on the assumption of a pure ballistic motion of electrons. In addition the Γ and X valley transfer has been ignored in the present work which can explain an excess valley current experimentally observed in the out of resonance region due to leakage currents via the X tunnelling path. At this stage it remains a major issue to discuss concerning the fact that this distortion of I-V curve observed experimentally results from either an intrinsic effect or an extrinsic one.

Schematically, several distortion-related phenomena can be distinguished for resonant tunnelling devices. Therefore, it was demonstrated over the past that trapping of carriers within the quantum well of a DBH can explain an hysteresis effect [16][17][18][19].

Indeed, at increasing bias voltage one can expect that a large amount of electrons are trapped within the well shifting by a space charge effect the bias voltage to a higher value. In contrast, under decreasing voltage condition, this space charge effect is of minor concern and voltage is pushed down to a lower value. Moreover, this effect is magnified by the asymmetry of the structure [18].

In our case, this phenomena should not play a significant role because the current magnitude and hence the trapping charge is low. Besides, the asymmetry in the potential profile is solely induce by the applied voltage.

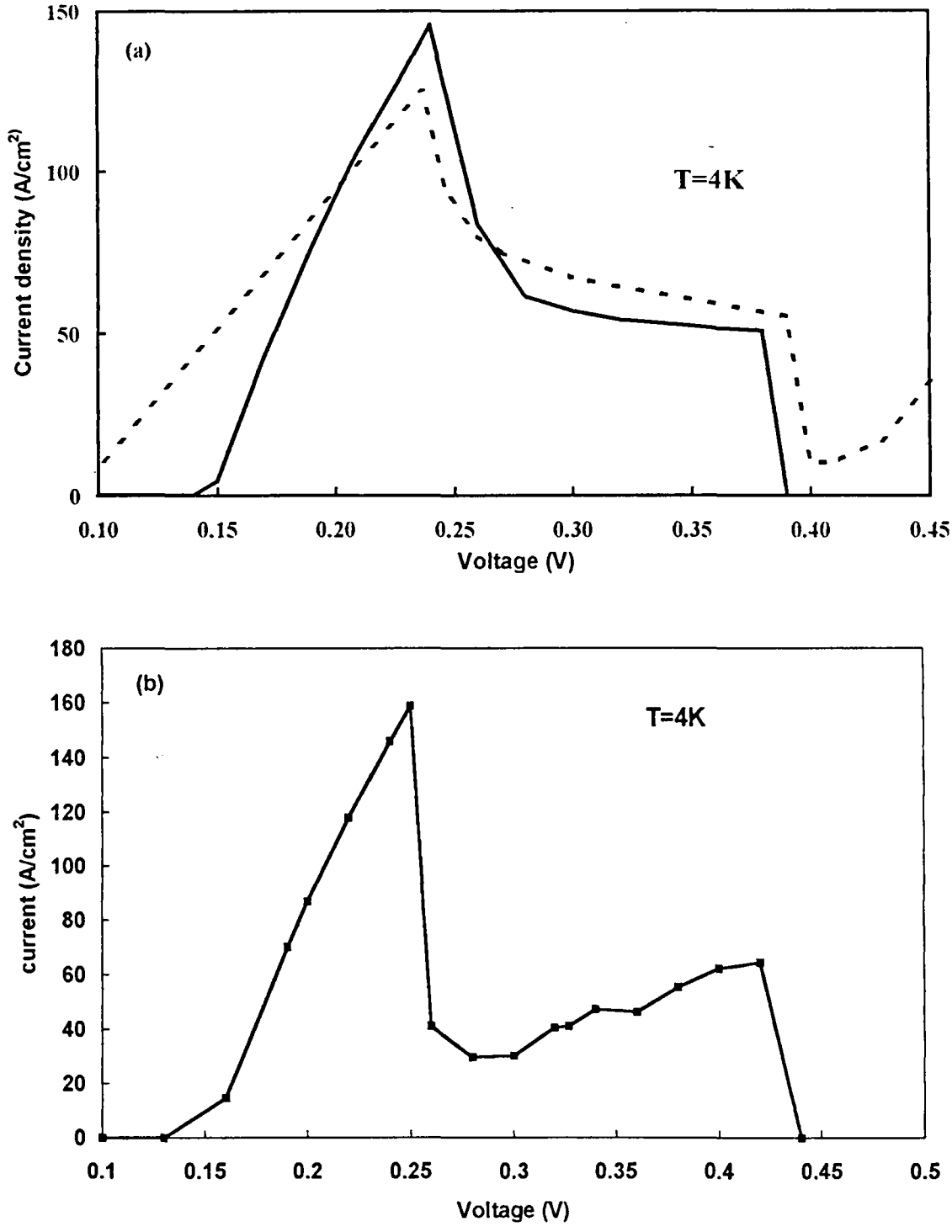


Figure 2.14 I-V simulations (a) numerical model and a comparison with experiment of [11], (b) analytical model without effective mass re-normalisation.

A more general issue which applies to NDC devices concerns the occurrence of self oscillations which are spontaneously developed due to the NDC effect. These oscillations distort the current-voltage characteristics which often exhibits a plateau-like or shoulder-like variation. Graphically, this can be simply understood assuming a single frequency and /or harmonic self oscillations. In Figure 2.15 we plot in dashed line a typical intrinsic I-V characteristics. In the same figure, the relevant distorted I-V curve which will be measured at each DC bias voltage (V_o) is also reported in solid line.

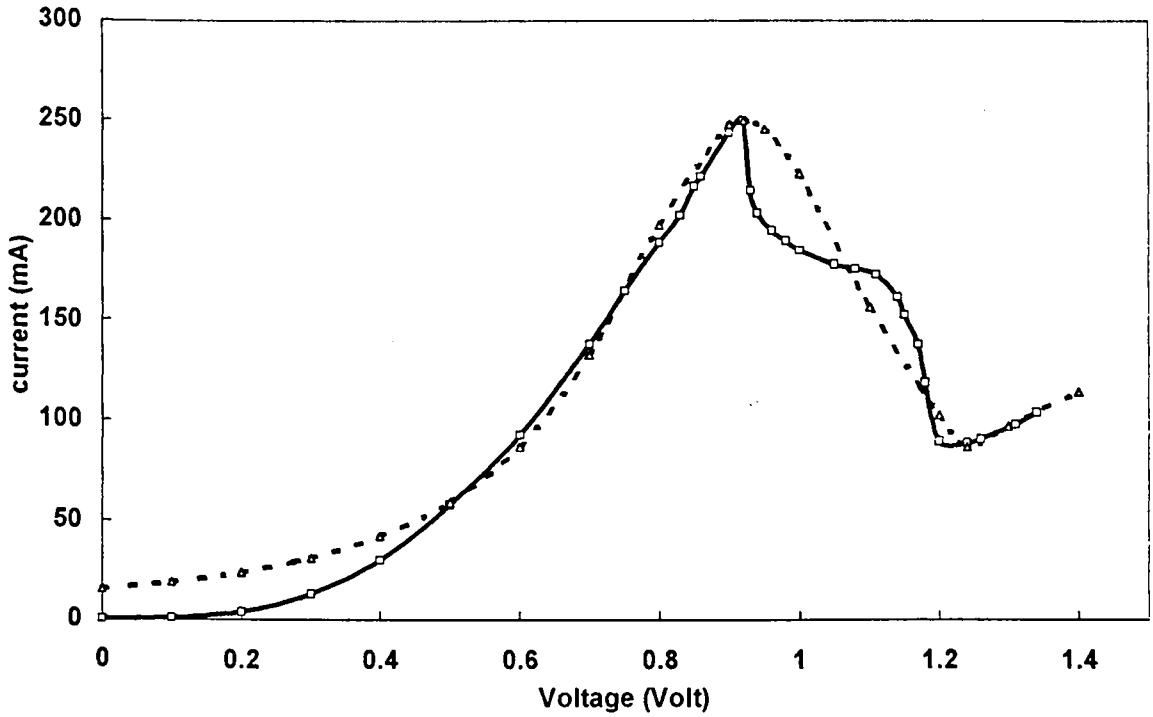


Figure 2.15 a typical intrinsic I-V characteristics (dashed line) and the relevant distorted I-V curve (solid line).

We now assume that an oscillating voltage contribution is superimposed over V_o, resulting from a spurious self oscillation. In this case, we will measure, under statistic condition, the average current value. Simple considerations on the influence of the bias point shows that this time averaged current is lower when the bias voltage is below the mid bias point and vice-versa. Qualitatively, the current variation against voltage shows a shoulder-type distorted form. It is clear that in real devices the parasitic oscillations can be non-harmonic and/or multi-frequency. One can find in the literature

several papers addressing this issue [20][21][22]. More interestingly in practice, is the oscillation condition which define the threshold limit between a stable and an unstable operation [19][24][25]. These stability criteria were established a long time ago, notably by the Bell-laboratory research staff. They working at that time on injection and transit time devices. One of the most simple analysis is based on the derivation of a lumped equivalent circuit (Figure 2.16) which consists of two intrinsic elements namely the NDR (R_d) and the diode capacitance (C_d) and two extrinsic elements. For the latter, this consists in a series resistance (R_s) and a reactive inductance (L_s). The extrinsic elements depicts the influence of the interconnecting section.

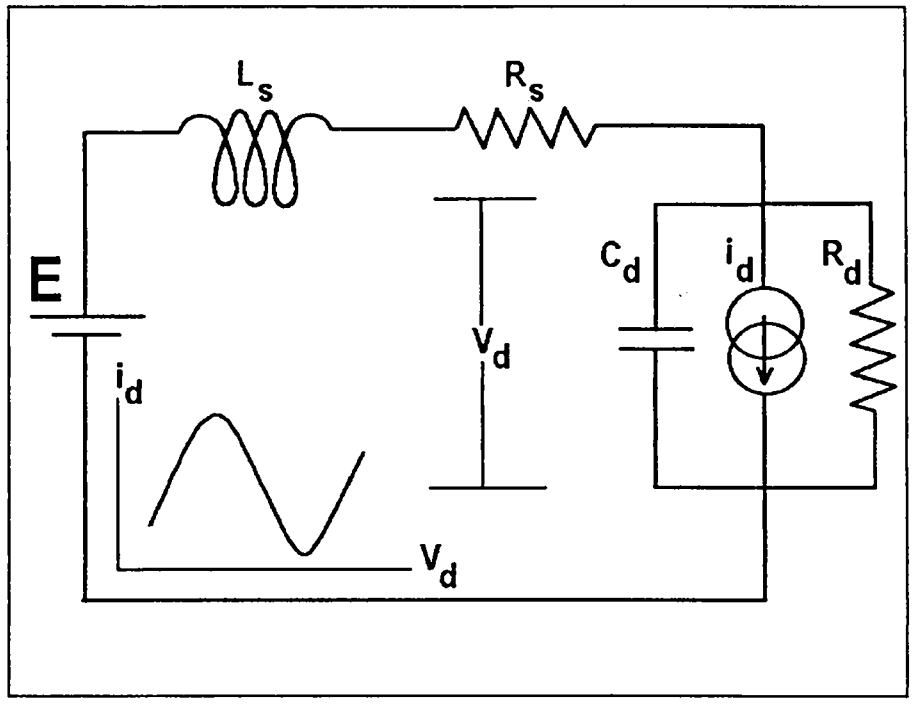


Figure 2.16 lumped element equivalent circuit of a RTD.[19].

Based on this equivalent circuit, the criteria of stability can be read,

$$L_s \leq C_d R_s |R_d| \qquad \textit{Eq. 2-30}$$

$$R_s \leq |R_d| \qquad \textit{Eq. 2-31}$$

This means that the way we are performing the measurements is critical about the achievement of a stable or unstable I-V characteristic notably through the influence of L_s. In addition, it can be seen that the second term of Eq. 2-30 scales as the inverse of the device area. In practice, this motivates the development of very small area devices vehemently discussed in the comment of G. Sollner [23].

Also the resistance level (R_d) can be used for satisfying the stability criteria. This explains why the I-V curve is found stable for a bias voltage close to the valley voltage due to an increase in the diode resistance level. In our case, the current flow through the device is quite low ($\cong$120 A/cm^2 over 20x20 μm^2) and hence the corresponding value R_d is sufficiently high for satisfying the stability condition. Moreover, this high impedance condition prevents us to face the problem of voltage drops in the measurement set-up (mainly due to the tip-shaped probes). Indeed, when a high current is flowing through the device there exists a proportional discrepancy between the test voltage and the voltage applied to the intrinsic diode. As a last argument about the non implication of parasitic oscillations for the present tested device, let us recall the wide range for NDC, surprising at first glance, discussed in details in the introduction of this chapter. For further readings see also [26][27][28].

In conclusion, it is believed that the signature effect extensively studied here is the underlying phenomenon responsible of the shoulder type characteristics of this device.

The fact that the conduction is sustained with a quite constant level is a direct consequence of the supply function. In order to illustrate this issue, the supply function is plotted for various bias voltages in Figure 2.17. One can note that the signature level is practically constant.

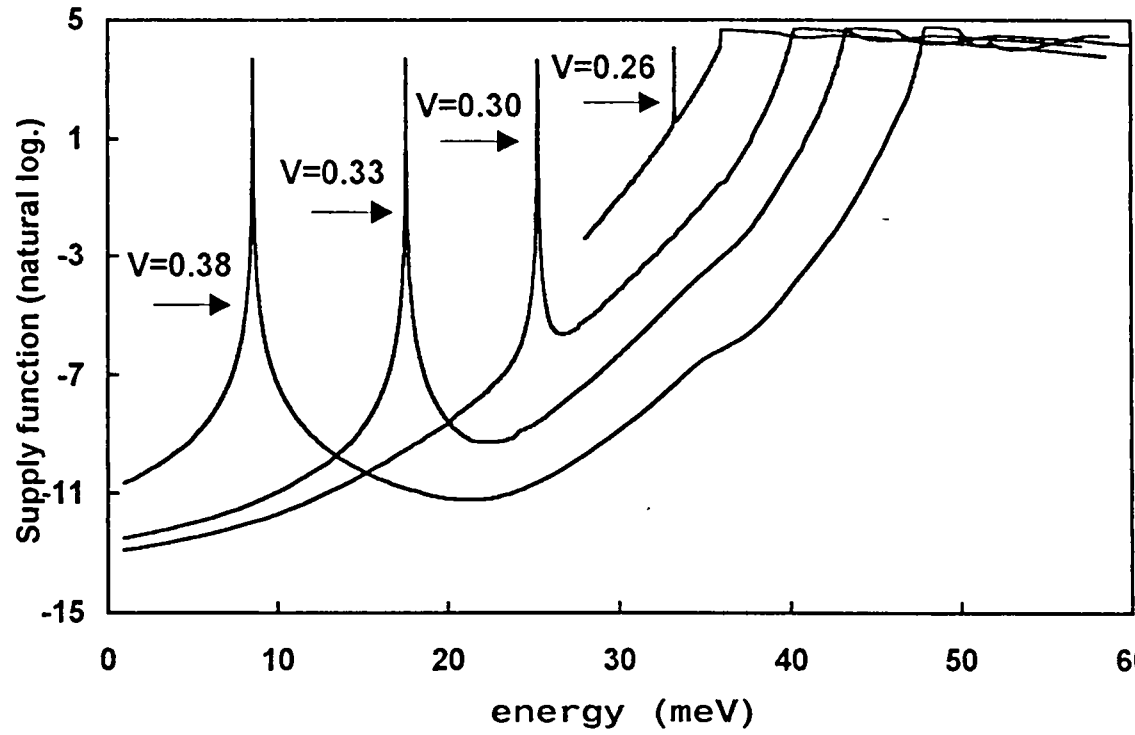


Figure 2.17 supply function plotted for different bias voltages.

2.8 Conclusion :

In this chapter, we have addressed the general problem of charge transfer between two systems of different dimensionality. This was applied to the tunnelling transition between the accumulation layer and the collector region through the resonant path afforded by a double barrier heterostructure (DBH). We found by this means new effects, here referred to as signature effect, resulting from the coupling between the quantum well and the injection zone. In short, beyond the bias point for anti-crossing of quantum states attached to the accumulation and quantum well regions respectively the tail of the wave function is sufficiently high to induce a finite density of states.

To our knowledge this is the first time that a such result is mentioned with a dramatic consequence on the conduction mechanisms under out-of-resonance conditions. In order to investigate this effect, a new tunnelling model was proposed based first on the derivation of the local density of states and second on the definition of a novel supply function which permits us describing how the available states are

populated without using rate equations. It is also found that the electron effective mass has to be re-normalised in the accumulation layer.

On this basis, the current-voltage characteristics have been calculated in details under various assumptions and shows a very broad voltage range for NDC effect with a plateau-like shape in a very good agreement with experimental data. It is believed that, this better understanding of the notions of both the local density of states and the supply function is not limited to the DBH-related phenomena but also could be generalised to the case of open quantum system.

2.9 References

[1] Yoo H.M., Goodnick S.M., Arthur J.R., Influence of spacer layer thickness on current-voltage characteristics of AlGaAs/GaAs and AlGaAs/InGaAs resonant tunnelling diodes, Appl. Phys. Lett. 56 (1), p84 ; 1990.

[2] Koenig E.T., Huang C.I., Jogai B., Independence of peak current from emitter spacer layer width in AlGaAs/GaAs resonant tunnelling diodes, J.. Appl. Phys. 68 (11), p5905, 1990.

[3] Leroux H., Mounaix P. and Lippens D., Investigations of quantum states in the cathode region of resonant tunnelling diodes, Inst. Phys. Conf., serie No.112, chapter 6, Paper presented at Int. Symp. GaAs and Related Compounds. Jersey 1990.

[4] Mounaix P., Vanbésien O., Lippens D., Effect of cathode spacer layer on the current-voltage characteristics of resonant tunnelling diodes, Appl. Phys. Lett. 57 (15), p1517,1990.

[5] Thomas D., Chevoir F., Barbier E., Goldner Y., Vieren J.P., Magnetotunnelling studies of charge build up in double barrier diodes, Proc. Of 4th. Int. Conf. On Superlattices, Microstructures and Microdevices 1988.

[6] Wu J.S., Chang C.Y., Le C.P., Chang K.H., Liv D.G., Liou D.C., Resonant tunnelling of electrons from quantized levels in the accumulation layer of double barrier heterostructure, Appl. Phys. Lett. 57 (22), p2311, 1990.

[7] Koenig E.T., Jogar B., Paulas M.J., Huang C.I., Bozada C.A., Charge quantization effects on current-voltage characteristics of AlGaAs/GaAs resonant tunnelling diodes with spacer layers, Journ. Appl. Phys. 68 (7), p3425, 1990.

[8] Bruggers A., Meiner O., Wölk C., Denfel R., Marten A., Rossmanith M., Klitzing K.V., Saver R., Pseudomorphic two dimensional electron gas emitter resonant tunnelling devices, Microelectronic Engineering, 15, p663, 1991.

[9] Kapre R.M., Madhukar A., Guha S., Highly strained GaAs/GaInAs/AlAs resonant tunnelling diodes simultaneously high peak current densities and peak to valley ratios at room temperature, Appl. Phys. Letters, 58 (20), p2255, 1991.

[10] Lassning R., Boxleitner W., Tunnelling from quasi two dimensional space charge layers, Solid State Communications, Vol.64, No.6979, 1987.

[11] H.Fukuyama, T.Waho and T.Mizutani J.Appl. Phys. 79, pp1801-1806, 1996

[12] Albert Thomas Fromhold, Jr. Quantum Mechanics for Applied Physics and Engineering, ACADEMIC PRESS p248, p85, 1981

[13] Frensley W. R., Effect of inelastic processes on the self consistent potential in the resonant tunnelling diode, Solid State Electronics, Vol.32 No.12, p1235, 1989.

[14] R. Dingle, for the review of early work up to 1975, Festkoerperprobleme 15, 21 (1975).

[15] Claude Weisbuch and Borge Vinter, Laboratoire Central de Recherches Thomson-CSF, Quantum Semiconductor Structures: Fundamentals and Applications, ACADEMIC PRESS, BOSTON.

[16] T.J. Foster, M.L. Leadbeater, L. Eaves, M. Henini, O.H. Hughes, C.A. Payling, F.W. Sheard, P.E. Simmonds, G.A. Toombs, G. Hill and M.A. Pate Current bistability in double-barrier resonant-tunnelling devices, Phys. Rev. B, Vol.39, No.9, p.6205., 1989.

[17] V.J. Goldman, D.C.Tsui and J.E. Cunningham, Observation of intrinsic bistability in resonant-tunnelling structures, Phys. Rev. Lett., Vol.58, No.12, p. 1256., 1987.

[18] A. Zaslavsky, V.J. Goldman, D.C. Tsui and J.E. Cunningham, Resonant tunnelling and intrinsic bistability in asymmetric double-barrier heterostructures, Appl. Phys. Lett. 53 (15), p.1408, 1988.

[19] D.D. Coon, K.M.D.V. Bandara and H. Zhao, Is intrinsic bistability really intrinsic tristability ?, App. Phys. Lett. 54 (21), p.2115, 1989.

[20] T.J. Shewchuk, J.M. Gering, P.C. Chapin, P.D. Coleman, W. Kopp, C.K. Peng and H. Morkoc, Stable and unstable current-voltage measurements of a resonant tunnelling heterostructure oscillator, App. Phys. Lett. 47 (9), p.986, 1985.

[21] H.C. Liu, Simulation of extrinsic bistability of resonant tunnelling structures, App. Phys. Lett 53 (6), p.485, 1988.

[22] E.S. Hellman, K.L. Lear and J.S. Harris, Jr., Limit cycle oscillation in negative differential resistance devices, J. App. Phys. 84 (5), p.2798, 1988.

[23] T.C.L.G. Sollner, preceding comment, Phys. Rev. Lett. 59, p.1622, 1987.

[24] L.A. Davidson, Optimum stability criterion for tunnel diodes shunted by resistance and capacitance, Proceedings of IEEE (Correspondence), p.1233 , 1963.

[25] L.I. Smilen and D.C. Youla, Stability criterion for tunnel diodes, Proc. IRE (Correspondence), Vol.49,p1206,1961.

[26] M.E. Hines, High-frequency negative-resistance circuit principles for Esaki diode applications, The Bell System Technical Journal, May 1960.

[27] C.S. Kim and A. Brändli, High-frequency high-power operation of tunnel diodes, IRE Trans. On Circuit Theory, p.416, 1961.

[28] J. Nagumo and M. Shimura, Tunnel diode loaded by a shorted transmission line, Proc. of the IRE, 1961.

CHAPTER 3

3. THEORETICAL DESCRIPTION OF SCATTERING EFFECTS ON THE HETEROSTRUCTURE DOUBLE BARRIER RESONANT TUNNELLING

3.1 Introduction :

In the previous chapter we studied extensively a DBH structure with highly doped cladding layer. As a consequence, the energy range of injected electrons is broad due to the high degeneracy level. In counterpart, it appears quite problematic to investigate fine scattering-assisted tunnelling transitions which are believed to be one of the key parasitic mechanisms controlling the valley current.

In this chapter, we will address the effects of phonon scattering on the current-voltage characteristics. Our test vehicle will be an AlGaAs/GaAs resonant tunnelling diode fabricated and characterised in our group of research [1].

The main advantage of the structure under consideration is the fact that the doping concentration is low ($\approx 10^{16} cm^{-3}$). This yields a well-defined energy injection under rather low temperature measurements (77 K). In the following, we will take advantage of these properties to assess the validity of the theoretical description of phonon scattering effects.

3.2 Epitaxial material :

Table 3-1 displays the structural parameters of the resonant tunnelling diode which was fabricated and tested in [1]. The DBH consists in three-layered structure involving AlGaAs barriers with an Aluminium content less than 30 % in such a way that the inter-valley (Γ - X) scattering can be avoided. Besides, the barrier height (0.3 eV). As seen later this fact minimises the contribution of interface scattering mechanisms. On each side the extended cladding layer is doped to 2×10^{16} cm^{-3}. They are followed by highly doped emitter and collector regions. The Fermi-level in these adjacent cladding layer which is a key figure of operation as seen in the introduction, depends

significantly on temperature. However, at 77 K the Fermi-level is at the degeneracy limit. In conjunction with a low temperature characterisation the resulting narrow injection energy window is very favourable to discriminate between ballistic and scattering assisted-tunnelling contributions. The fabrication details can be found in reference [1].

Layer	dimension (Å)	doping (cm^{-3})	material
emitter	500	n+	GaAs
cladding	500	$2x10^{16}$	GaAs
first barrier	50	$2x10^{16}$	AlGaAs
quantum well	50	$2x10^{16}$	GaAs
second Barrie	50	$2x10^{16}$	AlGaAs
cladding	500	$2x10^{16}$	GaAs
collector	500	n+	GaAs

Table 3-1 : the epi-layer structure of resonant tunnelling diode fabricated in our group of research by P. Mounaix et al [1],

Briefly, the epitaxy was grown by means of Solid Source Molecular Beam Epitaxy SSMBE system at a constant temperature of 600 ^{o}C starting from a GaAs semi-insulating substrate without growth interruption. For the fabrication of the test samples we used a mesa-etched technology in a quasi-planar configuration. Conventional technology was used for that purpose with AuGeNi metalisation for the Ohmic contacts and NH_4OH-based etchants for defining the active zone laterally. The samples were contacted by means of very fine gold wires, diced and subsequently mounted into a low-temperature fixture. The I-V measurements were carried out at liquid Nitrogen temperature by means of a conventional cryostat.

3.3 Context and main assumptions :

The earlier studies devoted to the influence mechanisms of scattering-assisted tunnelling were published a decade ago. They recognised that inelastic scattering was responsible of a loss of coherence in the resonant tunnelling process. Recently, the work of F. Chevoir and B. Vinter [2][3] was a key contribution in this field by introducing the various types of scattering which can occur between two tightly-coupled quantum wells. Notably, the influence of phonon-assisted processes was clearly identified in reference [4], starting from a very basic situation involving two square quantum-wells. By investigating the coupling probability when the relevant eigenstates are separated by an energy in the vicinity of the characteristic phonon energy ($\hbar\,\omega$), they showed a possible improvement in the well-coupling probability.

Our aim here is to investigate this kind of transport mechanism in a real device. We have seen previously that a similar 2D- transfer of charge is also encountered when we focus our attention on the 2D injection zone and the DBH. This means that we have only to consider the wave function coherence within these zones. It is obvious that in the contact and cladding layers far from the heterojunction there is no mean to preserve a coherent transport due to the very long distance. The straightforward consequence is that the emitter as well as the collector regions are regionally in quasi-equilibrium.

Ideally a microscopic approach such as that carried out by means of Monte-Carlo procedure could take all the scattering effects into account. It is now well known that such a numerical method based on free flight and scattering events with a random trials of their occurrence is very powerful to have a physical insight into the transport properties of low but not quantum sized dimensional devices. In particular the non-stationary dynamic effects such as velocity overshoot or quasiballistic transport. To our knowledge, there is now an increasing effort to develop theoretically a quantum Monte-Carlo approach based notably on the Bohm trajectories [5][6]. In this approach the Hamiltonian eigenstates have been shown to be non suitable and hence time-dependent wavepackets are required [7].

These kinds of models are now able to describe the transport properties after an injection of hot electron through a DBH injector. The next step is to introduce the scattering mechanisms in the adjacent layer as well as in the quantum region.

In our work we choose an intermediate approach which consists to determine the frequency of collision of various types of scatterings. This is carried out similarly as that in Monte-Carlo simulations. This is the first step and in section 3.4 we review the possible types of scattering mechanisms.

3.4 Possible types of scattering mechanisms :

Scattering mechanisms can be subdivided into two main categories namely bulk scattering mechanisms and interface scattering mechanisms. There exists another classification based on the resulting-in energy exchange, elastic scattering mechanism results in no energy exchange unlike the inelastic scattering mechanism in which the electron may gain or lose a certain amount of energy.

3.4.1 Bulk scattering mechanisms :

In the following we will consider that the scattering events are non-correlated. This assumption of statistically independence permits us to consider each interaction separately. Let us recall that this assumption is also used in the Monte-Carlo simulations where no memory effect is the considered. Another issue which can be discussed is the validity of a bulk rates for short dimension devices. Analogous situation is encountered between the study of transport properties in a bulk or a real short dimension device. In the former one electron can be studied during a long time whereas an ensemble Monte-Carlo code has to be used in the second case. This means that, by ergodicity principle there is equivalence between the two approaches.

3.4.1.1 Phonon scattering (lattice vibrations):

The physical picture of phonon scattering in crystalline lattice is depicted in Appendix(A) by means of the well-known spring approach. First considered in a monatomic lattice extended to a second step by considering diatomic crystals. On this basis the scattering rate for acoustic phonon is calculated in details taking the compression and tension motion of atoms into account. Here, we show that this acoustic phonon rate can be derived by another way which can be summarised in the following equation [8].

$$\nu(k) = \frac{V}{8\pi^3} 2\pi \int_0^{\pi} \left[s(\vec{k},\vec{k}') \, k'^2 \, dk' \right] \sin\beta' \, d\beta' \qquad \text{Eq. 3-1}$$

where,

- $s(\vec{k},\vec{k}')$ is the probability density of scattering from state $\vec{k}$ to state $\vec{k}'$
- $\frac{V}{8\pi^3}$ denotes the density of states in the momentum space.
- the factor 2π reflects the symmetry of the scattering relative to the momentum.

It can be shown that the probability density can be written :

$$s(\vec{k},\vec{k}') = \frac{2\pi}{\hbar} B_a(\vec{k},\vec{k}') N_a \delta\left(|\vec{k}| - |\vec{k}'|\right) \qquad \textit{Eq. 3-2}$$

where,

$$B_a(\vec{k},\vec{k}') = 2\frac{\hbar \Xi^2}{2\rho S V}|\vec{k} - \vec{k}'| \qquad \textit{Eq. 3-3}$$

With S and ρ the speed of sound and the specific density of the semiconductor material respectively and N_a is the Bose- Einstein distribution function.

$$N_a = \frac{1}{\exp\left(\frac{S\hbar|\vec{k} - \vec{k}'|}{k_B\theta}\right) - 1} \cong \frac{k_B\theta}{S\hbar|\vec{k} - \vec{k}'|} \qquad \textit{Eq. 3-4}$$

substituting from Eq. 3-2, Eq. 3-3 and Eq. 3-4 into Eq. 3-1 we get,

$$\nu = \frac{\sqrt{2}\, m^{*3/2} k_B \theta\, \Xi^2 \sqrt{\varepsilon}}{\pi \rho\, S^2\, \hbar^4} \qquad \textit{Eq. 3-5}$$

Eq. 3-5 is similar to the scattering rate equation derived in Appendix(A). Eq. 3-1 can be applied to any other type of scattering mechanism notably for calculating the scattering rate by optical modes [8].

3.4.1.1.1 Lattice scattering by optical modes :

Applying the same procedure (Eq. 3-1) outlined above we get for the lattice scattering by optical mode[8],

$$s(\vec{k},\vec{k}') = \frac{2\pi}{\hbar}\, \beta_o(\vec{k},\vec{k}') \begin{cases} N_o & \delta\left[\varepsilon(k') - \varepsilon(k) - \hbar\omega\right] \\ (N_o + 1) & \delta\left[\varepsilon(k') - \varepsilon(k) + \hbar\omega\right] \end{cases} \qquad \textit{Eq. 3-6}$$

where,

$$\beta_o(\vec{k},\vec{k}') = \frac{2\pi q^2 \hbar\omega}{4\pi \in_o \left|\vec{k} - \vec{k}'\right|^2} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_s}\right) \qquad \textit{Eq. 3-7}$$

ε_∞ and ε_s are the relative dielectric constants, for the infinity limit and the zero limit frequencies respectively and $\in_o$ is the dielectric constant of free space.

$$N_o(\omega,\theta) = \frac{1}{\exp\left(\hbar\omega / k_b\theta\right) - 1} \qquad \textit{Eq. 3-8}$$

Substituting from Eq. 3-6, Eq. 3-7 and Eq. 3-8 into Eq. 3-1 we get,

$$\nu_o = \frac{q^2 m^{*1/2} \omega}{\sqrt{2}\,\hbar\ 4\pi \in_o} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_s}\right) \frac{1}{\sqrt{\varepsilon}}\ \ln\left|\frac{\sqrt{\varepsilon} + \sqrt{\varepsilon \pm \hbar\omega}}{\sqrt{\varepsilon} - \sqrt{\varepsilon \pm \hbar\omega}}\right| \begin{cases} N_o \\ N_o + 1 \end{cases} \qquad \textit{Eq. 3-9}$$

In reference [9] by referring to the well-known data of GaAs Eq. 3-9 becomes

$$\nu_o \approx 6x10^{12}\begin{cases} N_o & \Leftrightarrow \text{absorption} \\ 2N_o + 1 & \Leftrightarrow \text{emission} \end{cases} \qquad \text{Eq. 3-10}$$

3.4.1.2 Scattering by ionised impurities :

Scattering by the ionised impurities is a result of the Coulomb field Figure 3.1 illustrates an electron of velocity v approaching the ionised charge (+Ze).

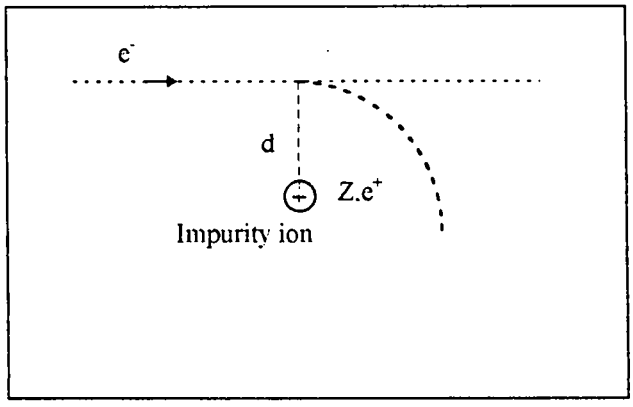


Figure 3.1 simple model for scattering by ionised impurity.

We will neglect the electron deviation while approaching the positive ion and hence one can assume that the smallest separation to be the distance (d). The maximum attraction force is

$$F_{max} = m^* a_\perp = \frac{Ze^+}{4\pi\ \varepsilon\ d^2} \qquad \textit{Eq. 3-11}$$

where $a_\perp$ is the perpendicular acceleration.

The interaction time might be approximated by (d/v) with v the initial tangential velocity, so the perpendicular velocity component might be given by

$$V_\perp = a_\perp \frac{d}{v} = \frac{Ze^2}{4\pi\,\varepsilon\ m^* v\ d} \qquad \textit{Eq. 3-12}$$

if $(V_\perp \approx v)$ then the deflection angel is 45^O. Also if the magnitude of the attraction is high enough so that $(V_\perp \geq v)$ a collision is said to be taken place. This is equivalent to say that,

$$d \le d_c = \frac{Ze^2}{4\pi\varepsilon\, m^* v^2} \qquad \textit{Eq. 3-13}$$

We say then that the collision cross section is ($\pi\, d_c^2$). By denoting (N) the volume density of ionised impurities, the average scattering rate is given by

$$\nu = \left\langle \pi\, d_c^2 N v \right\rangle \qquad \textit{Eq. 3-14}$$

By further assuming that the statistical average of product equal the statistical product of averages and also that the source of electron energy is only thermal, then by simple algebraic manipulation one find,

$$\nu = \frac{N\ Ze^4}{16\pi\ \varepsilon^2 (m^*)^{1/2} (3kT)^{3/2}} \qquad \textit{Eq. 3-15}$$

Eq. 3-15 states that the ionised impurity scattering rate is most effective at low temperatures provided that N is sufficiently high.

3.4.2 Interface scattering mechanisms :

3.4.2.1 Interface roughness scattering :

Interface roughness scattering will be treated in our model as a factor independently affecting the transmission probability of each barrier. Therefore its overall effect in a double barrier structure is described by replacing the transmission probability of a perfectly smooth barrier by that of an equivalent rough one. So our goal in this section will be to calculate the effect of surface roughness on a single barrier structure.

The interface roughness is usually modelled as a statistical distribution of terraces of monatomic layer thickness Δ. The probability density which governs the statistical distribution of their size is usually chosen to be of Gaussian type.

Following Leo and MacDonald [15] the interface roughness scattering potential is described [16] as,

$$V^{IR} = V_b \delta(z - z_i) \exp\left(-\frac{(r_i - r_\perp)^2}{2\sigma^2}\right) \qquad \textit{Eq. 3-16}$$

where, the kronecker symbol was used for addressing the interface location at z_i and r_i the lateral position of the terrace with respect to the coordinate $r_\perp$ and V_b is the potential height at the barrier. Z_i and r_i. are illustrated in Figure 3.2 and σ is the average terrace width.

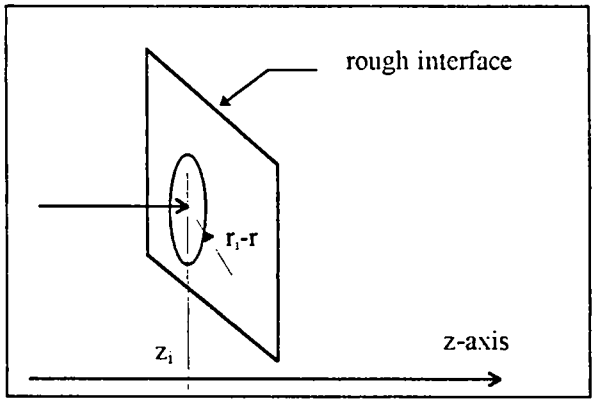


Figure 3.2 illustration of axis notation

Eq. 3-16 states that as the distance between the terrace centre and the position of electron incidence is increased the effect of scattering potential decreases exponentially with a characteristic decay length equal to twice the terrace average area.

If we now consider an ensemble of terraces mutually uncorrelated in position and width, it can be shown that the ensemble average interaction is [16],

$$M^{IR} = V_b^2 4\pi^2 D_T \sigma^4 \exp(-Q_\theta^2 \sigma^2) \qquad \textit{Eq. 3-17}$$

D_T is the interface surface density of terraces and ($Q_\theta = K'_\perp - K_\perp$), where $K'_\perp$ and $K_\perp$ are the initial and final transverse momentum (parallel to the interface).

If one assumes that ($\Lambda = 2\sigma$) and ($D_T = 4/\pi\Lambda^2 \cong 1/\Lambda^2$) i.e the average separation between terraces is equal to average terrace width we get, [3][16]

$$M^{IR} = V_b^2 \pi \Lambda^2 \exp\left(\frac{-Q_\theta^2 \Lambda^2}{4}\right) \qquad \textit{Eq. 3-18}$$

The angular integration with respect to θ yields,

$$M^{IR} = V_b^2 \pi^2 \Lambda^2 \exp\left(-\frac{K^2\Lambda^2}{2}\right) I_o\left(\frac{K^2\Lambda^2}{2}\right) \qquad \textit{Eq. 3-19}$$

where I_o is the modified Bessel function of the first type of order zero. By applying the analysis described in details [3] we finally obtain the capture probability rate per unit overlap integral as,

$$\partial W^{IR} = V_b^2 \pi^2 \Lambda^2 \exp\left(-\frac{K^2\Lambda^2}{2}\right) I_o\left(\frac{K^2\Lambda^2}{2}\right) \frac{2m^*}{\hbar^2} \frac{\partial K_\perp}{k_z} \qquad \textit{Eq. 3-20}$$

By knowing that the overlap integral or the form factor in [3] is,

$$F^{IR} = \Delta_{IR}^2 \left|\zeta_{k_z}(z_i)\right|^2 \left|\zeta_{rw}(z_i)\right|^2 \qquad \textit{Eq. 3-21}$$

where $\zeta_{k_z}(z_i)$ and $\zeta_{rw}(z_i)$ are the envelope wave functions at the interface (z_i) which correspond to the initial and final states (k_z, k_{rw}). The form factor described by Eq. 3-21 is the indication of the quantum probability of finding the electron in the neighbourhood (Δ_{IR}) of the interface.

Then to get the capture probability rate we must multiply Eq. 3-20 with Eq. 3-21. Let us now discuss the condition in which Vinter et al [3] derived the form factor. First they assume flat conduction band condition in the emitter region. Second they assumed that all the scattered electrons will be collected only in the well resonant level (the final state is always k_{rw}) which is not our case.

The form factor includes information about transmission and reflection probabilities. This is in the sense that when the reflection probability is maximised (unity value) the transmission probability is minimised (zero value) and the transmitted wave will be destructively interfere with the reflected one so $\zeta_{k_z}(z_i)$ will be minimised (zero value). In the other extreme case when the transmission probability is unity the electrons (quantum mechanically) do not feel the existence of the interface and hence the interface scattering is minimised. The last case is excluded in the analysis of Vinter et al [3] because they study the scattering only in the valley current range.

Besides a flat conduction band profile is used to calculate these envelope wave functions. In addition, the normalisation condition of the wave function in the analysis of [3] is not discussed.

In our case of a single barrier all the states on the other side are available and the structure is an open structure. Then the most plausible way to carry almost the same information which is carried out by the introduction of the form factor is just replacing Eq. 3-21 by

$$F^{IR} = \Delta_{IR}^2 \, T(\varepsilon_{zi}) \, R(\varepsilon_{zi}) \qquad \textit{Eq. 3-22}$$

The form factor as described by Eq. 3-22 has the advantage that it includes the back scattering effect and hence the current ($R+T=1$) is conserved.

Integration of Eq. 3-20 is performed as follows :

$$\partial W^{IR} = V_b^2 \pi^2 \Lambda^2 \frac{2m^*}{\hbar^2} \frac{\partial K_\perp}{k_z} \Sigma \qquad \textit{Eq. 3-23}$$

where

$$\Sigma = \int_0^\infty \exp\left(-\frac{K^2\Lambda^2}{2}\right) I_o\left(\frac{K^2\Lambda^2}{2}\right) \qquad \textit{Eq. 3-24}$$

using the change of variables

$$x = \frac{K_\perp^2}{2}, \quad dx = 2K_\perp \Lambda^2 dK_\perp, \quad dK_\perp = \frac{1}{2\Lambda\sqrt{2}} (x)^{1/2} \qquad \textit{Eq. 3-25}$$

by algebraic manipulations we get,

$$\Sigma = \frac{\sqrt{\pi, 2}}{2\Lambda} \int_0^1 (-\lambda^3 + 2\lambda^2 + 2\lambda)^{-1/2} \, d\lambda \qquad \textit{Eq. 3-26}$$

performing the above integral numerically we get,

$$W^{IR} = 1.278 \; V_b^2 \pi^2 \Lambda \frac{2m^*}{\hbar^2} \frac{1}{k_z} \frac{\sqrt{\pi/2}}{2} \qquad \textit{Eq. 3-27}$$

and hence the total transmission can be written as,

$$T_{tota} = T_c + W^{IR} \Delta_{IR} \, T_c R_c = T_c \, (1 + W^{IR} \, R_c \, \Delta_{IR}) \qquad \textit{Eq. 3-28}$$

where W^{IR} is the transmission via interface roughness, T_c and R_c are the transmission and reflection coefficient respectively and Δ_{IR} is the effective interface thickness.

3.5 Transmission coefficient:

Starting from the WKB approximation and apply it for a very approximated structure as shown in Figure 3.3 one can get an analytical expression describing the transmission probability [14].

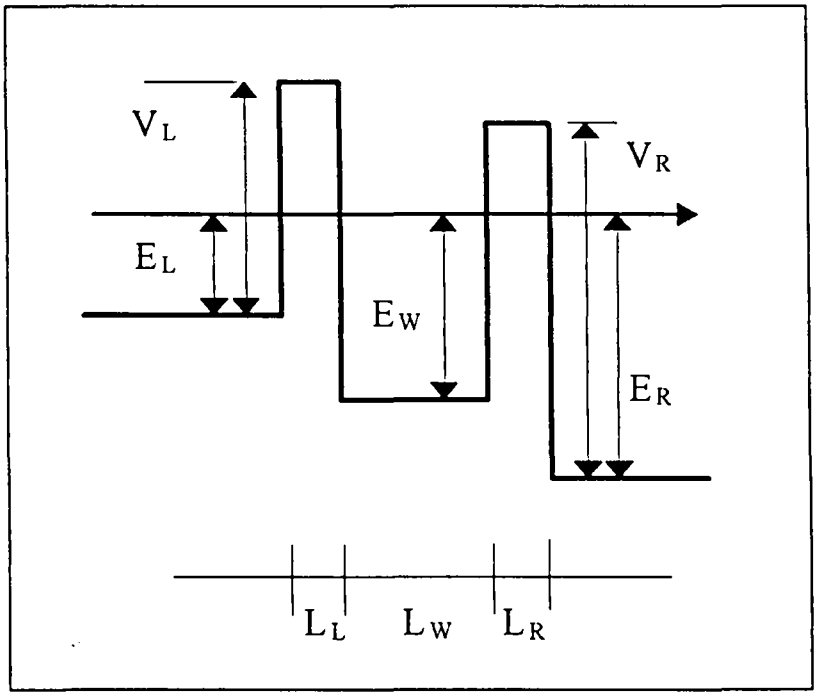


Figure 3.3 simplified energy band diagram of double barrier structure

$$T_c = \left| \frac{-|t_L||t_R| \exp[i(k_W L_W - \chi_{L1} - \chi_{R1})]}{1 - |r_L||r_R| \exp(i2\Delta)} \right|^2 \qquad \textit{Eq. 3-29}$$

where

L_W, L_R and L_L are the thickness of the well and of the two barriers with the subscript L and R standing for left and right.

We also define according to a specific region the attenuation coefficient ($\alpha_i \quad i = L, R$) and the propagation one ($k_j \quad j = L, R$) with the following expressions.

$$\alpha_L = \sqrt{2m(V_L - E_L)/\hbar^2} \quad , \quad \alpha_R = \sqrt{2m(V_R - E_R)/\hbar^2} \qquad \textit{Eq. 3-30}$$

$$K_L = \sqrt{2mE_L/\hbar^2} \quad , \quad K_R = \sqrt{2mE_R/\hbar^2} \quad , \quad K_W = \sqrt{2mE_W/\hbar^2} \qquad \textit{Eq. 3-31}$$

The coefficients denoted t and r describe the transmission and reflection coefficients on the left hand side and right hand side respectively. Their modulus can be expressed as follows,

$$t_L = \left| \frac{i2\cos(\theta_{IL})\sin(\theta_{OL})}{\cos(\theta_{IL}+\theta_{OL})\sinh(\alpha_L L_L) + i\sin(\theta_{IL}+\theta_{OL})\cosh(\alpha_L L_L)} \right| \sqrt{\frac{K_W}{K_L}} \qquad \textit{Eq. 3-32}$$

$$t_R = \left| \frac{i2\cos(\theta_{IR})\sin(\theta_{OR})}{\cos(\theta_{IR}+\theta_{OR})\sinh(\alpha_R L_R) + i\sin(\theta_{IR}+\theta_{OR})\cosh(\alpha_R L_R)} \right| \sqrt{\frac{K_R}{K_W}} \qquad \textit{Eq. 3-33}$$

In these relations θ coefficient has been introduced expressed as,

$$\theta_{IL} = \tan^{-1}(\alpha_L / K_L) \quad , \quad \theta_{OL} = \tan^{-1}(\alpha_L / K_W) \qquad \textit{Eq. 3-34}$$

$$\theta_{IR} = \tan^{-1}(\alpha_R / K_W) \quad , \quad \theta_{OR} = \tan^{-1}(\alpha_R / K_R) \qquad \textit{Eq. 3-35}$$

At last the phase terms χ and Δ respectively can be written as a function of θ, α and the thickness as follows,

$$\tan\chi_{L1} = \tan(\theta_{IL}+\theta_{OL})\coth(\alpha_L L_L), \ \tan\chi_{L2} = \tan(\theta_{IL}-\theta_{OL})\coth(\alpha_L L_L) \qquad \textit{Eq. 3-36}$$

$$\tan\chi_{R1} = \tan(\theta_{IR}+\theta_{OR})\coth(\alpha_R L_R), \ \tan\chi_{R2} = \tan(\theta_{IR}-\theta_{OR})\coth(\alpha_R L_R) \qquad \textit{Eq. 3-37}$$

$$2\Delta = 2K_W L_W - (\chi_{L1} - \chi_{L2}) - (\chi_{R1} + \chi_{R2}) \qquad \textit{Eq. 3-38}$$

In the above treatment the wave function in each region is assumed to be an exponential function with constant amplitude. This means that $(A/\sqrt{k})$ for propagation regions or $(A/\sqrt{\alpha})$ for attenuation regions were assumed to be independent of position within the corresponding region, which is an approximation

only valid for the flat or nearly flat conduction band structure (zero field approximation).

In addition, the above treatment would not be valid anymore if the x-energy value is near the conduction band edge such that $\left| k(x)/k'(x)^2 \right|$ .or $\left| \alpha(x)/\alpha'(x)^2 \right|$ is greater than unity which is the limit of validity of WKB approximation [17]. In other words, this model appears too simplified to give a realistic insight into the device physics and this fact motivates us to derive a more exact calculations.

3.5.1 Calculation of transmission coefficient :

Our goal is to overcome the key limitation of the flat conduction band approximation. It is now well-known that the Airy function is the exact solution in that case [18]. In the same time however it could be useful to conserve the simplicity of the exponential approximation of the WKB method. For this purpose the Airy function (exact solution [18]) is analysed by the product of two exponential functions one having a real exponent (attenuation responsible) and the other an imaginary one (propagation responsible). Therefore we will name this method as the Airy-Exponential Exact Solution (AEES) which will be one of the novel features of our model.

$$e^{-\alpha(x)} e^{i\beta(x)} = B_i[\zeta(x)] + iA_i[\zeta(x)] \qquad \textit{Eq. 3-39}$$

where

$$\zeta(x) = \left[2m/(eF\hbar)^2\right]^{1/3}\left[\varepsilon_0(x) - \varepsilon_x\right]$$

$\varepsilon_o(x)$ is the conduction band profile

F is the electrostatic field.

by simple algebraic manipulations we can get,

$$\alpha = \frac{A_i A_i' + B_i B_i'}{A_i^2 + B_i^2} \cdot \zeta' \qquad \textit{Eq. 3-40}$$

$$\beta = \frac{B_i A_i' - A_i B_i'}{A_i^2 + B_i^2} \cdot \zeta' \qquad \textit{Eq. 3-41}$$

In order to illustrate the difference between the WKB approximation and the AEES method we plotted in Figure 3.4 the propagation factor β and the attenuation factor α versus energy for a single barrier heterostructure under a constant electric field. The reference of energy is taken in the emitter region.

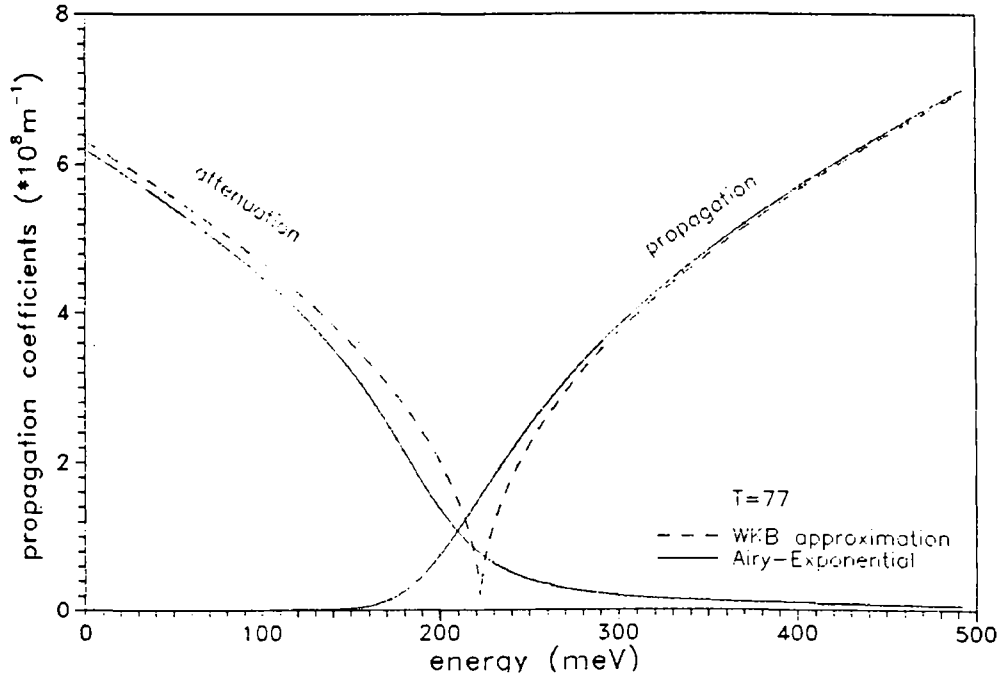


Figure 3.4 propagation factor (β) and attenuation factor (α) calculated for the resonant double barrier structure of Table 3-1.

In the above figure, far from the barrier height (220 meV) the WKB approximation fits quite well the Airy solution. In contrast, at energies near the barrier height the WKB approximation does not seem realistic with a breaking point at 220 meV.

The AEES method is not restricted to a single barrier and can be applied without difficulty to DBH structures while maintaining a constant electric field over the active quantum-region. The transmission probability calculated for the resonant double barrier whose epilayer sequence is listed in Table 3-1 is displayed in Figure 3.5. The bias voltage applied is close to the threshold voltage for NDC effect as seen by the energy position of the ground state close to the emitter band edge.

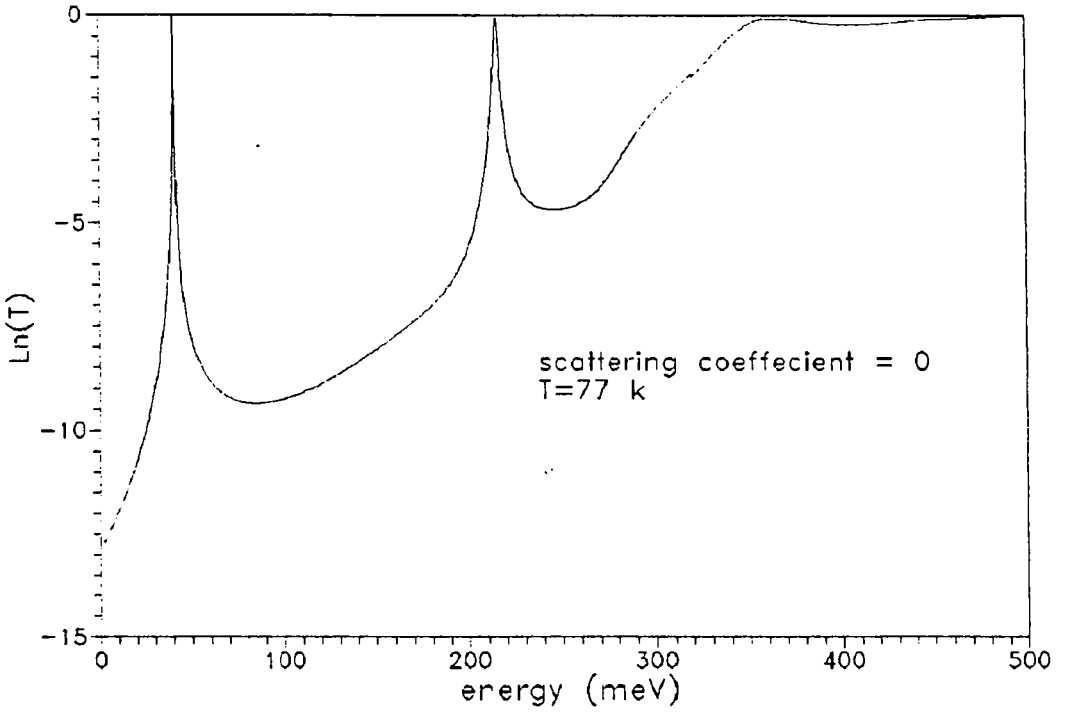


Figure 3.5 transmission probability for the resonant double barrier of Table 3-1 calculated using AEES method.

The transmission curve in the above figure was calculated at 77K without scattering and will be used here after for deriving the purely coherent quantum transmission probability.

In case of when using WKB approximation the model will fail to describe the enhanced resonant transmission through the first excited state which strongly influence the out of resonance level. The present AEES method does not require to consider a

constant electric field any where in the structure but could be generalised using piece-wise constant field approximation.

3.5.2 Scattering effects on transmission probability :

We learned in section 3.4.2 that the surface roughness-related scattering depends strongly of the barrier height and also of the parameter Δ. For the sample under test the barrier height is relatively low besides it is believed that surface quality is quite good in view of the high performance electrical characteristics. Therefore in a first stage we will focus our intention about phonon scattering effect and more particularly on longitudinal optical mode (see appendix A).

Physically, it can be understood that the life time of an electron within the quantum-well is a key parameter in the loss of coherence. This conclusion was early recognised by saying that the probability of finding an electron in a coherent state within a quantum-well of thickens L_W is reduced exponentially with an attenuation coefficient related to the scattering coefficient with the factor $\exp(-2\gamma L_W)$

A.D.Stone & P.A.Lee have introduced this idea for the first time [19] assuming γ as an adjustable parameter to be determined to get agreement with experiment. Later on Yasuhito Zohta et al [12][13][14] have extended the theory assuming $\gamma = 1/2l_f$ where l_f is the electron mean free path in the bulk material.

In this model, to kinds of transmission probabilities are introduced, refereed as coherent (T_c) and incoherent (T_i) transmission probabilities. Moreover, to conserve a unity probability we have to introduce an attenuation coefficient denoted her A_T which physically describes the incoherent part in the tunnelling process.

The coherent quantum transmission probability (T_c) and the coherent reflection probability (R_c) have the following expressions : [13]:

$$T_c = \left| \frac{-|t_L||t_R| \exp(-\gamma L_W) \exp[i(kL - \chi_{L1} - \chi_{L2})]}{1 - |r_L||r_R| \exp(-2\gamma L_W) \exp(i2\Delta)} \right|^2 \qquad \textit{Eq. 3-42}$$

$$R_c = \left| \frac{|r_L| - |r_R| \exp(-2\gamma L_W) \exp(i2\Delta)}{1 - |r_L||r_R| \exp(-2\gamma L_W) \exp(i2\Delta)} \right|^2 \qquad \textit{Eq. 3-43}$$

From them we can deduce the attenuation coefficient A_T :

$$AT = 1 - (T_c + R_c) \qquad \textit{Eq. 3-44}$$

by simple algebra A_T reads :

$$A_T = \frac{T_L \left[1 - |r_R|^2 \exp(-4\gamma L_W) \right] - T_L T_R \exp(-2\gamma L_W)}{\left| 1 - |r_L||r_R| \exp(-2\gamma L_W) \exp(i2\Delta) \right|^2} \qquad \textit{Eq. 3-45}$$

Now we have to derive the incoherent part of the transmission probability. Over the past our group, dealing with this issue proposed to assume that an electron after loosing its coherence escapes from the quantum-well prorata to the elementary transmission of the left and right barriers [11]. Mathematically, the incoherent transmissions from left to right and vice versa read,

$$T_i^{RL} = A_T \cdot \frac{T_L}{T_L + T_R} \qquad \textit{Eq. 3-46}$$

$$T_i^{LR} = A_T \cdot \frac{T_R}{T_L + T_R} \qquad \textit{Eq. 3-47}$$

It seems interesting to discuss here the validity of the above assumption which can appear at first glance relatively crude. In fact, it is believed that an electron which has just experienced an inelastic scattering loose the memory of resonant effect (phase-breaking). It could try to reconstruct a new phase memory but this implies several back

and forth bouncing motion within the quantum-well. In reality, it is believed that the escape time is too short to satisfy this condition. As a consequence, the electron experiences a very simple escaping process through a single barrier heterostructure.

Finally, the overall transmission (total) is the summation of coherent and incoherent contributions.

$$T_{TOTAL} = T_c + T_i \qquad \textit{Eq. 3-48}$$

Before using the above equation we are now discussing the energy conservation issue. This discussion is motivated by the large discrepancy obtained between the peak current calculated by means of the original model of Yasuhito Zohta et al [12][13][14] and the experiment (one order of magnitude). Moreover, the LO phonon peak has not any significant value on the I-V simulated results. While reviewing the assumptions made in references [12][13][14], it appears that the transverse energy was assumed to be equal zero which is not realistic. In fact, it was early recognised [2][3][4] that an elastic scattering in confined system, with thus the separation between the longitudinal and transverse energies, can be considered as an inelastic event in the direction of propagation. To take this fact into account we will assume, in a next section, that the probability of finding a scattered electron is uniformly distributed over an energy window $\hbar\omega$ optical phonon energy. The uniformly distributed probability density function is centred above (absorption event) or below (emission event) the injected electron energy level.

On the other hand, the authors of references [12][13][14] introduced a scattering factor $\exp(-2\gamma L_{II})$ which compares the scattering time to a transit time (simple crossing) within the quantum-well. This contradicts the fact that the electron occupies the quantum-well for a finite time suffering from multiple reflections. It is sure that we have to compare the ensemble averaged free-flight time within the quantum-structure to the resonant level life time.

Worth noting that the effect of considering the multiple reflections increases the probability of scattering and hence appreciable decrease in the peak current could be observed. We are now interesting to derive the life time under scattering condition.

3.5.2.1 Life time under scattering condition :

The basic idea is to take into account the number of finite reflections by considering a factor such that $\gamma = a\,\gamma_o$ where γ_o denotes the bulk value of γ and a is the number of finite reflections.

Thus multiplying γ_o by the number of finite reflections a such that,

$$\gamma = a\,\gamma_o \qquad \textit{Eq. 3-49}$$

which is exactly equivalent to replacing L_W in the analysis of Zohta et al [13] by,

$$L_{effective} = a\,L_W \qquad \textit{Eq. 3-50}$$

Knowing that,

$$a = \tau_{er}\,\frac{v}{L_W} \qquad \textit{Eq. 3-51}$$

On the other hand, it is well known that the life time τ_{er} is given by,

$$\tau_{er} = \frac{\hbar}{\Gamma(a)} \qquad \textit{Eq. 3-52}$$

where Γ is the full energy width at half maximum (FWHM) of the total transmission coefficient.

In this work Eq. 3-51 and Eq. 3-52 are self consistently solved by numerical techniques. However, to give a physical insight, an approximate analytical expression for the solution could be derived as follows :

For narrow quantum well structures, having T_L, T_R and $2a\gamma_o L_W$ much less than unity, an analytical expression for $\Gamma(a)$ could be described by,

$$\Gamma(a) = \frac{T_L + T_R}{2} \frac{\hbar v}{L_W} + 2\hbar v a \gamma_o \qquad \textit{Eq. 3-53}$$

The first term describes the coherent contribution whereas the second introduces the additional broadening resulting from scattering processes.

solving Eq. 3-51, Eq. 3-52 and Eq. 3-53 together we get,

$$a^2 (2\gamma_o L_W) + a\left(\frac{T_L + T_R}{2}\right) - 1 = 0 \qquad \textit{Eq. 3-54}$$

which is a second order algebraic equation in a and is one of the most significant contributions of the new model.

It is worth mentioning that Eq. 3-54 has to be applied separately for each (absorption or emission) mechanism. This is because each interaction process has a different γ_o value and hence a different life time.

In practice it is found that the difference between the analytical and the numerical one is minor at low temperatures under relatively strong confinement condition. However, the transmission probabilities which are displayed in Figure 3.6 associated with each scattering process are derived using numerical approach.

In Figure 3.6 the coherent (solid line) and the incoherent (dashed line) transmissions were plotted in logarithmic scale for the the LO phonon emission process (Figure 3.6a) and LO phonon absorption scattering (Figure 3.6b). For the former, it is seen that near the ground energy level the incoherent transmission is enhanced with respect to the coherent one. We have to stress that the scattering process here is particularly efficient such that electron suffering a phonon emission take advantage of a resonant tunnelling effect through the quantum well resonant level. In contrast, the phonon absorption is of minor concern.

At this stage, it remains two important issues to be discussed namely the injection condition investigated in section 3.6 and at last the space charge effect which will be studied in section 3.7.

Figure 3.6 depicts the different transmission probabilities associated with each scattering process.

It can be shown that $\exp\left(-2\alpha\gamma_o L_{//}\right) = \exp\left(-\tau_{er} / T_f\right)$ where T_f is the average time between two successive scatterings and τ_{er} is the electron life time.

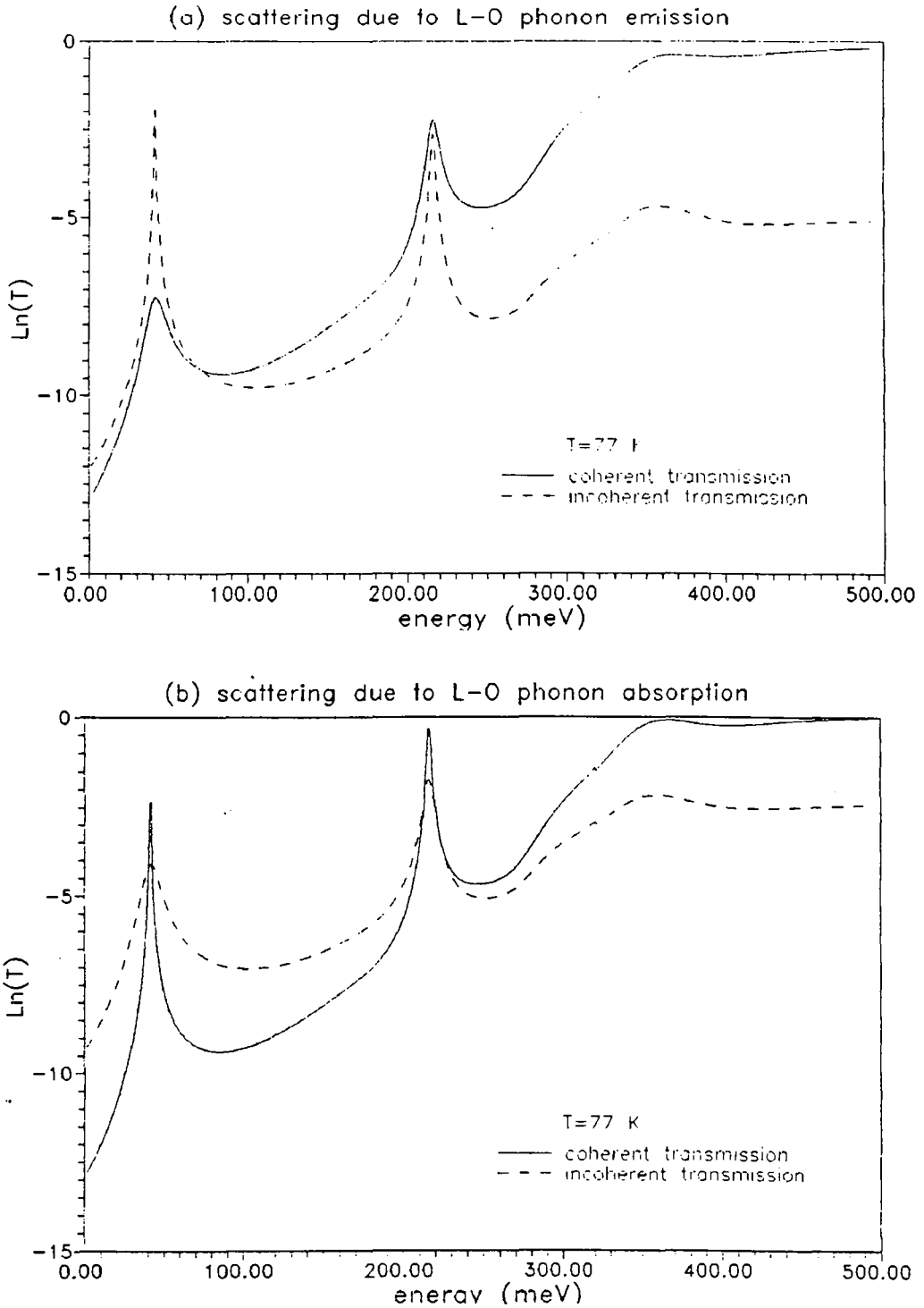


Figure 3.6 the transmission probability calculated using the AEES method (a) for the emission process (b) for the absorption process.

3.6 Scattering effect on injection conditions :

The effect of scattering on injection conditions could be summarised in two key points. First is to create a new set of local density of states (starting from the coherent one) which will be denoted as the incoherent local density of states. Second is to occupy these incoherent states by the appropriate probability of occupancy.

To describe the above mentioned incoherent density of states the following assumptions shall be adopted :

From a previous discussion we have shown that the phase memory of the scattered electron would be lost just after scattering and then the scattered electron try to recover a new phase memory in a finite transient time. In order to provide a time scale, it is reasonable to assume that this finite built-up time is much longer than the so called simple tunnelling time.

Also, we will assume that an incoherent local density of states can be derived from the purely coherent one by taking the 3D-carachter of interaction namely the sharing of the quantum energy $\hbar\omega$ between the longitudinal and the transverse directions into account.

To achieve that, the original coherent local density of states are maximally activated by assuming a unity probability of occupancy. In addition, as the scattering process occurs in the time domain in a completely random manner, the probability of scattering is assumed to be uniformly distributed over a range of energy equal to the LO phonon energy ($\hbar\omega_{op}$). Thus the scattered electrons will be distributed over the energy range ($\hbar\omega_{op}$) as described by the following equation,

$$n_{local}(\varepsilon,\sqrt{2m\varepsilon_x})=\begin{cases}\dfrac{1}{\hbar\omega_{op}}\cdot\displaystyle\int_{\varepsilon_x}^{\varepsilon_x+\hbar\omega_{op}} g(\varepsilon,\sqrt{2m\zeta})\,d\zeta\ldots\Leftrightarrow phonon.emission\\ \dfrac{1}{\hbar\omega_{op}}\cdot\displaystyle\int_{\varepsilon_x-\hbar\omega_{op}}^{\varepsilon_x} g(\varepsilon,\sqrt{2m\zeta})\,d\zeta\ldots\Leftrightarrow phonon.absorption\end{cases} \qquad Eq.\ 3\text{-}55$$

If a Fermi-Dirac statistic is applied for Eq. 3-55, the original Fermi (E_{f0}) level previously calculated for coherent electrons is not convenient for incoherent electrons, and a new quasi-Fermi level should be applied. The most plausible quasi-Fermi level to be used is described by the following equation,

$$E_f=\begin{cases}E_{f0}-\hbar\omega_{op}\ldots\Leftrightarrow phonon\ldots emission\\ E_{f0}+\hbar\omega_{op}\ldots\Leftrightarrow phonon.absorption\end{cases} \qquad Eq.\ 3\text{-}56$$

The density of electrons calculated in Eq. 3-55 is the incoherent local density of states calculated for unity probability of occupancy. Therefore, we have to weight the local electron concentration by a Fermi-Dirac statistics in which the quasi Fermi-level of Eq. 3-56 is applied.

It is worth-noting that the density of states calculated using Eq. 3-55 has to be devided by two to get only the positive x-momentum component (see also chapter 2).

In the above analysis the originally coherent local density of states are totally transformed into incoherent local density of states. This means that implicitly we assume a unity probability of scattering.

In order to find the various weighting terms, it is necessary to use again the comparison between the life time of electrons within the active zone (τ_e, τ_a) with respect to the characteristic time of scattering processes (T_e, T_a) which are directly derived from the scattering rate ($T_e = 1/\nu_e$, $T_a = 1/\nu_a$). Now, let us suppose that we have n_a and n_e electrons involved in the absorption and emission processes respectively per unit time.

Those, affected by the absorption process could be derived according to the following relation :

$$\frac{1}{T_a} = n_a\left[1 - \exp\left(-\tau_a / T_a\right)\right] \qquad \textit{Eq. 3-57}$$

In contrast, in case of avoidance we obtain,

$$\frac{1}{T_a^-} = n_a \exp\left(-\tau_a / T_a\right) \qquad \textit{Eq. 3-58}$$

Therefore, τ_a governs the weighting balance and one can check that when τ_a tends towards infinity all the involved electrons are scattered whereas the coherence is preserved when τ_a tends towards zero. From the above equation, the variable n_a can be eliminated and we get,

$$\frac{1}{T_a^-} = \frac{\exp\left(-\tau_a / T_a\right)}{T_a\left[1 - \exp\left(-\tau_a / T_a\right)\right]} \qquad \textit{Eq. 3-59}$$

Similarly for the emission process,

$$\frac{1}{T_e} = n_e\left[1 - \exp\left(-\tau_e / T_e\right)\right] \qquad \textit{Eq. 3-60}$$

$$\frac{1}{T_e^-} = \frac{\exp\left(-\tau_e / T_e\right)}{T_e\left[1 - \exp\left(-\tau_e / T_e\right)\right]} \qquad \textit{Eq. 3-61}$$

From T_a, T_a^-, T_e and T_e^- we deduce the various weighting coefficients denoted P_a^+, P_a^-, P_e^+ and P_e^-. The subscripts and superscripts referring to absorption (a), emission (e), collision (+) and avoidance (-).

$$p_e^- = \frac{\left(1/T_e^-\right)}{\left(1/T_e\right)+\left(1/T_e^-\right)+\left(1/T_a\right)+\left(1/T_a^-\right)} \qquad \textit{Eq. 3-62}$$

$$p_e^+ = \frac{\left(1/T_e\right)}{\left(1/T_e\right)+\left(1/T_e^-\right)+\left(1/T_a\right)+\left(1/T_a^-\right)} \qquad \textit{Eq. 3-63}$$

$$p_a^- = \frac{\left(1/T_a^-\right)}{\left(1/T_e\right)+\left(1/T_e^-\right)+\left(1/T_a\right)+\left(1/T_a^-\right)} \qquad \textit{Eq. 3-64}$$

$$p_a^+ = \frac{\left(1/T_a\right)}{\left(1/T_e\right)+\left(1/T_e^-\right)+\left(1/T_a\right)+\left(1/T_a^-\right)} \qquad \textit{Eq. 3-65}$$

Using the supply function model described in chapter 2 and the weighted local density of states discussed above, one can calculate the corresponding supply functions. In Figure 3.7 we plotted the purely coherent local density of states (a) calculated using the model discussed in detail in chapter 2 along with the associated momentum-less density of states (b). Also are depicted the coherent and incoherent supply functions with respect to the emission and absorption processes in Figure 3.8 and Figure 3.9.

Finally the tunnelling current is calculated according to following equation,

$$J = \sum_{a,e} \int SF_i^{LR}\, T_i^{LR} - SF_i^{RL}\, T_i^{RL} + SF_c\, T_c \quad d\varepsilon_x \qquad \textit{Eq. 3-66}$$

where the suffix i stands for incoherent and the suffix c for coherent. The summation is performed for both the absorption (a) and the emission (e) processes.

Before reporting the numerical results calculated in the framework of the model discussed in details above, let us now consider the space charge effect.

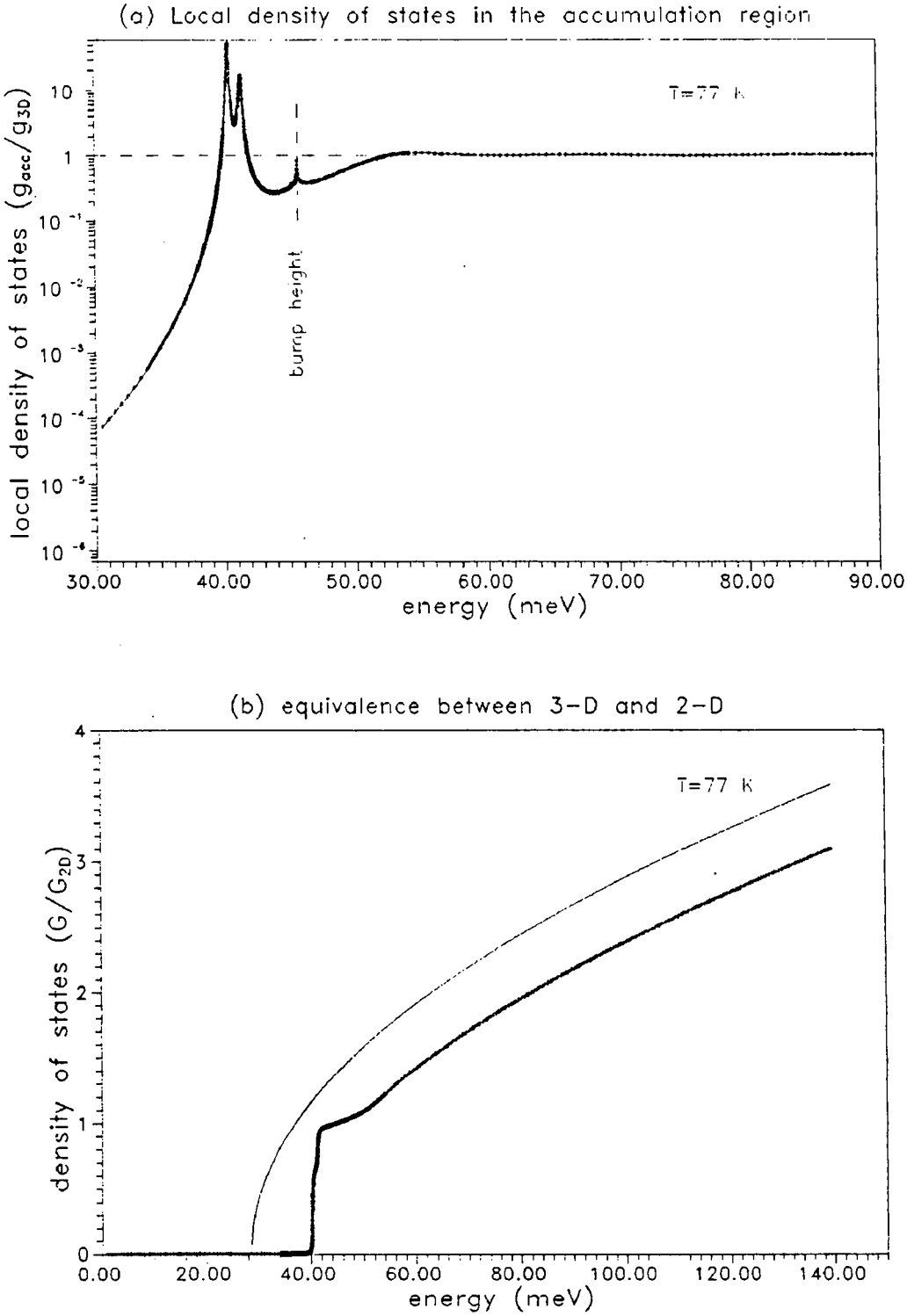


Figure 3.7 (a) the purely coherent local density of states calculated using the model discussed in the previous chapter, (b) the momentum-less density of states calculated by the integration of the above local density of states w.r.t the average momentum in the accumulation zone and a comparison with the equivalent 3-D one.

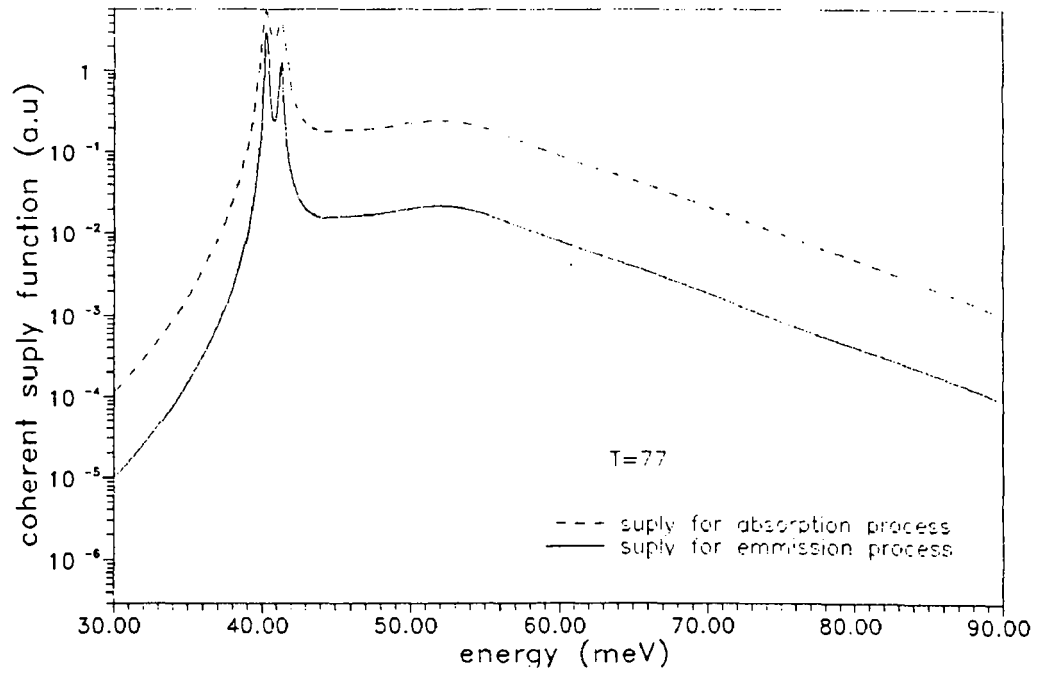


Figure 3.8 depicts the coherent supply function calculated using the model discussed above for both absorption and emission processes.

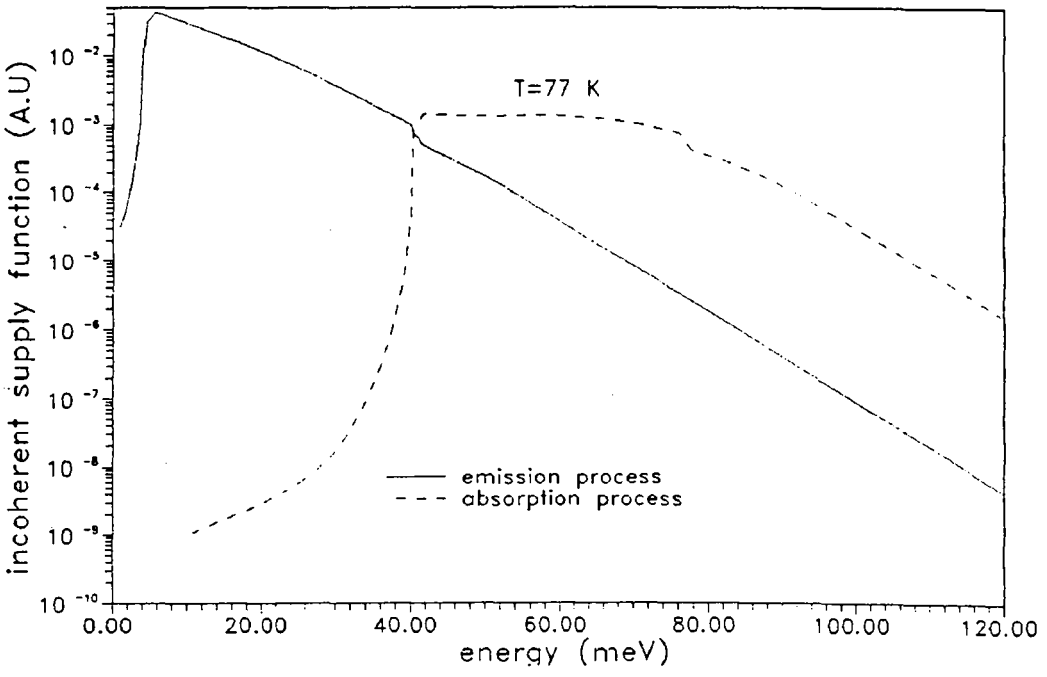


Figure 3.9 depicts the incoherent supply function calculated using the model discussed above for both absorption and emission processes.

3.7 Space charge effect :

The theoretical treatment of the space charge effect within a quantum structure is a tricky task. Ideally, a full quantum approach such as that used in Wigner Distribution Function (WDF) has to be performed. The reason are twofold, the first stems from the ability of such a theoretical approach to describe the conduction phenomena as a real transport process. Indeed, in quantum theory the difficulty of accounting for scattering is often alleviated by elaborating a perturbation approach on the basis of a pure coherent process. The WDF approach does not suffer from this limitation. In counterpart, it appears quite difficult to implement such a method for a device taken as a whole. In fact, such a method is very time consuming and the physics of phenomena is sometimes hidden by the complexity of the numerical treatment.

In the present work, in order to give a physical insight into the effect of the trapped charge we will made a crude assumption which permits us to derive a relationship between the trapped electrons and the current terms contribution including the scattering effect. We will show that, despite this approximation some of the main consequences of space charge effect notably the shift in the threshold voltage will be achieved.

It is also important to note that transport in resonant tunnelling structures with large spacers (which is our case) is space charge limited [16]. That is current and space charge are strongly coupled [16].

3.7.1 Calculation of the trapped charge:

Assuming that the electrons transfer after tunnelling, in the real space of the collector region with the ballistic speed $v(x,\varepsilon_x)=\sqrt{2m^*/\varepsilon_x(x)}$ in the x-direction where $\varepsilon_x(x)$ is the longitudinal energy at the position x.

From the definition of the current density we have,

$$dn(x,\varepsilon_x) = \frac{dJ_x(x,\varepsilon_x)}{q\,v(x,\varepsilon_x)} \qquad \textit{Eq. 3-67}$$

where $dJ_x(x,\varepsilon_x)$ is the x-component of the incremental current density.

The basic idea of the space charge model is to link the density of charge trapped in the well and consequently to introduce subsequent conduction band bending solved self-consistently with the local current. Previously, special attention was paid to derive forward and backward current contributions in the region where the electrons are bouncing back and forth between the confining barriers. In contrast, it is reasonable to assume that there is no wave function reflection in the collector region because there is no material discontinuity at this zone. For instance, such an assumption of a progressive wave is made to solve the Schroedenger equation using Runge-Kutta method [20]. The obvious linking between the region where a standing wave pattern is observed and those which can support a progressive wave is the current conservation.

Our primary goal is to calculate quantitatively the density of trapped charge which is subsequently introduced to Poisson's equation solver. We have the shape of this charge by means of the wave function, it remains to evaluate a normalisation constant which is the key issue. For this purpose we proceed by the following method :

3.7.1.1 Wave function normalisation to the current density:

Anywhere in the collector region the conservation of probability current density can be written as,

$$K(x,\varepsilon_x).\left|A^2(\varepsilon_x).\Psi^2(x,\varepsilon_x)\right| = C(\varepsilon_x) \qquad \textit{Eq. 3-68}$$

where $C(\varepsilon_x)$ corresponds to the probability current increment independent of x-position, $K(x,\varepsilon_x)$ is the wave vector, $\Psi^2(\varepsilon_x)$ is proportional to the electron density and $A^2(\varepsilon_x)$ is the coefficient of normalisation.

On the other hand, a quite general formula permits one to calculate the current intensity at ε_x from the corresponding transmission and supply function. Therefore, for a given temperature θ and bias condition V_c we have,

$$C(V_c,\varepsilon_x)=\frac{qm^*K_B\theta}{2\pi^2\hbar^3}T(V_c,\varepsilon_x)F(V_c,\varepsilon_x)\,d\varepsilon_x \qquad \textit{Eq. 3-69}$$

where the other symbols have their usual meaning.

If the effect of scattering on the space charge trapped in the quantum well could be neglected then, from Eq. 3-67 and Eq. 3-69 we get,

$$A^2(\varepsilon_x)=\frac{qm^*K_B\theta}{2\pi^2\hbar^3}T(\varepsilon_x)\,F(\varepsilon_x)d\varepsilon_x \,/\, K(x,\varepsilon_x)\left|\Psi^2(x,\varepsilon_x)\right| \qquad \textit{Eq. 3-70}$$

Since the product $K(x,\varepsilon_x)\left|\Psi_x^2(x,\varepsilon_x)\right|$ is constant everywhere in the collector region, It is arbitrary calculated at the collector electrode to be simply the propagation constant $K_c(\varepsilon_x)=\sqrt{2m^*(\varepsilon_x+V_c)/\hbar^2}$ at the collector electrode where the wave function is arbitrary chosen such that $\psi^2(x_{collector},\varepsilon_x)=1$.

Referring to Eq. 3-67 and the quantum definition of the probability density current $dJ(x,\varepsilon_x)$ we get,

$$dn(x,\varepsilon_x)=\frac{K(x,\varepsilon_x)\,A^2(\varepsilon_x)\left|\Psi^2(x,\varepsilon_x)\right|}{q\,v(x,\varepsilon_x)} \qquad \textit{Eq. 3-71}$$

Let us consider now a quantum region sandwiched between two heterointerfaces. Constructive and destructive interference between the forward and backward propagating waves will take place and the central issue is now whether such phase related phenomena affect the local electron distribution function. Mathematically, this means we have to distinguish between the modulus of the wave function as a whole or separately.

In the former we thus get,

$$\left|\psi^2(x,\varepsilon_x)\right| = \left|\Psi\ (x,\varepsilon_x)_f + \Psi\ (x,\varepsilon_x)_b\right|^2 \qquad \textit{Eq. 3-72}$$

where as in the latter we get,

$$\left|\psi^2(x,\varepsilon_x)\right| = \left|\Psi\ (x,\varepsilon_x)_f\right|^2 + \left|\Psi\ (x,\varepsilon_x)_b\right|^2 \qquad \textit{Eq. 3-73}$$

where the suffix (f) and (b) denote the forward and backward propagation.

The later (Eq. 3-73) stems from the fact that Eq. 3-68 is defined only for the regions of no-eddy currents or in other words in the zones of no-wave function reflections [17] (collector region).

In the following we are comparing systematically the two assumptions discussed above namely the role played by interference effect in the built-up of the trapped charge. Figure 3.10, is a plot of the space charge profile with and without interference effect calculated at a temperature 77 K. It can be seen that the main difference comes from the spatial profile. Accounting for phase interference effect, the maximum density ($n \cong 1.7 \times 10^{17} cm^{-3}$) is located approximately in the middle of the well reflecting the presence probability. Suppressing the phase interference effect leads to a rather constant charge profile. Figure 3.11 permits us to investigate the effect of biasing potential in the vicinity of the threshold voltage for NDC effect here between 0.4 and slightly above 0.5 Volts. In terms of carrier concentration it can be seen that the discrepancy are relatively low. Moreover, as expected the carrier density increases versus bias voltage. This is a direct consequence of the increase in the current density which is the origin of the supply.

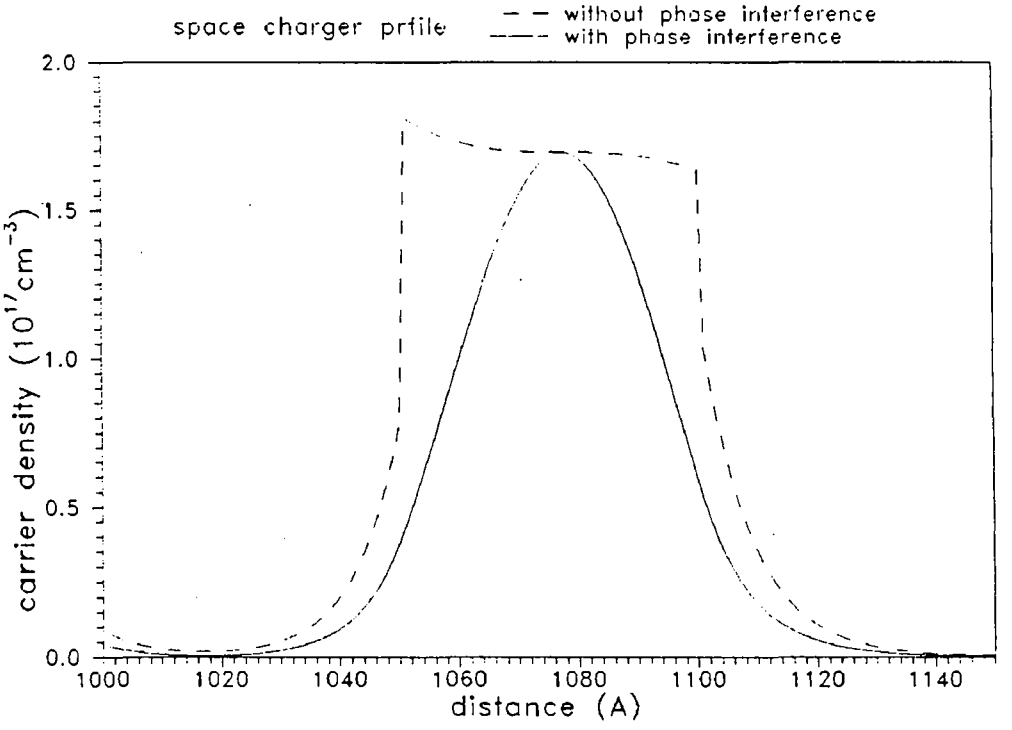


Figure 3.10 depicts the space charge profile with phase interference effect and without phase interference effect (T=77K).

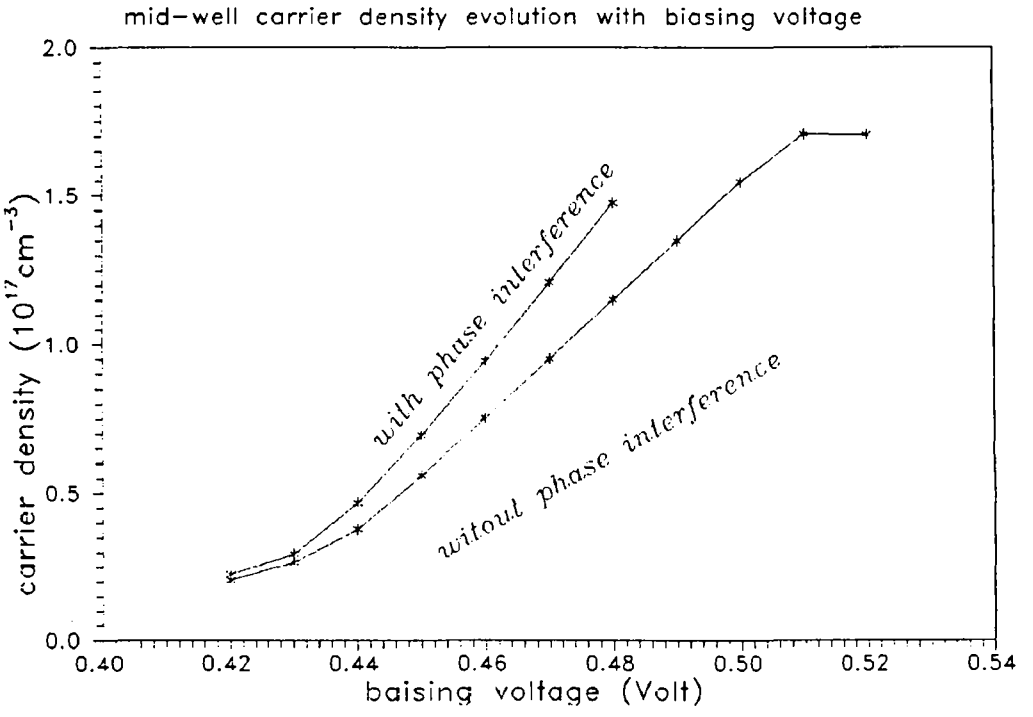


Figure 3.11 depicts the space charge evolution with applying potential with phase interference effect and without phase interference effect (T=77K)

Beyond the differences pointed-out in Figure 3.10 and Figure 3.11 the fact of considering the phase interference or not is still an open question.

3.8 I-V results and comparison with experiment :

In this section, we are now reporting the results according to various degree of assumptions. Figure 3.11 depicts the calculated and measured I-V characteristics when the coherence breaking is included with (dashed) and without (solid) space charge effect. As a matter of illustration, Figure 3.13 shows the measured I-V curve with a compressed voltage scale to give a clear view of the LO replica. The first remark concerns the relatively good agreement between the measured and the calculated peak current densities with a value typically of 500 A/cm^2. On the other hand, it can be seen that the phonon replica shown about a 0.65 Volts is also well described theoretically. It is seen that the agreement between threshold voltage V_{peak} is only achieved provided the space charge effect is included. In counterpart, for the both situations, it seems that a broadening of the resonant current curve is slightly under estimated. At last, a kink effect is apparent at low voltages ($\approx 0.35\ V$) both for the experimental and theoretical curves.

Figure 3.14 shows the various contributions to the calculated total current plotted in solid line. The first current anomaly at ($\approx 0.4\ V$) reflects the current contribution resulting from the cross-energy between a very weak resonance in the quasi-continuum and the quantum-well resonant energy (ε_{rw}). This feeding mechanism is reproduced owing to the new definition of both the local density of states and the supply function. The later is well defined over all the states namely the strictly confined and highly delocalised states.

Also in Figure 3.14, it can be seen that the incoherent absorption plotted in long-dashed line is not of sufficient magnitude to be responsible of this conduction anomaly. As a general rule, its influence is of second order. This is not the case of the incoherent emission which is responsible of bump in the current density noted at ($\approx 0.65\ V$). Finally, considering the coherent component, it is seen that this

contribution dominates the peak current whereas in current valley region, its contribution is of minor concern but is not zero. Obviously, the description of the peak to valley current ration is now much more realistic as compared with a fully coherent approach.

It remains to explain how a closer agreement could be achieved if the space charge effects are included. To this aim Figure 3.15 is a close-up of the voltage range of interest. One can note that the general trends by taking space charge into account is to shift the peak voltage to a higher value. Indeed, by accumulating a non-negligible charge density within the well, we subsequently introduce an electric field gradient which minimises the voltage drop over the quantum region. As a consequence, it needs more voltage to satisfy the threshold condition. Moreover, this shift is more important when no phase effect is introduced due to the higher sheet carrier density (see Figure 3.11). Nevertheless, it seems hazardous to select one of the two schemes on the basis of experimental agreement. The temperature of 77 K was chosen for illustrating mainly the phonon-replica.

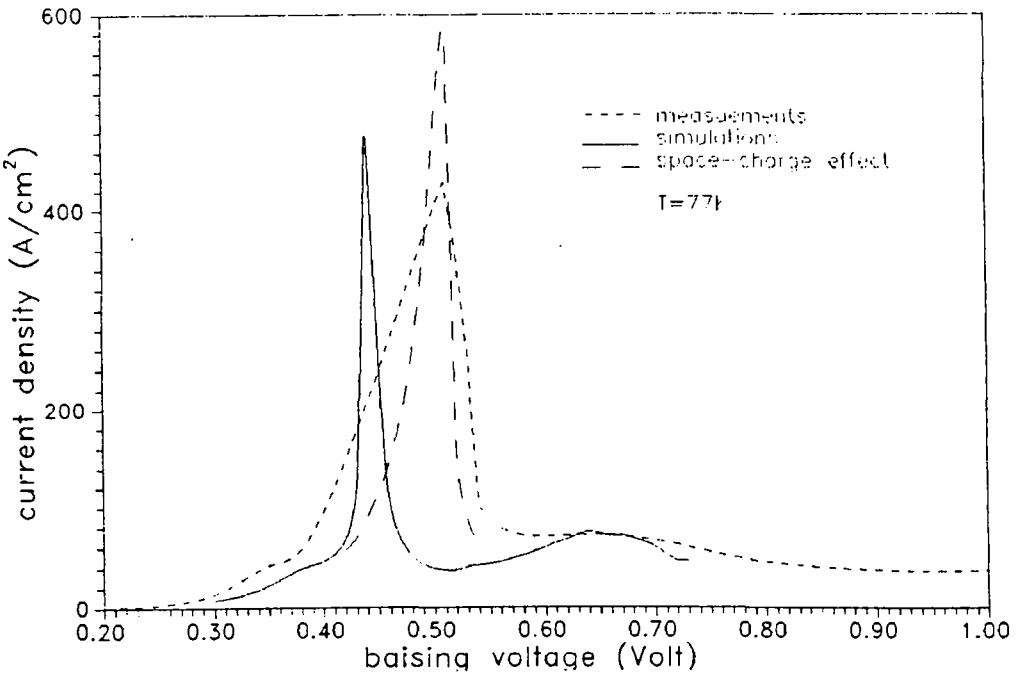


Figure 3.12 depicts both the measured and the simulated I-V characteristics calculated according to the above discussed model.

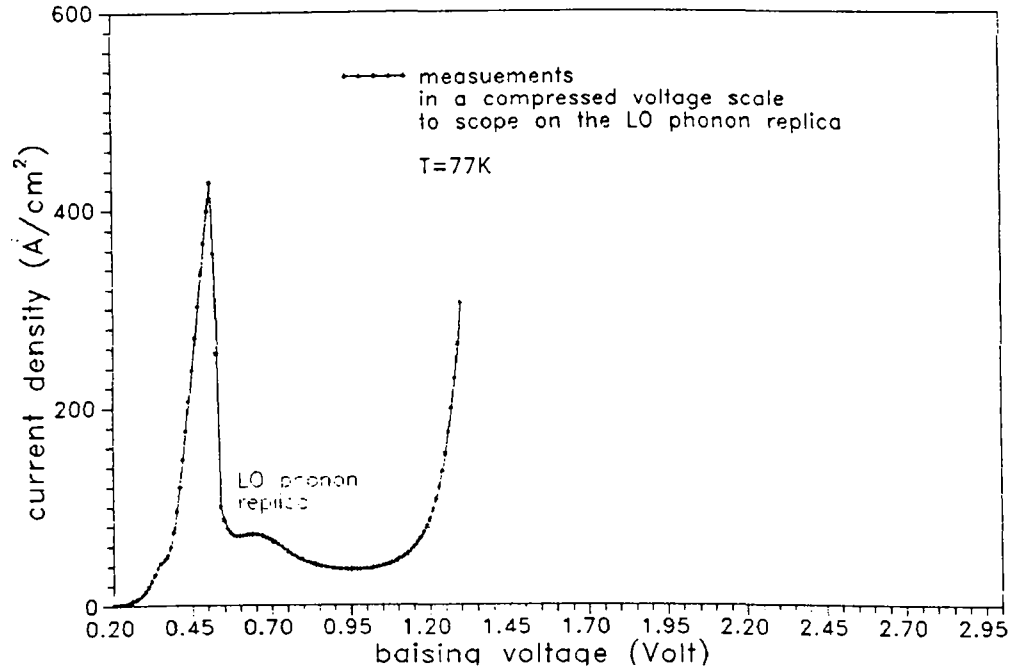


Figure 3.13 measured I-V characteristics [1].

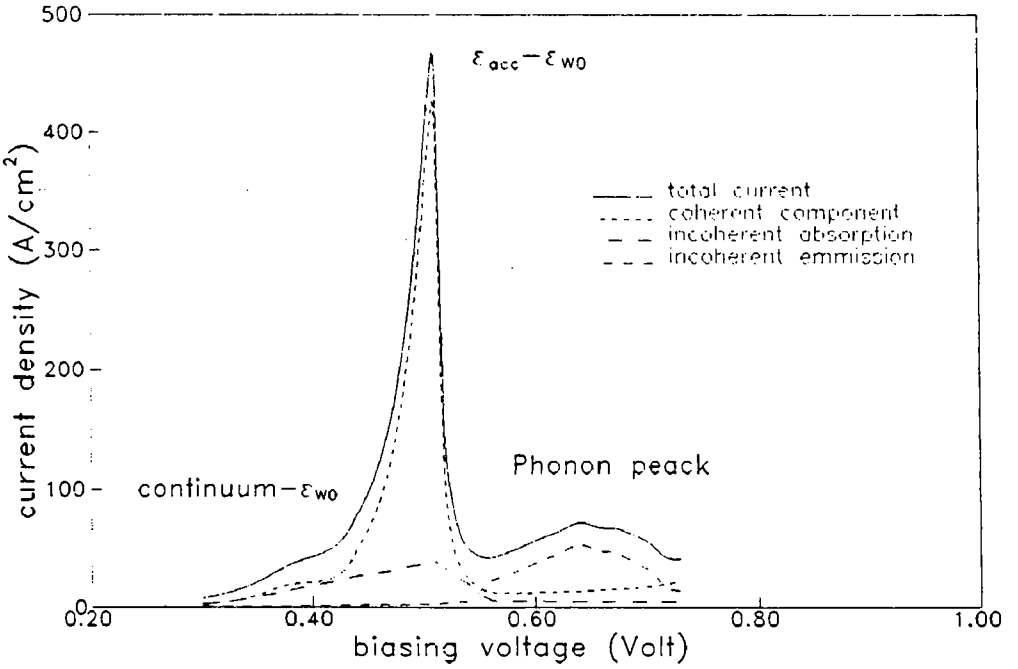


Figure 3.14 depicts the different current components (with space-charge effect at 77K)

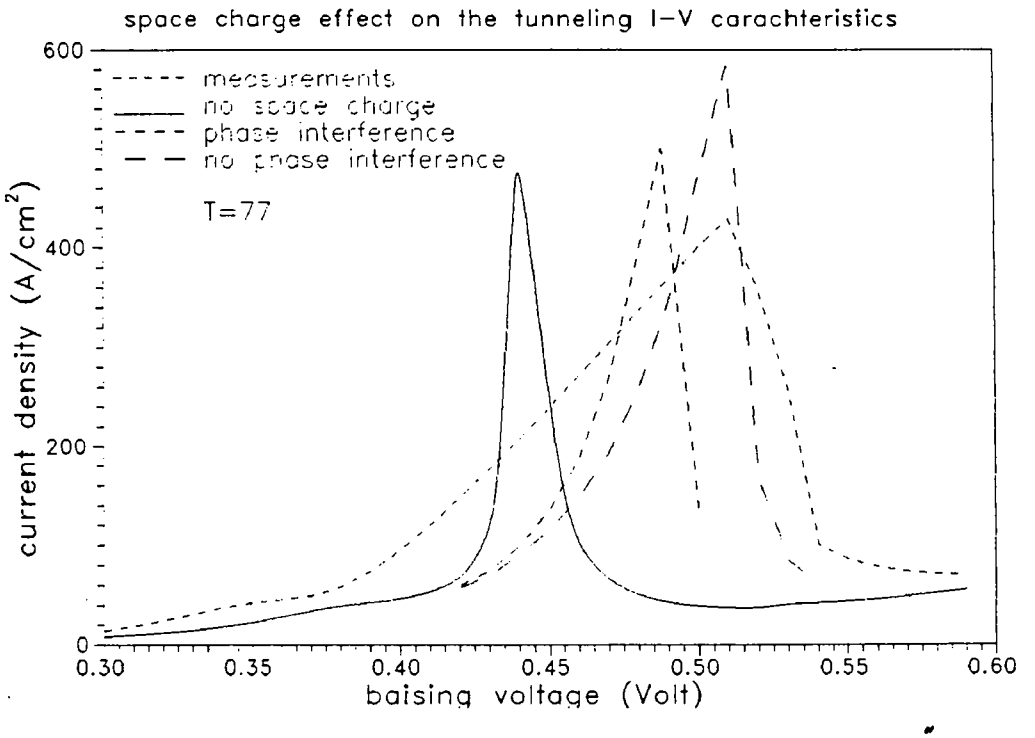


Figure 3.15 depicts the effect of space charge on the I-V characteristics at 77K.

With the aim to investigate the temperature dependence it seems interesting to calculate the I-V characteristics under various approximation degrees at 300 K. The results are shown in Figure 3.16 and Figure 3.17. As expected, and in agreement with experiment we observed a dramatic degradation in the peak to valley ration with the disappearance of the phonon-replica. Otherwise, we observed the same kind of trends with perhaps a little-bit disagreement concerning the peak voltage. for the plots of space charge effect (with and without phase interference) the curves in the negative differential conductance region were not plotted for sake of clarity but they follow the same evolution of the solid line in the considering figure.

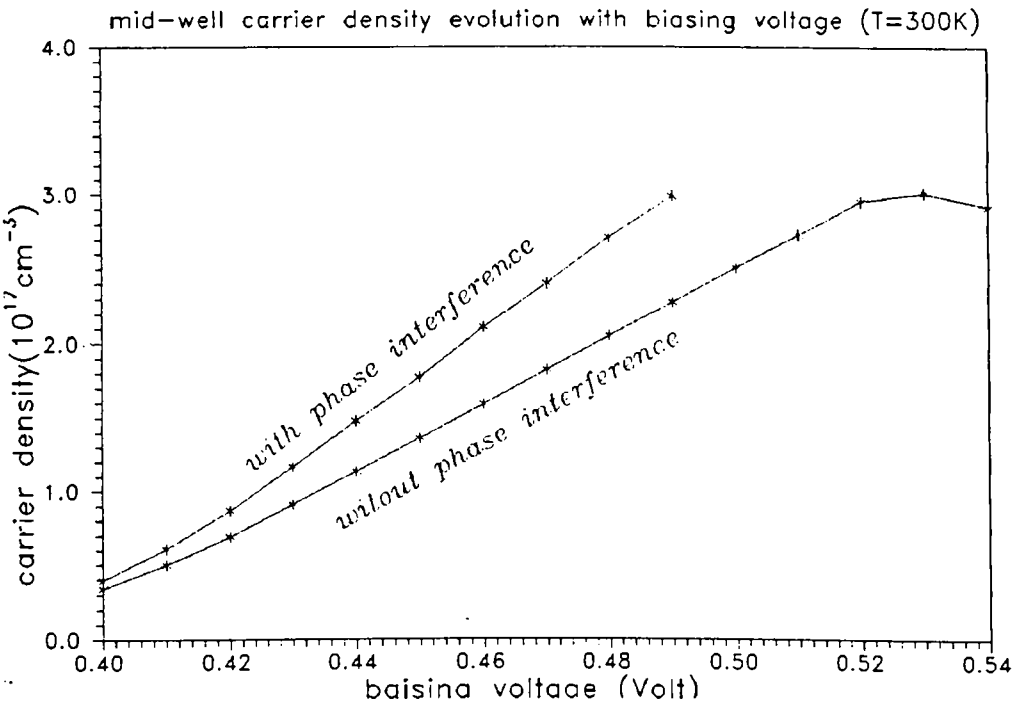


Figure 3.16 depicts the space charge evolution with applying potential with phase interference effect and without phase interference effect(T=300K)

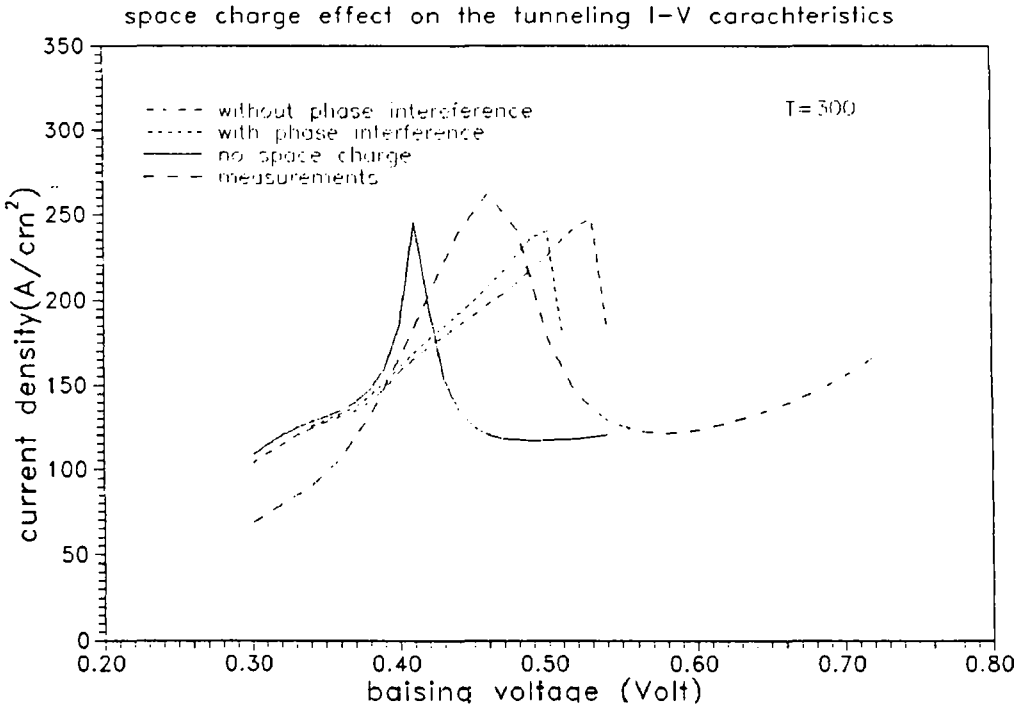


Figure 3.17 depicts the effect of space charge on the I-V characteristics at 300K.

3.9 Systematic study with temperature :

Before comparing simulated and measured data, let us briefly consider the experimental conditions. In contrast to the measurements performed solely at liquid Nitrogen temperature by dipping the samples, systematic measurements versus temperature requires to mount the device onto a test fixture cooled down in a cryostat. In Figure 3.18 we plotted in solid lines the peak and valley currents as a function of temperature. Initially, the peak to valley current ratio is about 10 : 1. At increasing temperature, this ratio degradates gradually to reach about 3 : 1 at room temperature. These dc data compare favourably to the best results achieved so far with similar $GaAs/Al_{0.3}Ga_{0.7}As$ DBH's. In further details, it can be shown that the valley current increases following approximately a linear relationship as a result of increasing scattering probability with temperature. In contrast, the peak current exhibits a quasi-exponential decrease.

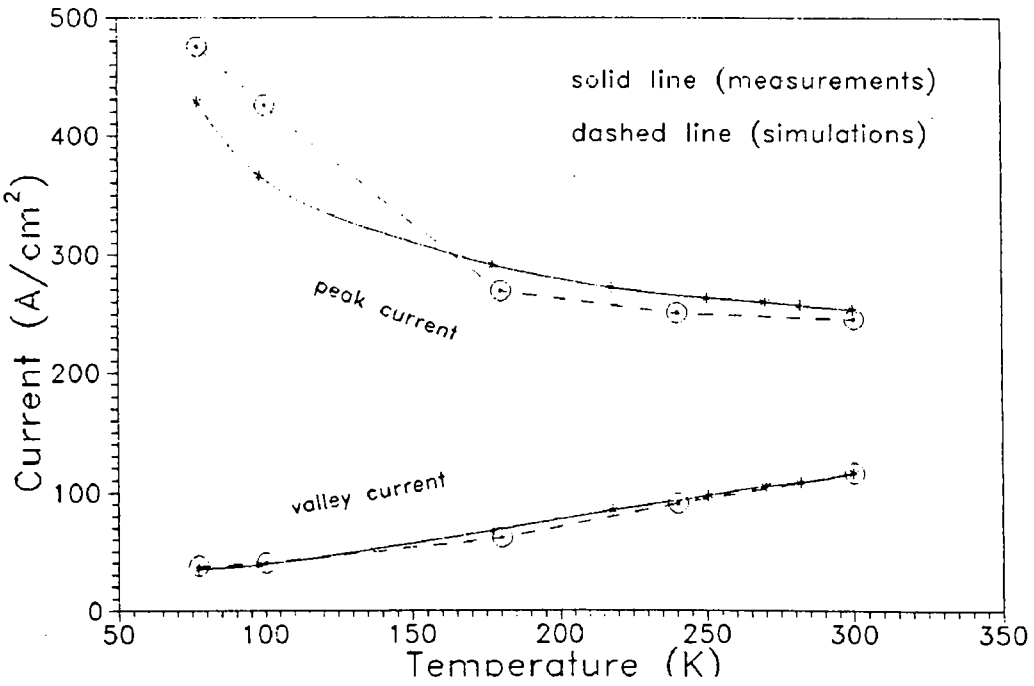


Figure 3.19 depicts the evolution of the peak and valley currents with temperature

The simulated data are plotted in a dashed lines. Overall, a good agreement is achieved. It is important to state that the above calculation are performed without concerning the space charge effect and the effective mass re-normalisation. We previously learned that the main effect of space charge concerns the relevant peak voltage without significant modification in the current contrast. To our knowledge, this is the first time that the peak to valley current ratio is described satisfactorily over a rather broad temperature range.

Another issue in connection with temperature studies, concerns the use of the second derivative of the I-V relationship. We address this issue in the next section.

3.10 Second derivative of the I-V relation :

One of the main drawbacks of characterisation technique based on the measurement of the I-V characteristics stems from the fact that there is no direct information about the quantum transmission probability. Let us recall that the total current implies an integration over energy of the supply function times the transmissivity. Recently, a very elegant way to overcome such a limitation was proposed by Sakaki [21]. The basic idea was to use the second derivative of the current with respect to the voltage at the onset of resonance process. Moreover, if the temperature is sufficiently low in order to minimise the thermal tail of the supply function, the authors of [21] have shown that some indications of the degree of broadening of transmission can be estimated. Previously, we have discussed about the relation between scattering probability and the broadening in transmission. Ideally, the characterisation should be conducted at liquid Helium temperature (4.2 K). In the present work, unfortunately, it was not possible to cool down the sample at so low temperatures. Nevertheless, it seems to us interesting to use the same approach based on the second derivative to have a first insight into the importance of the incoherent process.

The results of this study are summarised in Figure 3.20 and Figure 3.21. For the former we compare the bias dependence of the second derivative of current density with respect to voltage at 77 K (Figure 3.20 a) and at 300 K (Figure 3.20b) for the measured (dashed line) and calculated (solid line) data. The general trends are almost described. The main discrepancy stems from a voltage shift between 30 and 50 mV. The magnitude fits quite well for the first extrema whereas the error is within 30% for the third anomaly. We have to stress that this is the first part of the curves typically below 350 mV which is interesting on the basis of the arguments outlined before.

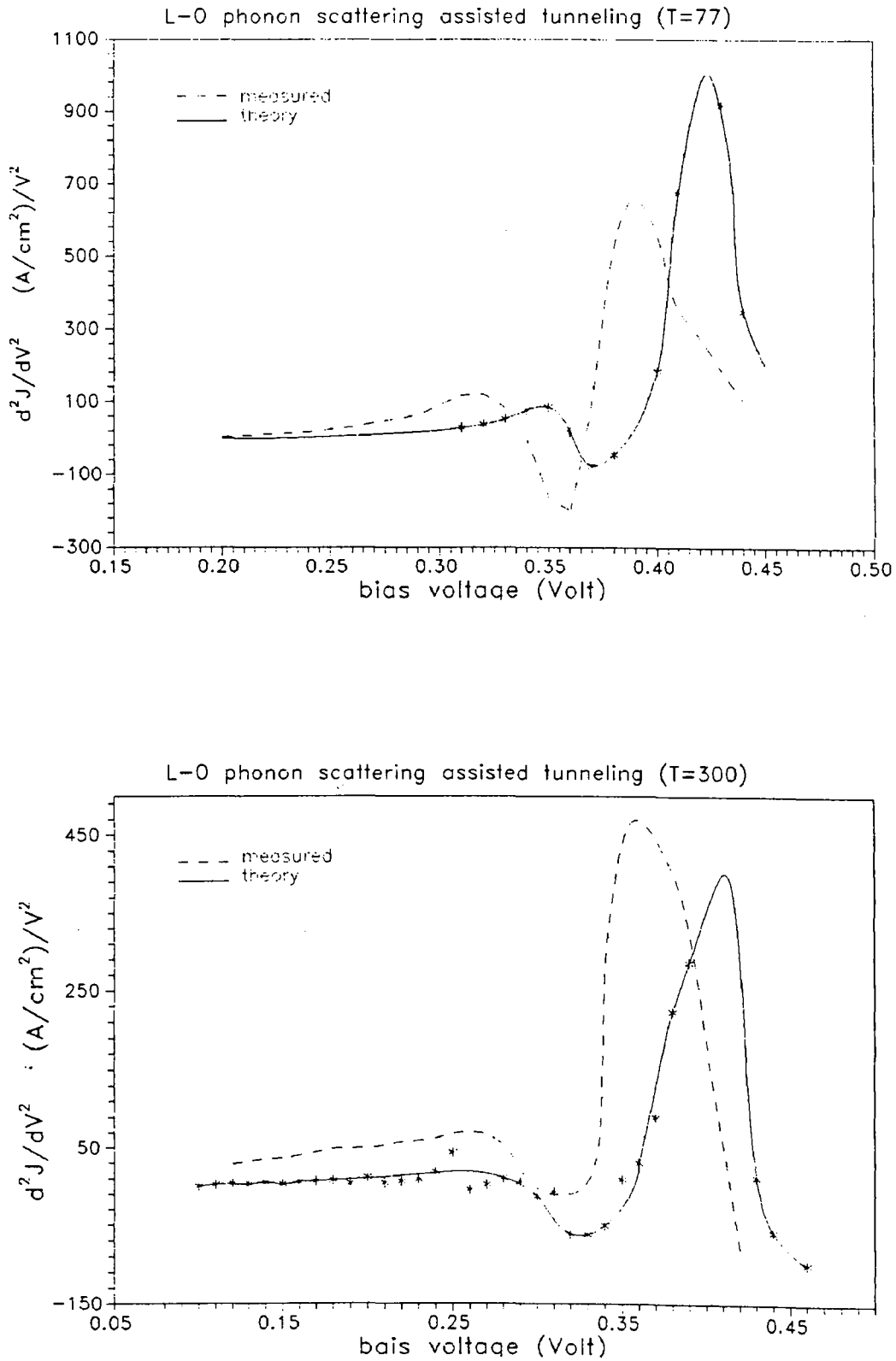


Figure 3.20 depicts a comparison between the measured and the calculated second derivative of the I-V characteristics (a) at T=77 K (b) at T=300 K

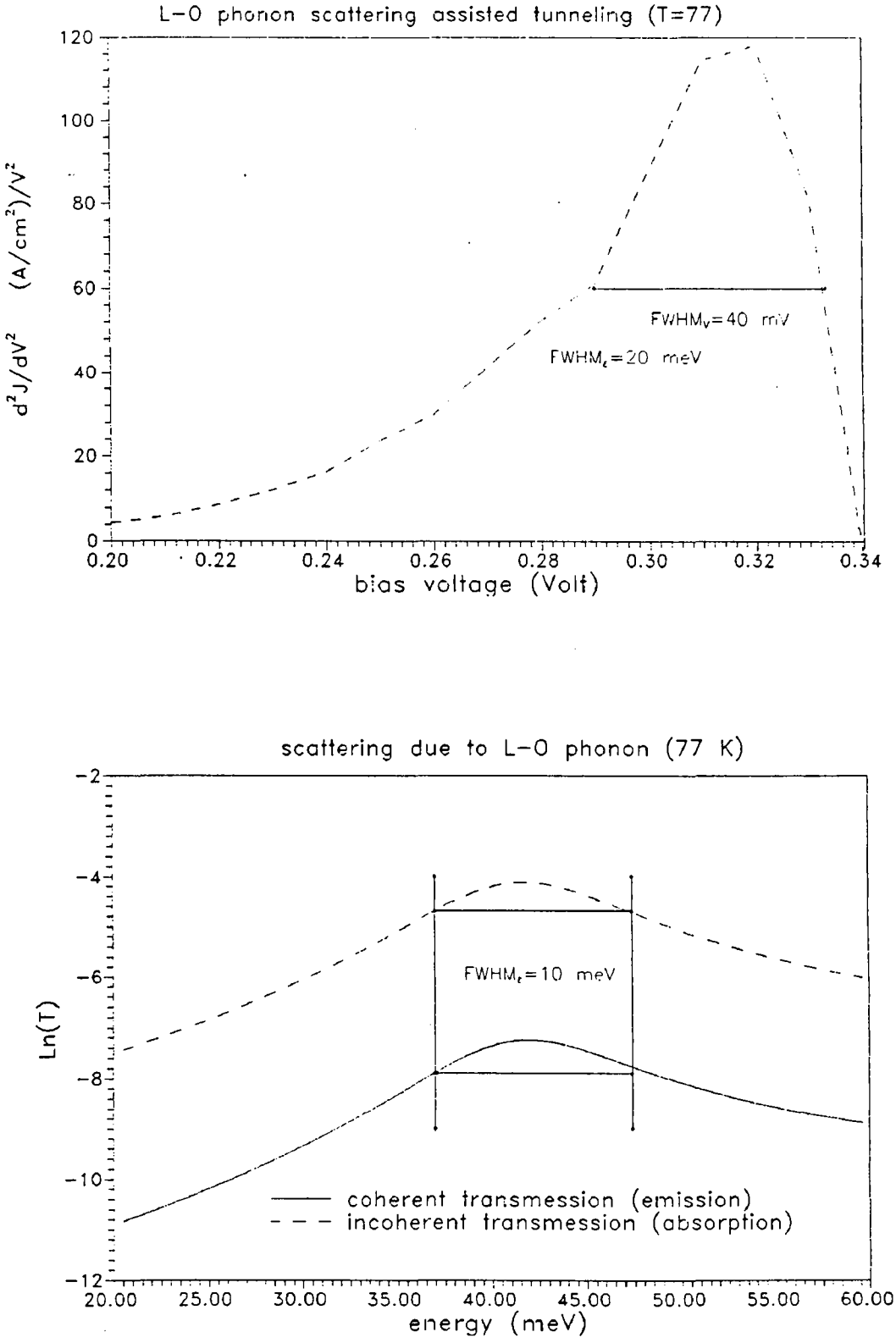


Figure 3.21 (a) measured transmission width determined by the second derivative method [21], (b) is derived from our simulation results.

The variation of d^2J/dV^2 with a sign change results directly from the bump pointed out in the I-V curves at the onset of the conduction which is attributed to the transition between the state of quasi-continuum and the 2D injection. At room temperature the same behaviour is achieved with an evolution which is somewhat less pronounced. It is certain that Sakaki analysis under this temperature condition is no longer valid. Therefore, we plotted in Figure 3.21 a zoom-in to the bias range between 200 and 340 mV at liquid Nitrogen temperature.

The Full Width at Half Maximum (FWHM) with respect to bias voltage taken as the independent variable is 40mV. It remains to correlate this voltage broadening to the energy one. A simple rule, concerning bare DBH's with symmetric barriers, is to assume that a shift in the voltage drop across the structure results in one-half the shift in quantum energy level (in electron volts). From this scaling rule the FWHM with respect to energy ($FWHM_\varepsilon$) 20 meV. This experimental broadening can be compared to the value calculated from theoretical transmission. In order to illustrate this issue, two cases have been considered, (i) coherent transmission component in an overall process involving phonon emission and (ii) incoherent transmission component in phonon absorption process. Both curves exhibits a similar broadening with the same order of magnitude (10 meV). The measured and the calculated data differ of solely a factor of two which is very encouraging in view of assumptions made both in theory and experiment.

3.11 Conclusion :

In this chapter, we have studied the scattering assisted tunnelling. The effect of scattering is included not only in the transmission probability but also into the injection conditions through a novel probabilistic model. In connection to the new formalism described in details in chapter 2, we have succeeded for the first time to define four probability rates associated with each possible scattering assisted supply condition. The associated phonon energy is partitioned equally-likely between transverse and longitudinal energies. The scattering effect is assumed to causes a partial loss in the wave function coherence. The transmission probability is self-consistently solved along with the equation defining the finite life time in the well. The later time reflects the finite reflections characteristic time required to construct the resonant wave function in the quantum well.

In addition, the transmission probability is calculated using a new method we called AEES based on Airy function solutions. This is done to achieve both the required accuracy and generality while preserving the simplicity of the exponential method.

The above model, is systematically applied to simulate the resonant tunnelling diode fabricated in our group of research and which characterised in the present work as a function of temperature. The comparison between the simulated results and the measured data are surprisingly in a very good agreement. This good agreement is also shown to be sustained for a broad range of operating temperature conditions.

Using the above novel model, we have described successfully the main anomalies in the I-V characteristics notably, the onset-bump, the main resonant current and the post-resonant phonon replica. The first reflects an injection condition corresponds to an intermediate state between the extended 3D and the confined 2D-injection conditions.

3.12 References :

[1] P. Mounaix, J. M. Libberecht and D. Lippens, electron transfer between two coupled quantum wells in a resonant tunnelling diode structure, Solid-State Electronics vol. 38, no.11, pp. 1899-1904, 1995.

[2] Francois Chevoir and Borge Vinter, Calculation of phonon-assisted tunnelling and valley current in a double-barrier diode, American Institute of Physics, 1859, (1989).

[3] Francois Chevoir and Borge Vinter, Scattering-assisted tunnelling in double-barrier diodes : Scattering rates and valley current, The American Physical Society Vol.47, Number 12, 7260, (1993).

[4] T. Wiel and B. Vinter, Calculation of phonon-assisted tunnelling between two quantum wells, J. of Apll. Phys. 60(9), p.3227, (1986)

[5] C. R. Leavens and G. C.Aers in : scanning tunnelling microscopy III, (Springer Berline 1993) p. 105.

[6] X. Oriols, F. Martin and J. Suñé, implications of the non-crossing property of Bohm trajectories in one-dimensional tunnelling configurations, Phys. Rev. A, 54, p. 2594, 1996.

[7] X. Oriols, F. Martin and J. Suñé, oscillatory Bohm trajectories in resonant tunnelling structures, Solid State Comm. Vol. 99, pp. 123-127, 1996.

[8] D. R. Rode, Semiconductors and semimetals, Vol. 10, Academic Press, 1975

[9] K. Hess, IEEE Trans. Electron Dev. ED.28, 937, 1981

[10] A.D.Stone and P.A.Lee, Effect of inelastic processes on resonant tunnelling in one dimension, Phys. Rev. Lett. 54, 1196 (1985).

[11] O.Vanbesien and D.Lippens, Solid State Electron. 32, 1533 (1989).

[12] Yasuhito Zohta and Hiroshi Ezawa, Effect of inelastic scattering on resonant tunnelling studied by the optical potential and path integrals, J.Appl. Phys. 72(8), 3584, (1992)

[13] Yasuhito Zohta and Tetsufumi Tanamoto, Improved optical model for resonant tunnelling diode, J.Appl. Phys. 74(11), 6996, (1993).

[14] Yasuhito Zohta and Tetsufumi Tanamoto, Valley current of three-dimensional resonant tunnelling diode studied by the improved optical model, J.Appl. Phys. 76(4), 2343, (1994).

[15] J. Leo and A. H. MacDonald, Phus. Rev. Lett. 168, 779 (1968).

[16] Patrick Roblin, Robert C. Potter and Ayub Fathimulla, Interface roughness scattering in AlAs/InGaAs resonant tunnelling diodes with an InAs sub-well, J. Appl. Phys. 79 (5), 1996.

[17] Albert Thomas Fromhold, Jr. Quantum Mechanics for Applied Physics and Engineering, ACADEMIC PRESS p248, p85, 1981

[18] D.J. Fishfer and C. Zhang, Electron tunnelling lifetime of a quasi-bound state in a double-barrier resonant tunnelling structure, J. Appl. Phys. 76(1), 606, (1994).

[19] A.D.Stone and P.A.Lee, Effect of inelastic processes on resonant tunnelling in one dimension, Phys. Rev. Lett. 54, 1196 (1985).

[20] Louik De ST Pool, theses

[21] Masahiro Tsuchiya and Hiroyuki Sakaki, Tunnelling spectroscopy of resonant transmission coefficient in double barrier structure, J. J. App. Phys., Vol. 30, No. 6, p.1164, 1991.

CHAPTER 4

4.20 Table of contents :

4. Study on tunnelling transport in lateral field-effect-induced quantum wells :

4.1 Introduction :

In the previous chapters we have discussed the tunnelling phenomena in vertical heterostructures, investigating the two dimensional injection effects along with the scattering assisted tunnelling. All the previous analysis are based on the solution of Schroedinger equation in one dimension. The character of uni-dimensionality of the solution stems from the assumption of neglecting the fringing effects. The later permits the suppression of the spatial dependence of the potential function in the transverse direction with respect to current direction. On the contrary, the lateral structures, results in an electrostatic induced potential whose space variation depends at least on two directions. For example, the planar resonant-tunnelling field-effect transistor (PRESTFET) [1][2][3][4] results in X and Y dependence. On the other hand, for a grid-gate lateral surface super lattice (Grid-Gate LSSL) [1][5], accounting for X, Y and Z dependence is necessary.

First of all, before investigating the tunnelling transport in these lateral structures, a special attention has been paid to find out the most realistic potential function, the relevant Fermi-level and the injection-mode. Also, it is important to mention that the PRESTFET fabricated by K. Ismail and co-workers [1] will be the test vehicle through out the study of this chapter.

This test structure is a dual nano-gate structure in such a way that a resonant tunnelling transport can be demonstrated. Beyond the description of the 2D character, the novelty of the study, presented here, stems from accounting for the quantum transport effect. Indeed, localisation effect in MODFET's, yielding a discretisation of energy levels, was early recognised [6][7].

This effect could be taken into account in the present work but this is not our main goal. In the present work the *lateral* quantum-sized and tunnelling effect are more specially addressed.

4.2 Notes on device design and operating conditions :

We now introduce the typical parameters of the epi-layer from which the device has been fabricated. On the other hand, we discuss of various device guide lines along with operating conditions. First of all, it is worth mentioning that the operating temperature for observing quantum lateral effect has to be low. In most cases, the measurement were performed at liquid Helium but further cooling the device could be useful. The reason is quite simple. In fact, in contrast to heterostructure potentials, the electrostatic induced potential exhibits a quite opened parabolic-like shape. From the technological side, this can be understood by the difficulty to fabricate dual gates on the nano-meter scale also in close proximity. Accordingly, the quantum levels corresponding to the quantisation in the lateral direction are closer in energy and it is imperative to decrease the operating temperature to discriminate them.

On the other hand, any additional scattering effect tends to smear-out the expected resonant feature. We mentioned in chapter 3 that the ionised impurity scattering dominates the mobility of carriers at very low temperature. Also, by using modulation doping concept in heterostructures, further improvements in the asymptotic mobility at low temperature can be achieved.

In brief, the modulation doping is achieved by the growth of a doped wide bandgap material over the top of an undoped narrow bandgap material. Figure 4.1 shows schematically the growth sequence of a typical MODFET in the GaAs based system. The difference in electron affinity, associated with the two materials, results in a conduction band discontinuity at the heterointerface. Due to doping gradient, the electrons diffuse from the wide to the narrow bandgap material. This diffusion process gives rise to a strong space charge effect resulting from the formation of a dipole. The resulting induced electrostatic force counterbalance the diffusion process and an

equilibrium is reached. The latter is characterised by a constant Fermi-level across the heterointerface.

The separation of the diffused electrons from their parent donors minimises the ionised donor scattering mechanism on one hand, and creates as seen previously, a strong electrostatic field normal to the heterointerface on the other hand. This field, results in conduction band bending forming a quasi-triangular potential well. The diffused electrons whose location is displayed in dashed line in Figure 4.1 are consequently confined in close proximity to the heterointerface in the form of a quasi two-dimensional electron gas (2DEG).

The minimisation of scattering mechanisms maximises the electron mobility along with the quantum coherence in the active region. In addition, a further enhancement in the 2DEG transport properties is achieved by further separating the supplying- and the current carrying-layers by a wide bandgap undoped spacer layer.

Finally, the epitaxial sequence is completed by the growth of a highly-doped capping layer for subsequent Ohmic contact deposition.

Concerning the choice of a material system, ideally an undoped very narrow gap such as InAs should be selected with AlSb forming the wide band gap. However, the epitaxial quality of heterointerface is also of major concern. Therefore, most of successful reports that have been so far obtained, make use of AlGaAs/GaAs compounds. The rules of optimisation of Aluminium concentration at such low temperatures are out of the scope of this study. Nevertheless, let us recall that owing to DX-centres in connection with the cross over of $\Gamma-$ and $X-$valley, samples with x ranging between 0.25 and 0.3 have to be grown. With respect to the thickness and the doping of the highly doped GaAlAs layers there exists numerous degrees of freedom.

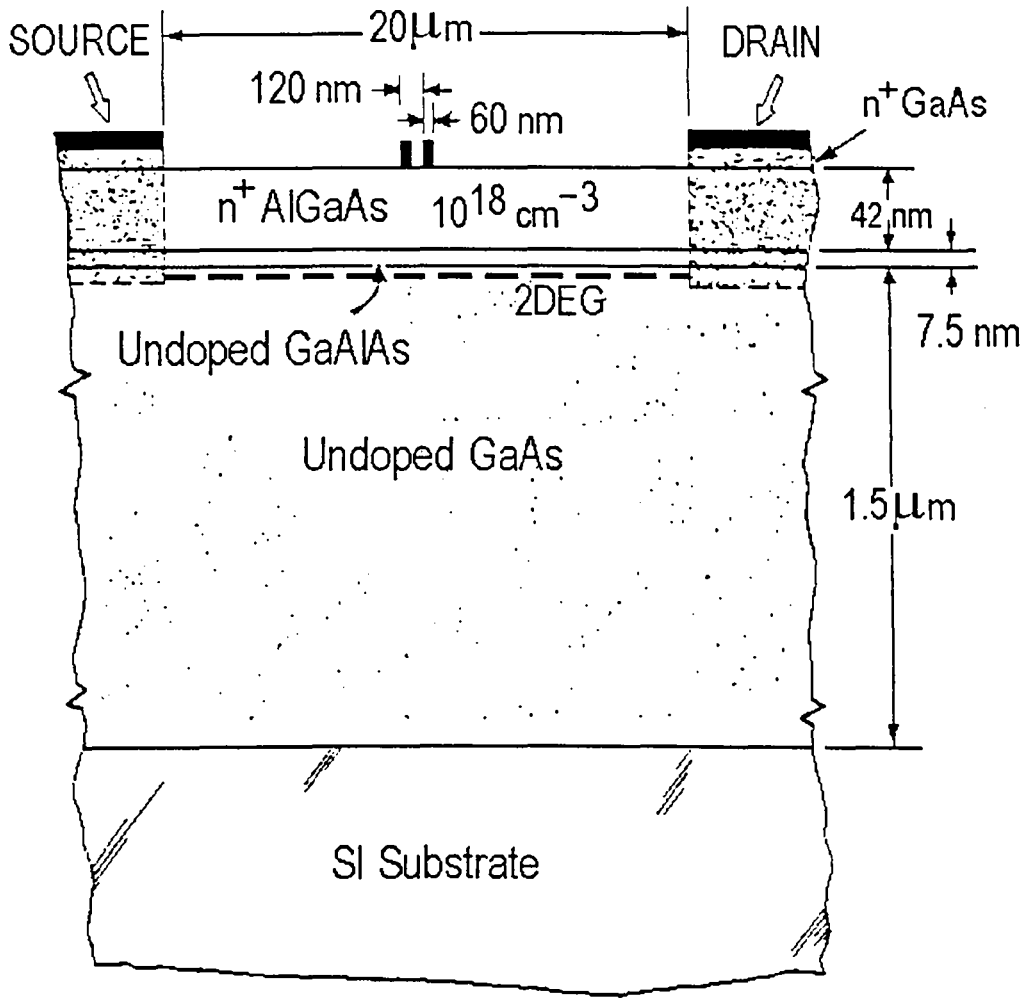


Figure 4.1 schematic cross-section of the resonant tunnelling dual-gate MODFET fabricated and characterised in [1] (after [1])
The aluminium percentage is between 0.25 and 0.3 %.

In the following, we report on the PRESFET fabricated in [1] with the following design data namely :

- 42 nm-thick $Al_{0.25\sim0.3}Ga_{0.75\sim0.7}As$ doped to 10^{18} cm^{-3}.
- 7.5 nm-thick $Al_{0.25\sim0.3}Ga_{0.75\sim0.7}As$ spacer layer.
- 1.5 μm-thick GaAs undoped layer.

The lateral scale is determined by the gate length, here 60 nm for both gates. Also for the inter-gate separation the same dimension is kept. We would like to mention that, despite the advance in high resolution electron beam lithography, no further break through has been obtained with respect to the shrinking of the gate width. This means that in the present status of technology it seems difficult to overcome the need of low temperature operation. At last, the source drain separation was fixed to 20 μm. consequently much wider than the active region dimensions.

For the Device Under Test (DUT), $Al_xGa_{1-x}As$ and GaAs are used as the wide-bandgap and narrow-bandgap materials respectively. The doping level is optimised to 10^{18} cm^{-3}. The optimisation criteria are based on the maximum doping level for which the gate leakage tunnelling current and traps in the AlGaAs could not affect the device performance at 4.2 K to a very large extent.

4.3 Notes on device characterisation :

The cross-section of DUT is shown in Figure 4.1. Let us summarise hereafter the experimental I-V characteristics which will support the validity of the theoretical analysis performed in this chapter.

At room temperature, the device was just at the limit of a normally on condition. Indeed, only 50 mV is needed for setting on the conduction. The positive voltage swing on the gates was limited to 0.6 V by the Schottky-contact current leakage. The room temperature transconductance was around 150 mS/mm. At 4.2 K the device threshold voltage for conduction onset was shifted up to 0.15 V, and the peak

transconductance was increased to 350 mS/mm. There was no persistent photoconductivity observed, and hence the device can be tuned by changing the optical intensity of a red Light Emitting Diode (LED). The authors measured the source-drain current I_{DS}, as a function of gate bias V_{GS} with both gates connected together. The measurements of resonant tunnelling features (Figure 4.2) were performed at 4.2 K in the dark at a drain-source bias V_{DS} of 0.2 mV.

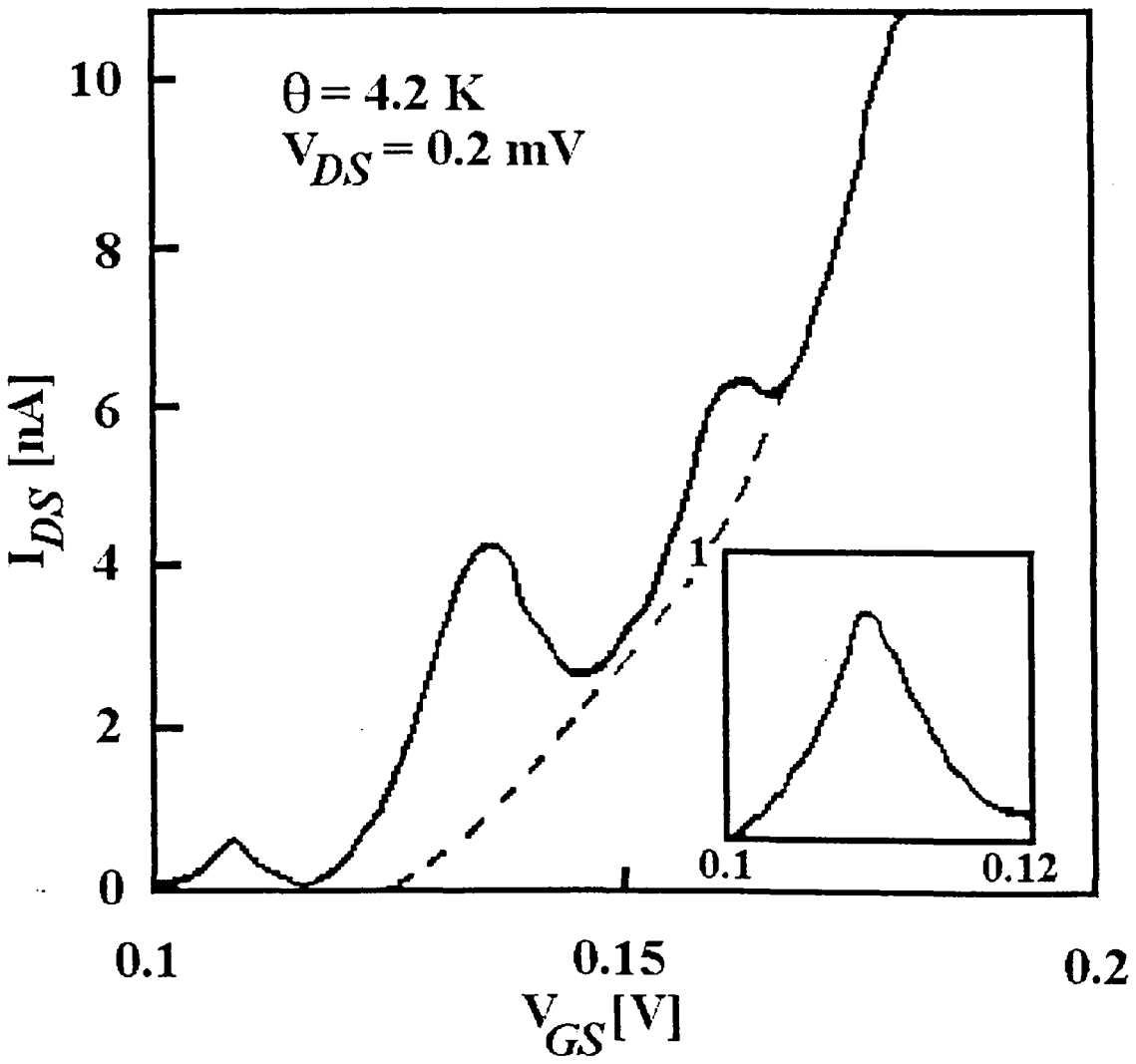


Figure 4.2 drain-source current as a function of gate bias, after [1].
(the inset is a zoom-in for the first resonant peak). Both gates are connected together.

No gate leakage current was detected down to the measurement limit (10 pA). In Figure 4.2, a clear structure was observed below threshold (~ 0.15 V). Three resonant peaks were superimposed over the exponential characteristic subthreshold regime (displayed schematically in dashed guide line). On the other hand, increasing the drain

bias V_{DS} resulted in [1] a gradual smearing of peaks and valleys. At V_{DS} = 5 mV no resonant structures were observed. This first conclusion is an indication about the quantum level separation. Let us recall that we can discriminate between two tunnelling paths via the various quantum levels provided that the injection energy window is less than energy offsets between two successive levels. On this criteria, this means that at 5mV two quantum levels can be simultaneously involved with the subsequent smear out of resonant tunnelling feature. With respect to operating temperature issue, this criterion also determines the temperature robustness of phenomenon by comparing the energy of 5meV to $k_B\theta$.

In addition, the device with constant bias condition corresponding to the onset of the second resonant peak (V_{GS} = 0.12 V and V_{DS} = 0.2 mV) was exposed to the light emission of a LED. The authors have found almost quite similar resonant tunnelling structures which are displayed in Figure 4.3 with the plot of drain-source current controlled by the bias LED current.

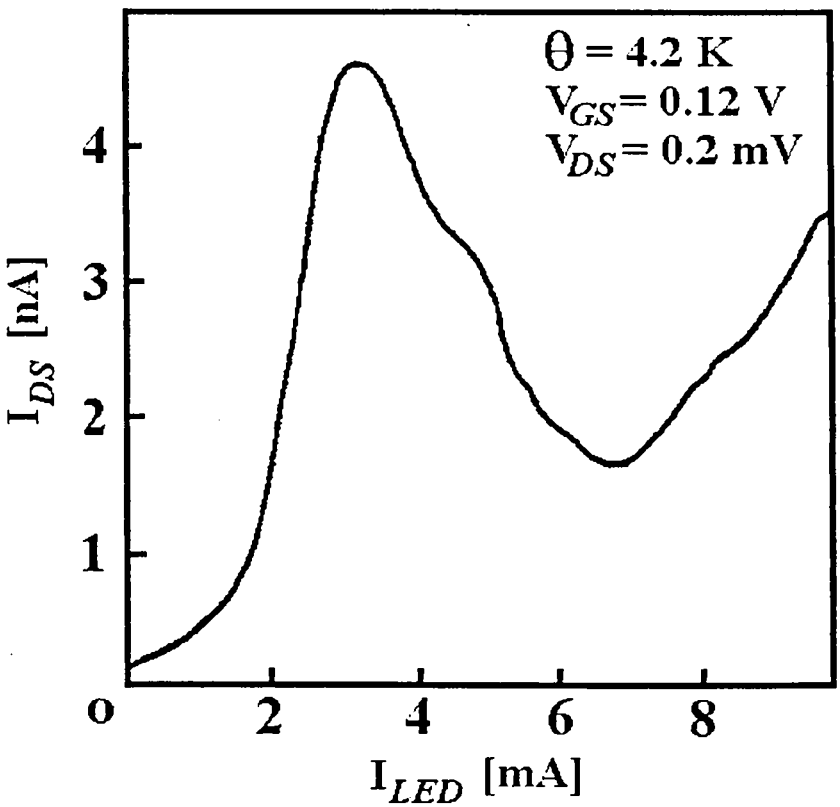


Figure 4.3 drain-source current as a function of LED current after [1].

After this outline of key experimental results let us introduce a brief review on the major modelling techniques applied to Field-Effect Transistors.(FET's) by pointing out the main assumptions.

First, we discuss the analytical modelling of MODFET's and the resulting charge control law, maximum 2DEG concentration and threshold voltage, ... On this bases we investigate the validity of those analytical approaches for nano-structure. Then we focus our attention to numerical techniques, namely Mont-Carlo, hydrodynamic, quantum mechanical and hybrid methods. Finally, we describe the major physical and numerical features of our model.

4.4 Review of FET's modelling techniques :

Both, analytical and numerical models can be devided into two classes namely classical or quantum mechanical approaches. On the other hand, subdivision can be established with respect to the fact that the transport properties are treated at the microscopic or macroscopic level. At last, there is some distinction which stems from the non stationary effects notably local or energy models. In this context, numerous theoretical approaches can be found in the literature which are briefly discussed below.

4.4.1 Brief review of analytical modelling techniques :

Analytical models are closed form expressions for the transistor DC and small signal parameters. A physical based analytical model is very useful to get a first insight into the device behaviour. Unfortunately, it often requires to make quite crude assumptions which limits their validity for quantum-sized structures. Starting from the conventional field effect control principle (details in section 4.4.1.1) we will discuss of its application to heterostructure FET's.

4.4.1.1 Analytical models for conventional FET's :

The first analytical model was proposed by Schockley [9] by introducing the gradual channel approximation and by assuming complete depletion. Furthermore, it was assumed that the depletion width is determined uni-dimensionally. It is clear that this assumption is no longer valid for very short gates where fringing field, not to mention non-stationary electron dynamics and non-isothermal transport, cannot be neglected.

Later on, the origin of the drain current saturation was clearly established with the works of Statz et al. [10] and Greben and Ghandi [11] by involving the carrier velocity saturation at the drain side of the gate.

Shur [12] introduced the Negative Differential Mobility (NDM) effect by assuming the formation of a high field dipole domain at the gate output region.

From the above introduction on the conventional FET analytical modelling, one can summarises the most common assumptions as follows :

- the transport parameters are instantaneously dependent of the local electric field (local models).
- the diffusion current is neglected (drift models).
- the substrate current is neglected.
- the gradual channel approximation is most frequently assumed.
- the zone under the gate are assumed to be completely depleted.
- the velocity-field (v-F) relation is often approximated by piece-linear model.

It can be concluded that the analytical models techniques are not able to describe the sub-micrometer gate and a-fortiori the nano-gate device behaviour where the non-stationary electron dynamics and the non-isothermal transport dominates the device performance.

As regards the heterostructure FET's, the analytical models are besides, usually based on the triangular-well approximation. In that case, the basic principle can be understood but often these models failed to quantitatively predict the relevant features. In addition, back injection, unintentional doping, screening effect are always neglected.

4.4.1.2 Analytical models for heterostructure FET's :

a. Charge control equation :

Figure 4.4 shows typical heterostructure FET (HFET) with the conduction band variation under the gate. Solving Poisson's equation in both sides of the heterointerface taking into account the continuity of the normal displacement vector, we get the following charge-control law [13],

$$n_s = \frac{\epsilon}{q\,d}\left(V_p - |\Phi| + V_{gs} + |\Delta E_c| - |E_{fn}|\right) \qquad \textit{Eq. 4-1}$$

where,

$V_p = \frac{qN_d}{2\epsilon}(d-e)^2$ is the pinch-off potential. *Eq. 4-2*

V_{gs} is the applied voltage.

Φ is the built-in potential.

ΔE_c is the conduction bend offset.

Other parameters are defined in Figure 4.4.

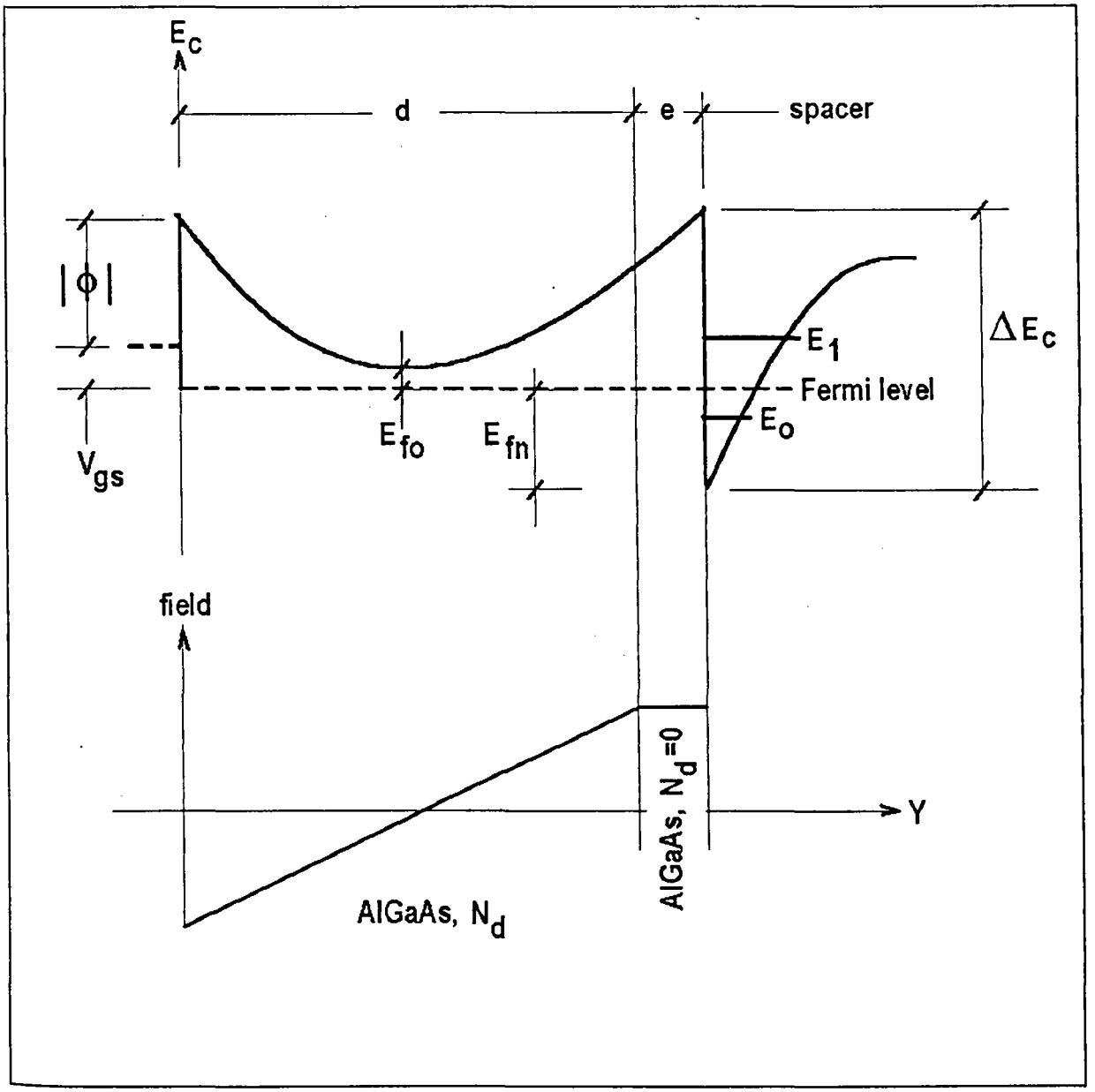


Figure 4.4 schematic diagram of MODFET conduction band profile along with the electric field.

b. Maximum 2DEG :

The maximum 2DEG concentration is obtained for the gate-less heterostructure when the gate is too far from the heterointerface to affect the 2DEG.

This is given by [14],

$$\tilde{n}_s = \sqrt{\frac{2N_d \in}{q}\left(|\Delta E_c| - |E_{fo}| - |E_{fn}|\right)^2 + N_d^2 e^2} - N_d e \qquad \textit{Eq. 4-3}$$

Where,

N_d is the doping concentration of the highly-doped material.

E_{fo} is the equilibrium Fermi level in the highly-doped material.

e is the spacer thickness.

It clear that e influences drastically the carrier density with subsequent impact on transport properties.

4.4.1.3 Classical versus quantum mechanical models :

a. Classical models :

In classical models, the Fermi-level is related to the 2DEG concentration through either, Fermi-Dirac :

$$n_s = \int_0^W N_c F_{1/2}\left[\frac{E_{fn} - E_c(y)}{k_B \theta_o}\right] dy \qquad \textit{Eq. 4-4}$$

or Boltzmann statistics :

$$n_s = \int_0^W N_c \exp\left[\frac{E_{fn} - E_c(y)}{k_B \theta_o}\right] dy \qquad \textit{Eq. 4-5}$$

where N_c is the 3D effective density of states and W is the effective quantum well width and the integration is over space in the growth direction (y).

b. Quantum mechanical models :

In these models, the Fermi-Dirac statistics are weighted by the existence probability resulting from the solution of the time-independent Schroedinger's equation, this gives,

$$n_s = \sum_{i=1}^{\infty} \int_0^W \frac{m^* k_B \theta_o}{\pi h^2} \ln\left[1 + \exp\left(\frac{E_{fn} - E_i}{k_B \theta_o}\right)\right] |\Psi(y)|^2 dy \qquad \textit{Eq. 4-6}$$

Where for the triangular well approximation, E_i and $\Psi_i(y)$ are given by,

$$E_i = \left(\frac{h^2}{2m^*}\right)^{1/3} \left[3\, q F_s\, \pi \left(i + 3/4\right)/2\right]^{2/3} ;\ i=0,1,2... \qquad \textit{Eq. 4-7}$$

$$\Psi_i(y) = A_i\left[\frac{2m^* q F_s}{h^2}\left(y - \frac{E_i}{q} F_s\right)\right] \qquad \textit{Eq. 4-8}$$

where A_i is the Airy function and F_s is the electric field at the heterointerface.

In both classical and quantum mechanical models, numerical techniques are required to get a self-consistent solution for the 2DEG concentration. This numerical approach shows that the $(n_s - E_{fn})$ relation can be expressed analytically by the following linear relation,

$$E_{fn} = a\, n_s + \delta E_f, \qquad 15x10^{11} \geq n_s \geq 5x10^{11}\ cm^{-2} \qquad \textit{Eq. 4-9}$$

Where δE_f denotes a constant value.

Substituting into Eq. 4-3 gives the maximum 2DEG,

$$\tilde{n}_s = \sqrt{\frac{2N_d \in}{q}\left(|\Delta E_c| - |E_{fo}| - |\delta E_f|\right)^2 + N_d^2 e1^2} - N_d e_1 \qquad \textit{Eq. 4-10}$$

where

$$e_1 = e + \frac{\in a}{q} \qquad \textit{Eq. 4-11}$$

By comparing Eq. 4-3 and Eq. 4-10, it can be seen that the physical meaning of (e_1) is an effective spacer thickness. In fact the factor ($\in a / q$) accounts for the shift of the effective centre of 2DEG charge from the heterointerface. This effective centre of 2DEG charge depends on the original used model i.e classical or quantum mechanical model.

4.4.1.4 Charge control threshold :

This is the minimum voltage value for a gate effect onset. In other words, for gate voltages higher than threshold voltage, the 2DEG is no longer controlled by the gate potential. This can be obtained by equating Eq. 4-1 and Eq. 4-3.to give,

$$V_{gth} = |\Phi| - |E_{fo}| - \left(\sqrt{\frac{qN_d d^2}{2\in}} - \sqrt{BB + \frac{qN_d e^2}{2\in}}\right)^2 \qquad \textit{Eq. 4-12}$$

where (BB) denotes the band bending and is defined as,

$$BB = |\Delta E_c| - |E_{fo}| - |E_{fn}| \qquad \textit{Eq. 4-13}$$

The above equations, address the electrostatic issue, assuming either semi-classical Boltzmann/Fermi-Dirac statistics or quantum-mechanical occupancy. In other words, the charge control and the relevant threshold voltage are now well understood including the effect of doping, spacer thickness and gate voltage.

In contrast, the lateral motion of electrons is treated under very simplified assumptions and to overcome such drawbacks, it is often interesting to make use of numerical modelling techniques. On the other hand, more accurate predictions about sheet carrier density can be derived by self consistently solving the Schroedinger's and Poisson's equations with realistic potential profile.

4.4.2 Numerical modelling techniques :

Numerical models are usually classified into three main categories, notably, the microscopic quantum approach, the Mont-Carlo technique (MC) and the hydrodynamic models. The microscopic quantum approach is an ab-initio modelling technique which consider the electron dynamics on the atomic scale. It is evident that such basic approach requires computing resources which are too important for parametric studies of a real device.

One can conclude that the microscopic quantum approach simulates the microscopic motion of electrons associated with their wave function propagation, under ballistic conditions along with the microscopic perturbation effects of different scattering mechanisms. In this quantum approach the most exact potential function created by the different atoms of the lattice has to be taken into account. This is why this technique is essentially devoted onto the study of material properties rather than device structures.

In this context, a so-called Mont-Carlo (MC) method is often performed by keeping the microscopic character of the simulation. In brief, the band structure of the semiconductor of concern is calculated using notably tight bending or pseudo-potential techniques. The output data are the dispersion relationships in the Brillouin zone. Subsequently, one can model the motion of electrons under both the deterministic effect of the applied electric field and the stochastic effect of different scattering mechanisms. These methods are expensive from the stand point of computational time and need very powerful computers. This is particularly true for describing transport phenomena which involve electron motion in the whole Brillouin zone. The most representative situation is the impact ionisation process which requires to take into

account all the high energy sites for permitting the electron getting the ionisation threshold energy.

Otherwise, different degrees of approximations have been proposed for alleviating the physical complexity along with the computational effort. In the previous discussion we spoke about a homogeneous material. Certainly, the description of a heterojunction while concerning the microscopic treatment is further a challenging issue. So far, the heterojunction is often treated by neglecting the different material mutual-effects.

On the other hand, the effective mass approximation is often assumed along with in some cases a non-parabolicity situation by some fitting parameter. Nevertheless, MC method can be considered as the most exact model able to intrinsically describes the hot electron effect such as velocity overshoot, quasi-ballistic motion .... On the other hand, by using this particle method, it is possible to predict all the key parameters of an electron gas subject to any field force. Therefore, the variation of mobility, the diffusion coefficient and both the energy and momentum relaxation times can be calculated as a function of electron energy.

Such a knowledge, is determinant in the scope of hydrodynamic energy models which are briefly discussed here after. The introduction of hydrodynamic treatment in the sense that the electron gas is considered as a continuum of electron streams is motivated by decreasing the computation effort. On the other hand, the numerical random fluctuations of the results due to the probabilistic nature of microscopic studies are avoided.

Reiser [15], K. Yamaguchi [16], and J. Barnes [17] developed numerical local hydrodynamic models for FET's where the electron motion is governed by hydrodynamic conservation equations in which the mobility and diffusion coefficients are assumed to be instantaneous functions of local electric field.

Blotekjaer [18] assumed the relaxation times, mobility and effective mass to be instantaneous functions of the average electronic kinetic energy. He derived a set of transport equations including energy conservation for two valley semiconductors based on Boltzmann transport equation (BTE).

Shur was one of the first who proposed the use of MC steady state simulations to get the energy-dependent transport parameters required to solve the hydrodynamic conservation equations. This was believed to give the same degree of precision as MC simulations but without the inherent numerical noise of MC techniques.

These numerical procedures were applied either in bulk situation or for uni-dimensional device structures. In order to treat the two dimensional character, several authors have introduced various level of approximations. Notably, A. Cappy [19] with quasi-two-dimensional approach and Cook and J. Fery. The later [20] suggested to solve the continuity equation in two dimensional whereas the energy equation was considered one dimension.

In the present work, we are going to base our analysis on the model originally developed by Ibrahim. This model published in reference [21] is a full 2D hydrodynamic energy approach with self-consistent treatment of continuity, energy conservation and Poisson's equation. As stated before the non-stationary effect comes from the energy dependent parameters derived from steady state MC simulations.

S. El-Azhary [22] and K. Ismail [23] exhaustively used the same model in MESFET analysis and simulations. T. Shawki [24][25][26] extended the model to HFET's modelling and simulations. K. Sherif [27] used the same model for MODFET simulation at cryogenic temperatures. We have extended the code to study the device performance of dual-gate MODFET's [28][29].

4.5 Model description :

We shall discuss the frame work of the model adopted here and then we introduce the necessary modifications for including the quantum tunnelling effects.

Particle, momentum and energy conservation equations are obtained by taking the first three moments in k-space of Boltzmann Transport Equation (BTE) :

$$\frac{\partial f}{\partial t} + \vec{v}.\nabla_r(f) + \frac{q\vec{F}}{h}.\nabla_K(f) = \left(\frac{\partial f}{\partial t}\right)_c \qquad \textit{Eq. 4-14}$$

This procedure leads to the following conservation equations for each valley of a multi-valley semiconductor denoted with the index i :

- conservation of particles, n_i

$$\frac{\partial n_i}{\partial t} = \left(\frac{\partial n_i}{\partial t}\right)_c - \nabla_r.(n_i\,\vec{v}_i) \qquad \textit{Eq. 4-15}$$

- conservation of momentum, $\vec{P}_i = (P_{xi}, P_{yi}, P_{zi})$

$$\frac{\partial P_{xi}}{\partial t} = \left(\frac{\partial P_{xi}}{\partial t}\right)_c - \nabla_r.(P_{xi}.\vec{v}_i) - \frac{\partial (n_i k_B \theta_i)}{\partial x} + q\,n_i F_x \qquad \textit{Eq. 4-16}$$

- conservation of total kinetic energy, W_i

$$\frac{\partial W_i}{\partial t} = \left(\frac{\partial W_i}{\partial t}\right)_c - \nabla_r.(W_i.\vec{v}_i) - \nabla_r.(n_i k_B \theta_i \vec{v}_i + \vec{Q}_i) + q\,n_i \vec{v}_i.\vec{F} \qquad \textit{Eq. 4-17}$$

The term $\nabla_r.(\vec{Q}_i)$ accounts for the outward heat flow rate by thermal conduction. For Maxwellian or any other symmetric distribution function the heat flow $\vec{Q}_i$ being an odd moment of order three vanishes.

The terms, $\nabla_r.(n_i\,\vec{v}_i)$, $\nabla_r.(P_{xi}.\vec{v}_i)$ and $\nabla_r.(W_i.\vec{v}_i)$ represent the rate of decrease of n_i, P_{xi} and W_i due to the outward flux i.e by convection.

The terms, $q\,n_i F_x$ and $-\frac{\partial\,(n_i \mathrm{k}_B \theta_i)}{\partial \mathrm{x}}$ represents the rate of increase of momentum due to the electrostatic force and the electronic pressure gradient respectively.

The term $q\,n_i \vec{\mathrm{v}}_i . \vec{F}$ represent the power gained by electrons from electric field, and $\nabla_r . (n_i k_B\, \theta_i \vec{\mathrm{v}}_i)$ represents the power lost by electrons in the form of electronic gas expansion (decrease in electronic gas pressure).

Finally the collision terms, $\left(\frac{\partial n_i}{\partial t}\right)_c$, $\left(\frac{\partial P_{xi}}{\partial t}\right)_c$ and $\left(\frac{\partial W_i}{\partial t}\right)_c$ represent the rate of change of n_i, P_{xi} and W_i due to different scattering mechanisms. Assuming full isotropy and the scattering rate to be dependent only on the average kinetic energy per electron

$$w_i = \frac{W_i}{n_i} = \frac{3}{2} k_B \theta_i + \frac{1}{2} m_i^* \vec{v}_i^2 \qquad \textit{Eq. 4-18}$$

we get [18],

$$\left(\frac{\partial n_i}{\partial t}\right)_c = \frac{-n_i}{\tau_{nij}(w_i)} + \frac{n_j}{\tau_{nji}(w_j)} \qquad \textit{Eq. 4-19}$$

$$\left(\frac{\partial \vec{P}_i}{\partial t}\right)_c = \frac{-n_i m_i^* \vec{v}_i}{\tau_{nij}(w_i)} + o + \frac{-n_i m_i^* \vec{v}_i}{\tau_{pi}(w_i)} \qquad \textit{Eq. 4-20}$$

$$\left(\frac{\partial W_i}{\partial t}\right)_c = \frac{-n_i w_i}{\tau_{nij}(w_i)} + \frac{n_j w_j}{\tau_{nji(w_j)}} + \frac{-n_i (w_i - w_o)}{\tau_{wi}(w_i)} \qquad \textit{Eq. 4-21}$$

where $w_o = \frac{3}{2} k_B \theta_o$ is the lattice energy. It is worth mentioning that despite the fact that there is some exchange of particles between the valleys i and j, the net effect for the incoming momentum point of view is zero. The later statement is based on the assumption that the effect of scattering on the incoming momentum is completely randomised. This explains the zero term in Eq. 4-20.

The above conservation equations are too complicated to be of any practical use. The difficulty mainly arises from the fact that one has to solve those equations in each valley separately then coupling the results by the collision terms. The solution is to find an equivalent single valley electron gas model.

4.5.1 The equivalent single valley electron gas model :

Such an equivalent single valley treatment can be achieved by taking the weighted average of the properties of the constituting electron gases.

This is done for extrinsic semiconductors with shallow donors and no traps while considering that the momentum relaxation time is one order of magnitude less than that of the energy. Also assuming that the statistical average of product is equal to the product of statistical averages, the following equivalent single electron gas conservation equations can be written [21][26] :

$$\frac{\partial n}{\partial t} + \nabla_r . (n \vec{v}) = 0 \qquad \textit{Eq. 4-22}$$

$$n \vec{v} = n \mu \vec{F} - \mu \nabla_r \left(n \frac{k_B \theta}{q} \right) \qquad \textit{Eq. 4-23}$$

$$\frac{\partial \varepsilon}{\partial t} + \vec{v} . \nabla_r (\varepsilon) = q \vec{v} . \vec{F} - \frac{1}{n} \nabla_r . (n k_B \theta \vec{v}) - \frac{(\varepsilon - w_o)}{\tau_\varepsilon (\varepsilon)} \qquad \textit{Eq. 4-24}$$

where now the energy conservation equation is written for the total (drift +thermal) average electron energy.

These conservation equations in addition to Poisson's equation represents the basic hydrodynamic energy model characterising electron transport in non-degenerate, multi-valley semiconductors and subjected to the above mentioned assumptions.

4.5.2 Simulation of different types of boundaries :

In this section, we discuss the different types of boundaries and the relevant boundary conditions. First, we shall describe in short the conventional procedures to simulate the various boundaries which are used by most of the authors. Second we shall investigate the validity of them and on this basis we introduce some modifications.

4.5.2.1 Simulation of the gate rectifying junction :

The gate rectifying junction is modelled by introducing a value potential ($V_g = V_{g\ externally\ applied} - \Phi$), in Poisson's equation ($\Phi = 0.8$ Volt). As regards the continuity equation, one of the following boundary conditions has to be used :

- setting the normal component of the conduction current intentionally to zero. This is only valid for the usually used biasing conditions where the gate conduction current should be zero. Under this condition the charge on the gate (n_g) is determined from the solution of the following particle equation ;

$$J_{cy} = \mu\left[n_g F_y - \frac{\partial}{\partial y}\left(n_g k_B \theta / q\right)\right] = 0 \qquad \textit{Eq. 4-25}$$

which yields :

$$n_g = n_{g-1} \exp\left(-\frac{qF_y}{k_B\theta} . \Delta y \right) \qquad \textit{Eq. 4-26}$$

It is found that this boundary condition yields unphysically a low gate conduction current. This current is due to the numerical noise which is amplified by numerical differentiation and exponentiation.

- the second boundary condition is based on the knowledge relevant to the Fermi-level. Given that the gate zone is practically non-degenerate the Boltzmann approximation holds and hence :

$$n_g = N_c \exp\left(qE_f / k_B\theta\right) \qquad \textit{Eq. 4-27}$$

Using the above boundary condition, it can be shown that n_g is much smaller than the doping level. In counterpart, very high free carrier gradient might be a source of numerical instabilities if this technique is applied on a coarse mesh size. However, the usage of the above equation results in a more stable convergence and much smaller numerical difference between the source and drain currents.

At cryogenic temperatures (4.2 K), both of the two boundary conditions result in a gate charge n_g practically equal to zero and both could be used without any preference. This implies a very small time step to preserve the stability of the program which increases the computational effort needed to achieve the convergence.

4.5.3 Source and drain shallow Ohmic contacts :

The source and drain shallow contacts are introduced to Poisson's equation solver as equi-potential surfaces with a fixed potential, zero for source contact and V_{ds} for the drain one. As there is neither accumulation nor depletion ($n = N_d$) across the biased Ohmic contacts then, a fixed level of mobile-charge carriers (n) is introduced. This level of charge on the contacts is dependent on the donor ionisation probability which is highly dependent on the temperature.

In addition, the energy of electrons at the two contacts are set to the lattice energy. Also, it is important to notice that usually there exists enough separation between the nearest gate and the drain, so that the electrons are thermallised when they reach the drain contact.

Usually the above boundary conditions are systematically applied in case of deep Ohmic contacts which shall be discussed in the following.

4.5.4 Electrode-free surfaces :

Since the boundaries of the region under study are separated enough from the active region, Neumann boundary conditions are usually applied :

$$\text{no diffusion} \Rightarrow \frac{\partial n}{\partial x} = \frac{\partial n}{\partial y} = 0 \qquad \textit{Eq. 4-28}$$

$$\text{no drift} \Rightarrow \frac{\partial V}{\partial x} = \frac{\partial V}{\partial y} = 0 \qquad \textit{Eq. 4-29}$$

$$\text{no energy dissipation} \Rightarrow \frac{\partial \varepsilon}{\partial x} = \frac{\partial \varepsilon}{\partial y} = 0 \qquad \textit{Eq. 4-30}$$

where $\frac{\partial}{\partial y}$ for lateral surfaces and $\frac{\partial}{\partial x}$ for the vertical ones.

4.5.5 Simulation of heterojunction :

It is well known that the energy gain or loss across the interface of a heterojunction can be described by means of a quasi-field. The magnitude of this field depends strongly of the expected finite distance over which the transition from one material to another is achieved. It is clear that for high conduction band offsets and rather perfect interface over a distance of one atomic layer (≈ 5 Å), the magnitude of this field becomes so high that this procedure is questionable. Nevertheless, over so short distance the net effect is the change between potential and kinetic energies across the heterointerface. In practice, this equivalent field is added to the electrostatic field, solution of Poisson's equation.

4.6 Solution of Poisson's equation :

Poisson's equation is directly solved using the Matrix Double Sweep (MDS) method [21]. The method is applied under a variable mesh size condition. The MDS method is the LU factorisation technique applied to five diagonal systems [30]. Although it is a direct and accurate method, it prohibits the use of adaptive mesh refinement techniques due to the unacceptable code slow down so that a suitable mesh should be chosen from the first beginning of time iterations. The same method is used in the previous chapters but in one-dimensional form.

4.7 Solution of the energy equation :

The energy equation is usually solved as follows :

Taking w_o as the energy reference, the energy in Eq. 4-24 yields,

$$\frac{\partial \varepsilon}{\partial t} + \vec{v}.\nabla_r(\varepsilon) = q\,\vec{v}.\vec{F} - \frac{1}{n}\nabla_r.(n\,k_B\theta\,\vec{v}) - \frac{\varepsilon}{\tau_\varepsilon(\varepsilon)} \qquad \textit{Eq. 4-31}$$

In finite differences Eq. 4-31 reads,

$$\frac{\partial \varepsilon_{ij}}{\partial t} = b^o_{i,j}\,\varepsilon_{ij} + b^1_{i,j}\,\varepsilon_{i,j+1} + b^2_{i,j}\,\varepsilon_{i,j-1} + b^3_{i,j}\,\varepsilon_{i+1,j} + b^4_{i,j}\,\varepsilon_{i-1j} + h_{i,j} \qquad \textit{Eq. 4-32}$$

Eq. 4-32 can be written for all mesh points (ij) in one single vectorial equation,

$$\frac{\partial \vec{\varepsilon}}{\partial t} = B\,\vec{\varepsilon} + \vec{h} \qquad \textit{Eq. 4-33}$$

Where B is a five diagonal matrix that does not depend explicitly on the energy vector ($\vec{\varepsilon}$). The vector ($\vec{h}$) which is the source term in the above vector equation, represents the energy generation/dissipation terms,

$$q\,\vec{v}.\vec{F}, \quad \frac{1}{n}\nabla_r.(n\,k_B\theta\,\vec{v}) \qquad \textit{Eq. 4-34}$$

In the energy conservation equation (Eq. 4-32), the general expression for the elements of B are given by,

$$b^{o}_{i.j} = \frac{-1}{\tau_{\varepsilon}(\varepsilon_{i.j})} + \frac{1}{\Delta x_j}\left(s_1 v_{x\,i.j+1/2} - s_2 v_{x\,i.j-1/2}\right) + \frac{1}{\Delta y_j}\left(s_3 v_{y\,i+1/2.j} - s_4 v_{y\,i-1/2.j}\right) \qquad \textit{Eq. 4-35}$$

$$b^{1}_{i.j} = \frac{-1}{\Delta x_j} s_1 \, v_{x\,i.j+1/2} \qquad \textit{Eq. 4-36}$$

$$b^{2}_{i.j} = \frac{+1}{\Delta x_j} s_1 \, v_{x\,i.j-1/2} \qquad \textit{Eq. 4-37}$$

$$b^{3}_{i,j} = \frac{-1}{\Delta y_i} s_3 \, v_{y\,i+1/2,j} \qquad \textit{Eq. 4-38}$$

$$b^{4}_{i,j} = \frac{+1}{\Delta y_i} s_3 \, v_{y\,i-1/2,j} \qquad \textit{Eq. 4-39}$$

The (s_k, $k = 1, 2, 3$ & 4) are introduced to the above formalism through the term $\vec{v}.\nabla_r(\varepsilon)$ in Eq. 4-31, which is originally a part of the energy convection term $\nabla_r.(W.\vec{v}) = \nabla_r.(n\,\varepsilon\,.\vec{v}\,)$ in the non-simplified form of the energy equation (Eq. 4-17). If one consider the temporal energy variation due to the energy convection only one can write,

$$\left.\frac{-\partial \vec{\varepsilon}}{\partial t}\right|_{\text{convection}} = \nabla_r\,.\,(\varepsilon.\vec{v}\,) = \varepsilon\,.\,\nabla_r\,.\,(\vec{v}\,) + \vec{v}\,.\,\nabla_r\,(\varepsilon) \qquad \textit{Eq. 4-40}$$

A general issue concerning the discretisation scheme stems from the fact that the relevant convection terms depends of scalar as well as vectorial quantities. In practice, all vectorial variables are defined at midway points between scalar-quantities mesh-points. This means that a careful strategy must be followed when one needs to evaluate $\vec{v}.\nabla_r(\varepsilon)$ at the (i,j) mesh point. The upstream method is chosen for its simplicity and above all for the stability it creates with time evolution.

Based on the upstream method the energy at the mesh point (i,j+1/2), ($\varepsilon_{ij+1/2}$) is defined by,

$$\varepsilon_{i,j+1/2} = \varepsilon_{i,j}\frac{1+\mathrm{sgn}(v_{x\,i,j+1/2})}{2} + \varepsilon_{i,j+1}\frac{1-\mathrm{sgn}(v_{x\,i,j+1/2})}{2} \qquad \textit{Eq. 4-41}$$

$$\varepsilon_{i+/2,j} = \varepsilon_{i,j}\frac{1+\mathrm{sgn}(v_{y\,i+1/2,j})}{2} + \varepsilon_{i+1,j}\frac{1-\mathrm{sgn}(v_{y\,i+1/2,j})}{2} \qquad \textit{Eq. 4-42}$$

where,

$$\mathrm{sgn}(v_x) = \begin{cases} +1 \Leftrightarrow v_x \geq 0 \\ -1 \Leftrightarrow v_x \leq 0 \end{cases} \qquad \textit{Eq. 4-43}$$

At last,

$$s_1 = \frac{1-\mathrm{sgn}(v_{x\,i,j+1/2})}{2} \qquad \textit{Eq. 4-44}$$

$$s_2 = \frac{1+\mathrm{sgn}(v_{x\,i,j-1/2})}{2} \qquad \textit{Eq. 4-45}$$

$$s_3 = \frac{1-\mathrm{sgn}(v_{y\,i+1/2,j})}{2} \qquad \textit{Eq. 4-46}$$

$$s_4 = \frac{1+\mathrm{sgn}(v_{y\,i-1/2,j})}{2} \qquad \textit{Eq. 4-47}$$

4.7.1 The source term of the energy equation :

The energy equation between two successive iterations k and k+1 can be written as,

$$\frac{\Delta \varepsilon_{i,j}^{k+1}}{\Delta t} = \frac{\Delta \varepsilon_{i,j}^{k}}{\Delta t} + \Delta\left(\frac{\Delta \varepsilon_{i,j}}{\Delta t}\right)^{k \to k+1} \qquad \textit{Eq. 4-48}$$

where in the above equation,

$$\frac{\Delta \varepsilon_{i,j}^{k}}{\Delta t} = A_{i,j}^{o} \qquad \textit{Eq. 4-49}$$

$$\Delta\left(\frac{\Delta \varepsilon_{i,j}}{\Delta t}\right)^{k \to k+1} = b_{ij}^{o} \Delta \varepsilon_{i,j}^{k+1} + \sum_{c=1}^{4} b_{i,j}^{c} \, \Delta \vec{\varepsilon}^{k} + \frac{\partial h_{i,j}}{\partial \varepsilon_{i,j}} \frac{\partial \varepsilon_{i,j}}{\partial t} \Delta t \qquad \textit{Eq. 4-50}$$

and,

$$\frac{\partial h_{i,j}}{\partial \varepsilon_{i,j}} \frac{\partial \varepsilon_{i,j}}{\partial t} \Delta t = \frac{\partial h_{i,j}}{\partial \varepsilon_{i,j}} \Delta \varepsilon_{i,j}^{k+1} \qquad \textit{Eq. 4-51}$$

Simple algebraic manipulations leads to,

$$\Delta \varepsilon_{i,j}^{k+1} \left(\frac{1}{\Delta t} - \frac{\partial h_{i,j}}{\partial t} - b_{i,j}^{o}\right) = A_{i,j}^{o} + \sum_{c=1}^{4} b_{i,j}^{c} \, \Delta \vec{\varepsilon}^{k} \qquad \textit{Eq. 4-52}$$

Where, ($b_{i,j}^{o}$, $b_{i,j}^{1}$, $b_{i,j}^{2}$, $b_{i,j}^{3}$, $b_{i,j}^{4}$ & $\frac{\partial h_{i,j}}{\partial \varepsilon_{i,j}}$) are semi-updated and $A_{i,j}^{o}$ is calculated directly using the analytical form of the energy equation with semi-updated values.

The physical effect of the term $\frac{\partial h_{i,j}}{\partial \varepsilon_{i,j}}$ in Eq. 4-52 is that it accelerates/decelerates the temporal variation of energy. This is simply as it sums over the term $\frac{1}{\Delta t}$ in the same equation.

Now we are addressing the evaluation of the term $\frac{\partial h_{i,j}}{\partial \varepsilon_{i,j}}$.

Given that, $$h_{i,j} = q\,\vec{v}.\vec{F} - \frac{1}{n}\nabla_r.(n\,k_B\theta\,\vec{v}) \qquad \textit{Eq. 4-53}$$

By simple algebra Eq. 4-53 yields,

$$h_{i,j} = \frac{v^2}{\mu} - V_{th}\,\nabla_r.\vec{v}\,, \qquad V_{th} = k_B\,\theta / q \qquad \textit{Eq. 4-54}$$

hence,

$$\frac{\partial h_{i,j}}{\partial \varepsilon} = \left(2\frac{\vec{v}}{\mu} - V_{th}\nabla_r\right).\frac{\partial \vec{v}}{\partial \varepsilon} - \frac{v^2}{\mu^2}\frac{\partial \mu}{\partial \varepsilon} - \frac{\partial V_{th}}{\partial \varepsilon}\nabla_r.\vec{v} \qquad \textit{Eq. 4-55}$$

A first order approximation is carried out by neglecting, $\frac{v^2}{\mu^2}\frac{\partial \mu}{\partial \varepsilon}$ and the drift component in $\frac{\partial \vec{v}}{\partial \varepsilon}$ thus,

$$\frac{\partial h_{ij}}{\partial \varepsilon} = \left(2\frac{\vec{v}}{\mu} - V_{th}\nabla_r\right).\frac{\partial \vec{v}_{dif}}{\partial \varepsilon} - \frac{\partial V_{th}}{\partial \varepsilon}\nabla_r.\vec{v} \qquad \textit{Eq. 4-56}$$

which is the formulation usually used in the simulation of Field-Effect devices, by taking into account the relaxation effects into account. Finally Eq. 4-52 is solved using successive over relaxation method. In most cases, the numerical stability is even achieved using a relaxation parameter equal one.

4.8 Solution of the continuity equation :

Let us recall the simplified form of the continuity equation (Eq. 4-22),

$$\frac{\partial n}{\partial t} + \nabla_r . (n \bar{v}) = 0 \qquad \textit{Eq. 4-57}$$

Using finite difference, the above equation reads,

$$\frac{\partial n_{ij}}{\partial t} = a^o_{i,j}\, n_{ij} + a^1_{i,j}\, n_{i,j+1} + a^2_{i,j}\, n_{i,j-1} + a^3_{i,j}\, n_{i+1,j} + a^4_{ij}\, n_{i-1,j} \qquad \textit{Eq. 4-58}$$

where,

$$a^o_{i,j} = \frac{1}{\Delta x_j}\left(s_1 v_{xi,j+1/2} - s_2 v_{xi,j-1/2}\right) + \frac{1}{\Delta y_j}\left(s_3 v_{yi+1/2,j} - s_4 v_{yi-1/2,j}\right) \qquad \textit{Eq. 4-59}$$

$$a^1_{i,j} = \frac{-1}{\Delta x_j} s_1 v_{x\,i,j+1/2} \qquad \textit{Eq. 4-60}$$

$$a^2_{i,j} = \frac{+1}{\Delta x_j} s_2 v_{x\,i,j-1/2} \qquad \textit{Eq. 4-61}$$

$$a^3_{i,j} = \frac{-1}{\Delta y_i} s_3 v_{y\,i+1/2,j} \qquad \textit{Eq. 4-62}$$

$$a^4_{i,j} = \frac{+1}{\Delta y_i} s_3 v_{y\,i-1/2,j} \qquad \textit{Eq. 4-63}$$

The upstream method is also used and hence,

$$n_{i,j+1/2} = n_{i,j}\frac{1+\mathrm{sgn}(v_{x\,i,j+1/2})}{4} + n_{i,j+1}\frac{1-\mathrm{sgn}(v_{x\,i,j+1/2})}{4} \qquad \textit{Eq. 4-64}$$

$$n_{i+/2,j} = n_{i,j}\frac{1+\mathrm{sgn}(v_{y\,i+1/2,j})}{2} + n_{i+1,j}\frac{1-\mathrm{sgn}(v_{y\,i+1/2,j})}{2} \qquad \textit{Eq. 4-65}$$

At last,

$$s_1 = \frac{1 + \mathrm{sgn}(v_{x\,i,j+1/2})}{2} \qquad \textit{Eq. 4-66}$$

$$s_2 = \frac{1 - \mathrm{sgn}(v_{x\,i,j-1/2})}{2} \qquad \textit{Eq. 4-67}$$

$$s_3 = \frac{1 + \mathrm{sgn}(v_{y\,i+1/2,j})}{2} \qquad \textit{Eq. 4-68}$$

$$s_4 = \frac{1 - \mathrm{sgn}(v_{y\,i-1/2,j})}{2} \qquad \textit{Eq. 4-69}$$

Finally Eq. 4-58 is also solved using successive over relaxation. In most cases, the code converges with a relaxation parameter equal to 1.2.

4.9 Hydrodynamic model, results and discussions :

Before addressing the core issues of quantum devices, it seems interesting to illustrate the capability of the numerical code described above. Primarily, this is done to give a physical insight into the 2-D character.

For this purpose, we simulated a MODFET device having the structural parameters depicted in Figure 4.5. With respect to the test vehicle discussed previously, here the highly doped wide gap thickness is shrunk to 20 nm for a gate inter-distance of 40 nm. Also, the simulations are conducted at room temperature to alleviate, in a first approach, the problems in connection with very low temperature simulations.

The steady state results of this simulations carried out at zero gate and drain bias are displayed in Figure 4.6 a and b. In practice, we started from the non equilibrium charge distribution $n(x,y) = N_d(x,y)$ at zero time step with zero energy anywhere.

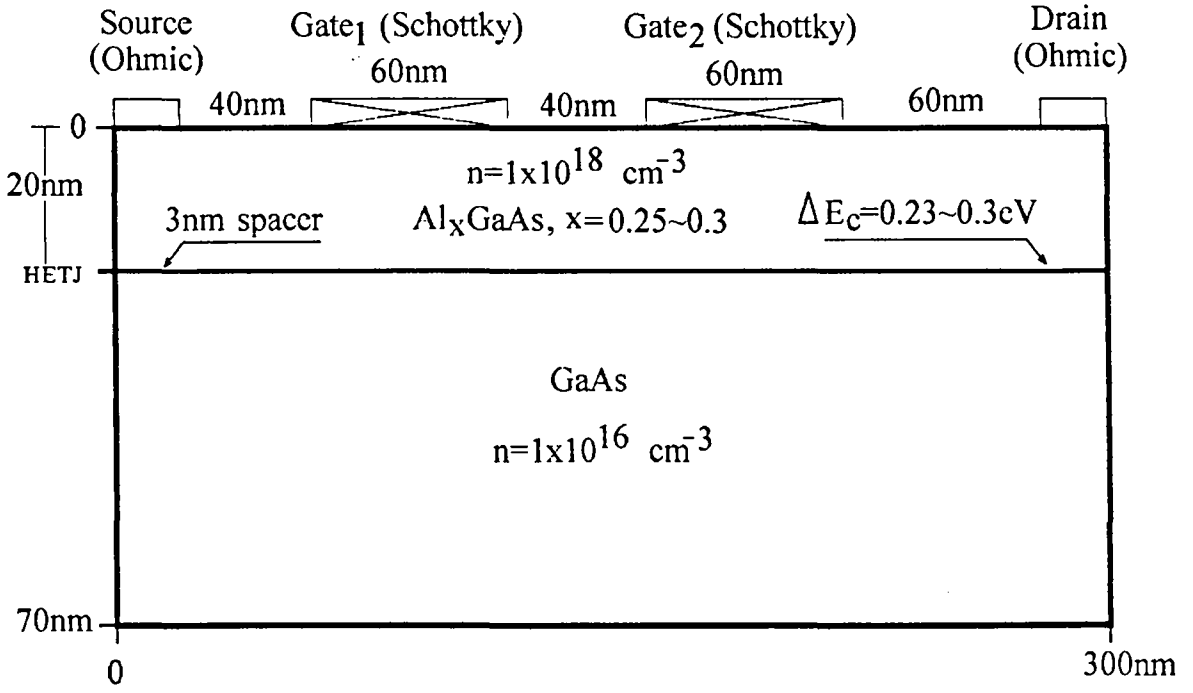


Figure 4.5 structural parameters and dimensions of the simulated MODFET device

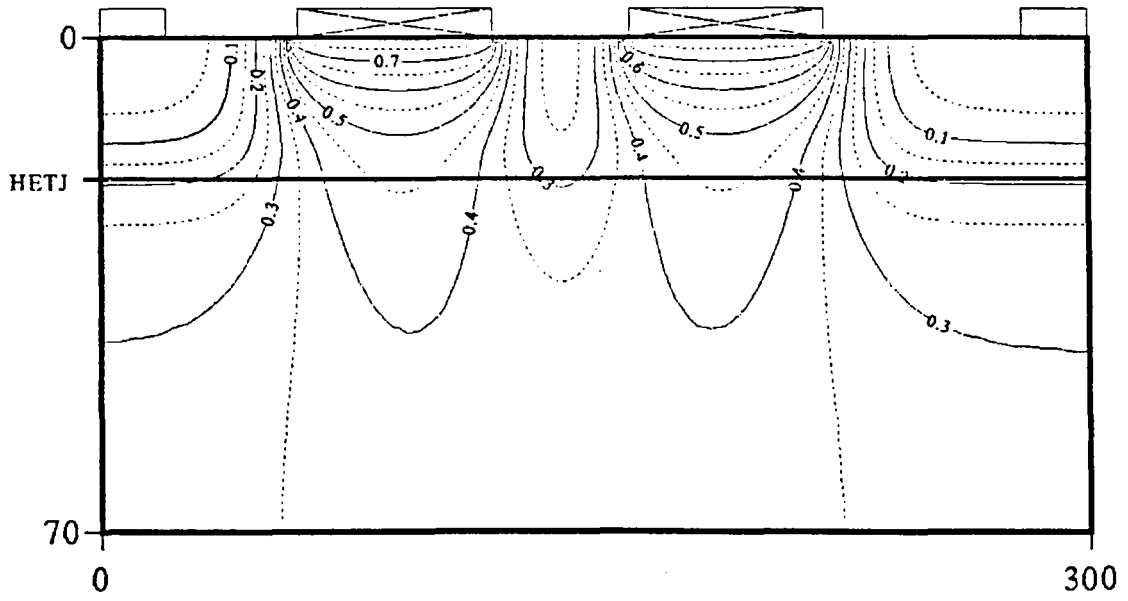


*Figure 4.6 (a) equilibrium distribution of potential contours for the device described in Figure **4.5***

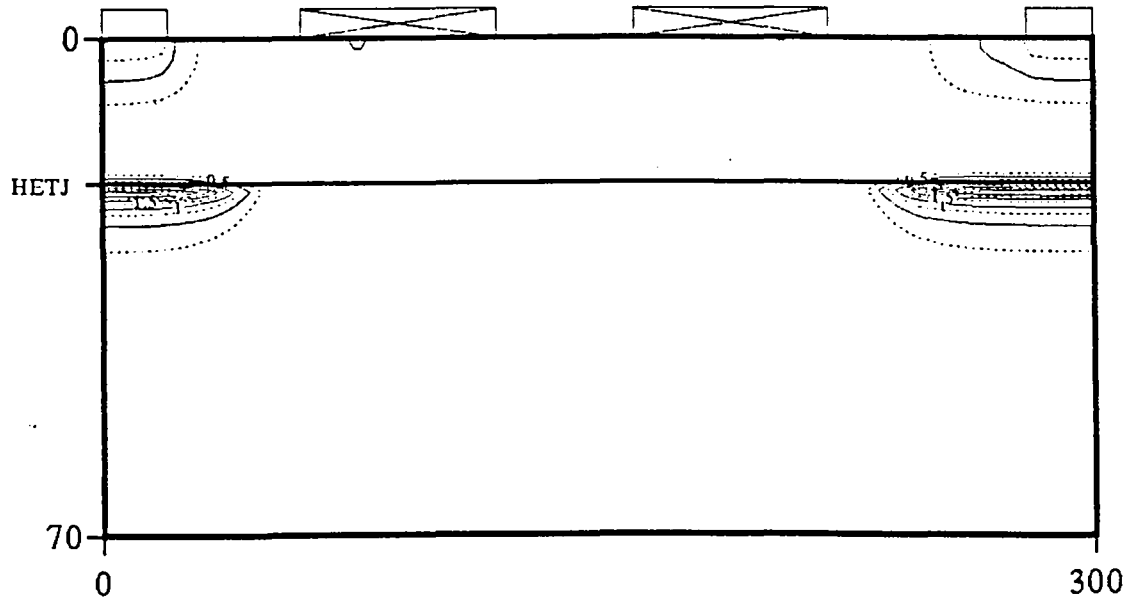


*Figure 4.6 (b) equilibrium distribution of charge contours for the device described in Figure **4.5***

It can be seen that the built-in potential of 0.8 Volts depletes the active region at the heterointerface beneath and between the two gates. This illustrate a situation where there exists a parabolic like quantum-well in the channel between the two gates but in the situation considered, it is empty despite the relatively high operating temperature. Otherwise, it can be noticed that the potential contours are highly symmetric with high fringing effect for this structure on nano meter scale as expected.

To go further in the analysis, it could be useful to investigate now the importance of the degeneracy effect on the resulting potential and Fermi-level distributions. For the latter we used alternatively both Boltzmann and Fermi-Dirac statistics.

Figure 4.7 depicts the x-variation of potential in the channel in close proximity to the heterojunction in the low-gap material, the Fermi-level in this figure is calculated using Boltzmann statistics (a) and Fermi-Dirac statistics (b). A constant Fermi-level is achieved using Boltzmann statistics (a) whereas a variation in E_f is obtained with Fermi-Dirac statistics (b). Let us recall that the choice between Boltzmann and Fermi-Dirac statistics depends on the doping level in bulk material. In contrast, in heterojunction the carrier concentration always reaches high values such that the Fermi-level is pushed in the conduction band.

In order to fix some magnitudes relative to the degeneracy issue while including both the carrier concentration and temperature effect, let us recall the well known degeneracy limit [31] :

$$n/(k_B\theta)^{3/2} \leq 10^{20}\ cm^{-3}.\ eV^{3/2} \qquad \textit{Eq. 4-70}$$

In fact, the above inequality (Eq. 4-70) is based on the definition of the so called degeneracy factor γ_d. The last is defined for the first time in reference [32] to be the ratio between the average electronic energy calculated with Boltzmann statistics ($3/2\ k_B\theta$) and the average energy calculated using Fermi-Dirac statistics for the same temperature and electron concentration. This definition can be described mathematically as,

$$\gamma_d = \frac{f_{3/2}(n, k_B\theta)}{f_{1/2}(n, k_B\theta)} \qquad \textit{Eq. 4-71}$$

where $f_{3/2}$ and $f_{1/2}$ are the Fermi integrals of orders 3/2 and 1/2 respectively. It is worth mentioning that the inequality of Eq. 4-70 insures that $\gamma_d \cong 1$, under this condition the Boltzmann statistics are valid.

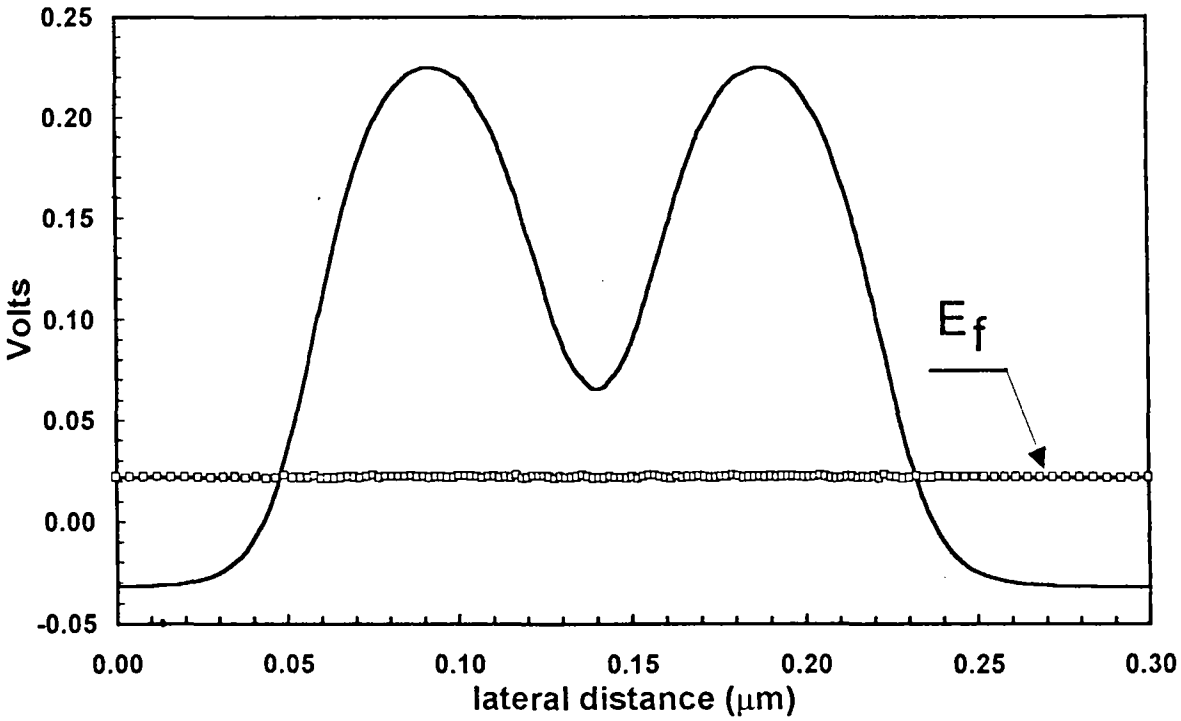


Figure 4.7 (a) lateral potential profile in the channel just after the heterojunction, the Fermi level is calculated using Boltzmann statistics.

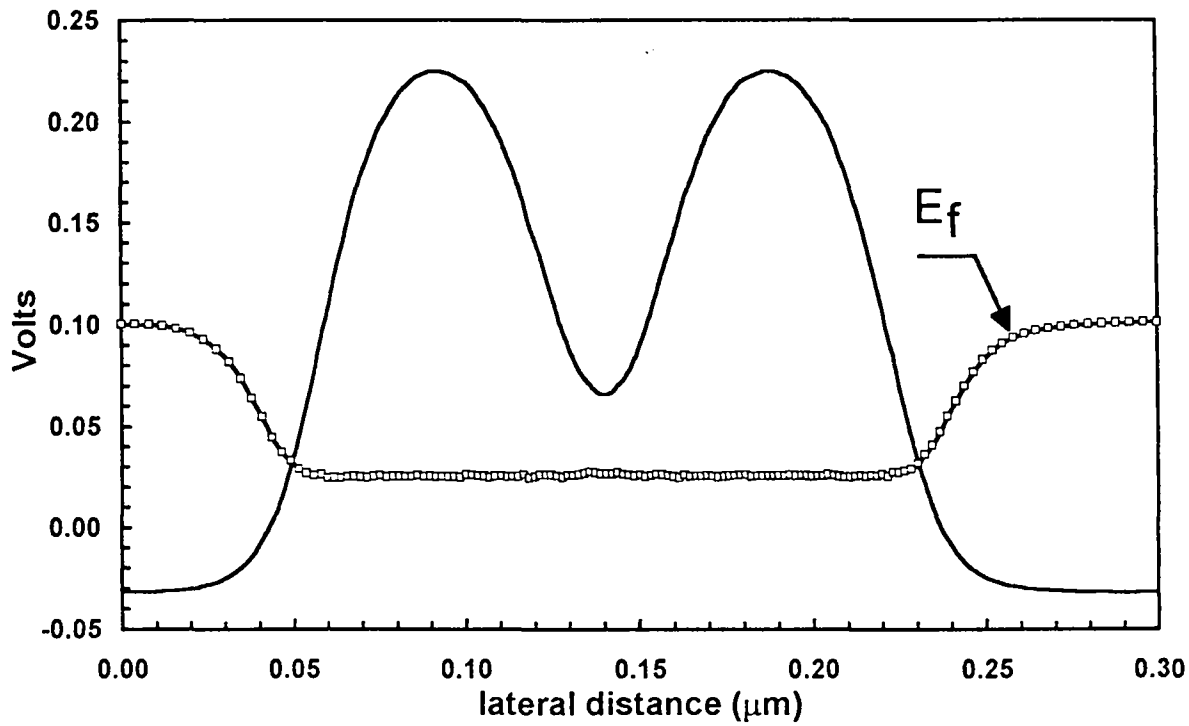


Figure 4.7 (b) lateral potential profile in the channel just after the heterojunction, the Fermi level is calculated using Fermi-Dirac statistics.

Turning now to the vertical variations (y-direction) plotted under both statistics in Figure 4.8, it can be seen that the Fermi-level exhibits some discontinuity particularly using a Fermi-Dirac statistics.

From the Boltzmann statistics it was checked out that the slight discontinuity (≈ 8.3 meV) is invariant versus the lateral position (x-direction). Consequently, the discontinuity is independent of the ratio between the charge concentration at both sides of the heterojunction. In contrast, for the Fermi-Dirac statistics the more pronounced discontinuity is a varying function of lateral position. Besides, the Fermi-level in (b) is not constant in both the wide-band gap and narrow band-gap materials. Indeed, the above mentioned conservation equations are not valid for the applications that having effective mass discontinuities.

To verify this effective mass dependence issue we plot in Figure 4.9 the vertical potential distribution associated with the calculation of Fermi-level for both the light and heavy effective masses. It is clear that the Fermi-level has a constant value independent of the value of the effective mass.

We are now discussing the inclusion of the degeneracy criteria in the conservation equations discussed above, then after we shall investigate the effect of effective mass discontinuity at the heterointerface on the continuation of the Fermi-level. We will start by the model proposed by Azoff in reference [32] which will be denoted hereafter by Azoff model.

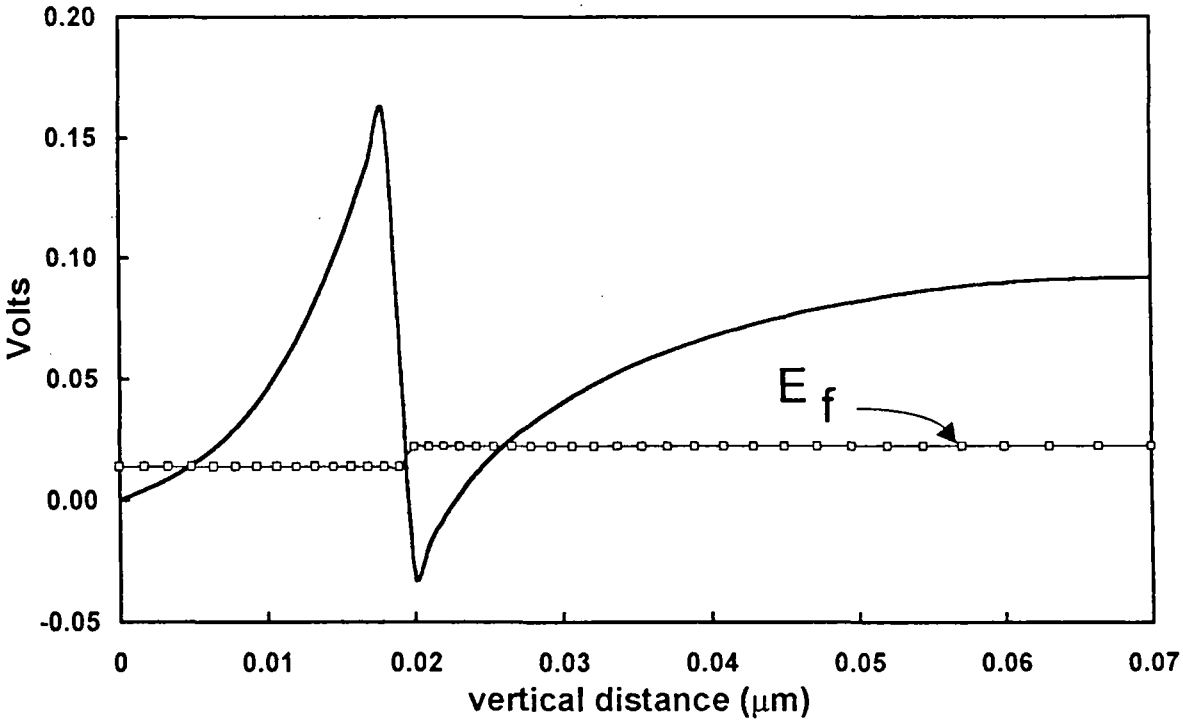


Figure 4.8 (a) vertical potential profile in the region under the source electrode, the Fermi level is calculated using Boltzmann statistics.

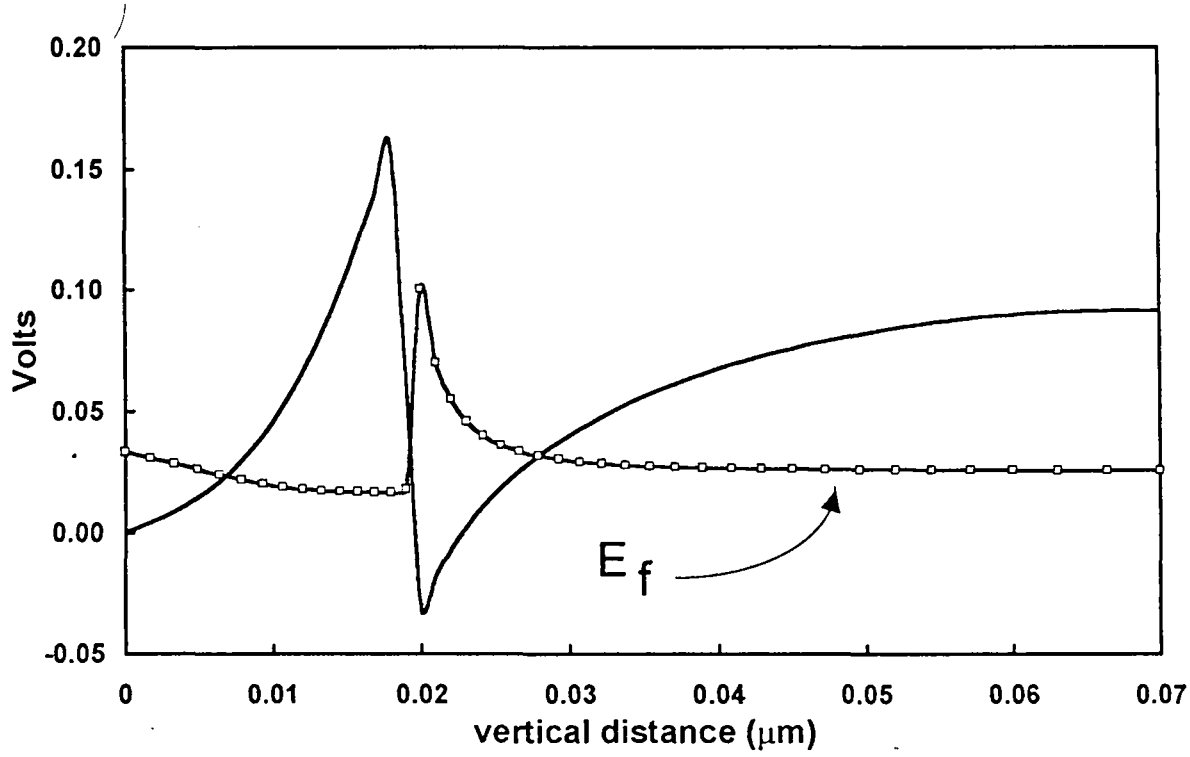


Figure 4.8 (b) vertical potential profile in the region under the source electrode, the Fermi level is calculated using Fermi-Dirac statistics

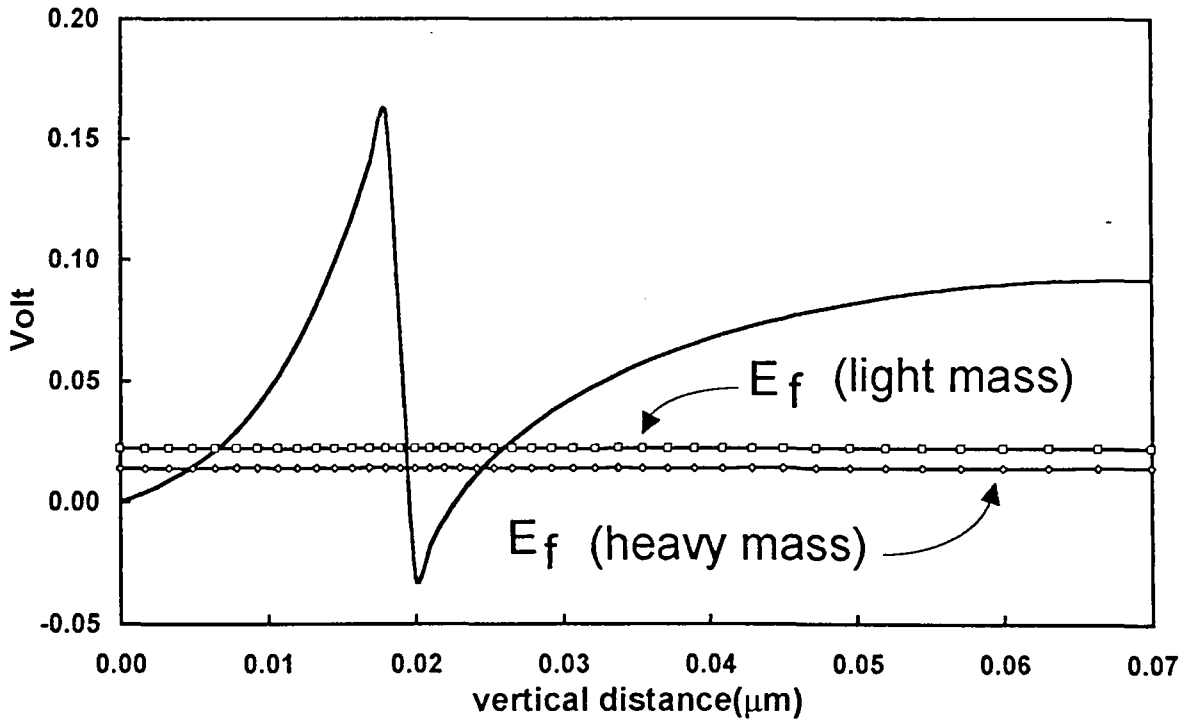


Figure 4.9 vertical potential profile in the region under the source electrode, the Fermi level is calculated using Boltzman statistics for both heavy and light masses.

4.10 Azoff degeneracy model :

The basic idea of this model is to replace the thermal energy ($3/2\ k_B\theta$) which is the average energy calculated at thermal equilibrium in Boltzman statistics distribution (Eq. 4-72) by the average energy calculated by means of Fermi-Dirac statistics.

$$3/2\,k_B\theta = \frac{\int_0^\infty G_{3D}\ \exp\left(\varepsilon_{fB} - \varepsilon / k_B\theta\right)\varepsilon\, d\varepsilon}{\int_0^\infty G_{3D}\ \exp\left(\varepsilon_{fB} - \varepsilon / k_B\theta\right)\ d\varepsilon} \qquad \textit{Eq. 4-72}$$

It can be shown that the relevant results can be expressed as a function of $k_B\theta$ by introducing a correction factor γ_d (Eq. 4-73) (see also Eq. 4-71).

$$3/2\, k_B \gamma_d \theta = \frac{\int_0^\infty G_{3D} f_{1/2}\, \varepsilon\, d\varepsilon}{\int_0^\infty G_{3D} f_{1/2}\, d\varepsilon} \qquad \textit{Eq. 4-73}$$

This means that the total energy (drift +thermal) in the above conservation equations is mathematically described in this approach by the following equation,

$$\varepsilon = \frac{1}{2} m^* v^2 + \frac{3}{2} k_B \gamma_d \theta \qquad \textit{Eq. 4-74}$$

In other words, this is equivalent to replacing the electronic temperature θ by an effective temperature ($\gamma_d\, \theta$) in the above conservation equations.

In order to investigate the validity of the well established E. M. Azoff model [32] (Eq. 4-74) also used by many others ([26], [31]), we applied Eq. 4-74 for simulating the same device described in Figure 4.5. The two dimensional steady state potential contours along with carrier concentration distributions are depicted in Figure 4.10.

The main difference in Figure 4.10 and Figure 4.6 is a spreading in carrier 2D patterns in the populated regions at the heterojunction. On the other hand, it can be seen that the potential contours are quantitatively modified with notably a significant decrease in the vertical field (y-direction) in the channel. This moderates the conduction band bending effect resulting in a less confinement.

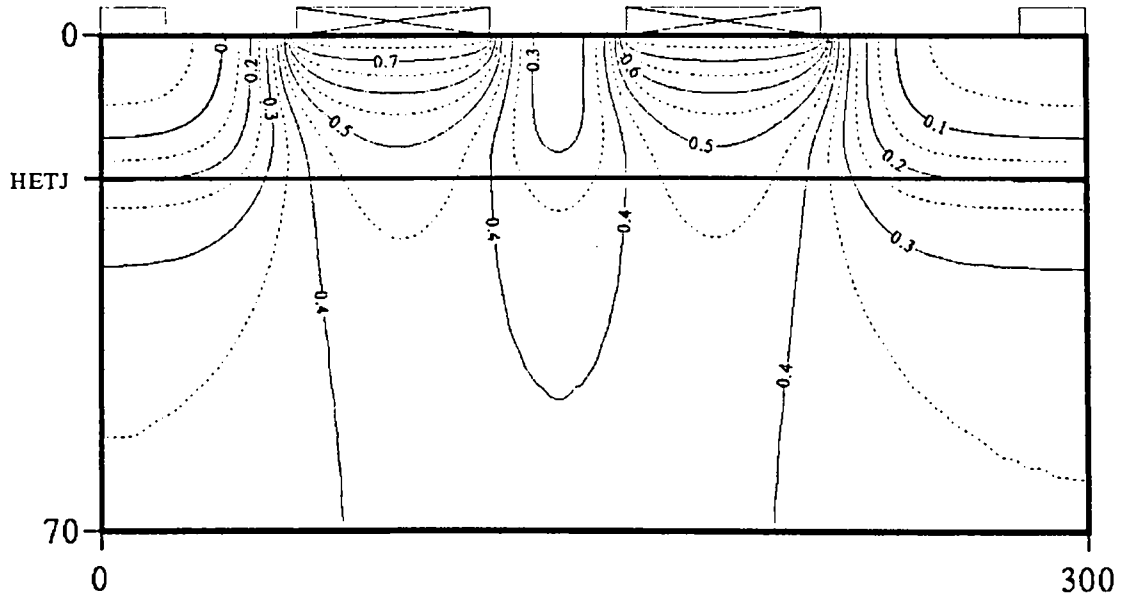


Figure 4.10 (a) equilibrium distribution of potential contours for the device described in Figure 4.5 with Eq. 4-74.

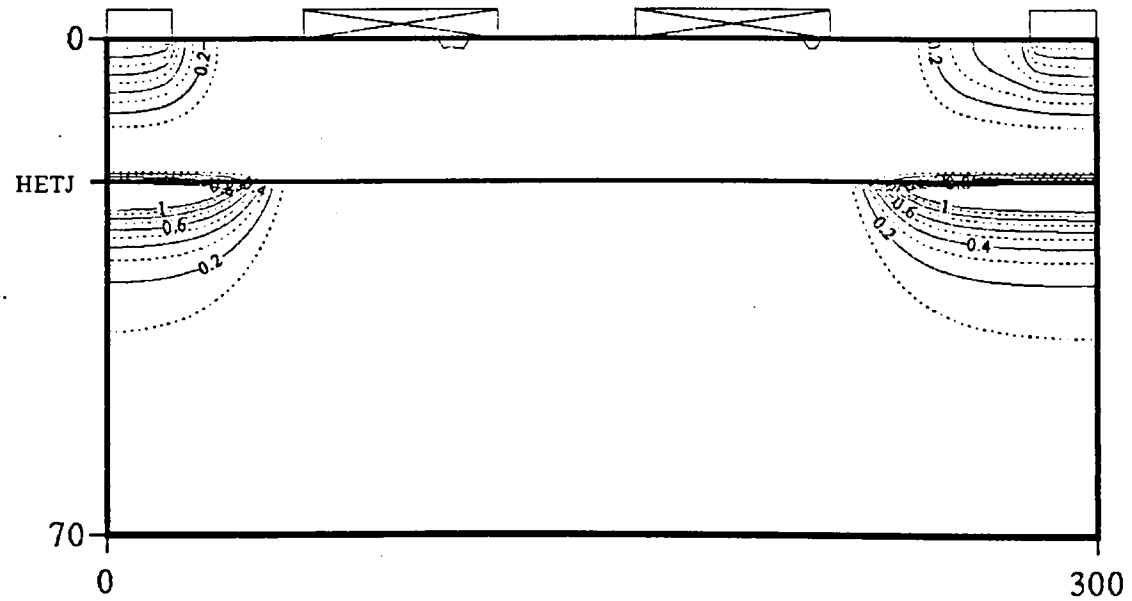


Figure 4.10 (b) equilibrium distribution of charge contours for the device described in Figure 4.5 with Eq. 4-74.

Let us now investigate the impact of Azoff correction (Eq. 4-74) on the Fermi-level profile. Starting first with the problem concerning the lateral profile, Figure 4.11 is a plot of the Fermi-level (E_f) profile from the source to the drain in the channel in close proximity to the heterojunction. This profile shows unphysical space dependence. On this basis, it seems interesting to investigate the variation of E_f at various depths in the vicinity of the heterojunction. Three cases are considered in Figure 4.12, where a, b and c situations corresponds to locations on the channel, very close to the heterojunction and in the spacer layer respectively. Whereas the space dependence is slight for a-case, big differences are detected for b- and c-case. Hence, it can be seen that Azoff model failed to achieve a perfect flat Fermi-level along the x-direction.

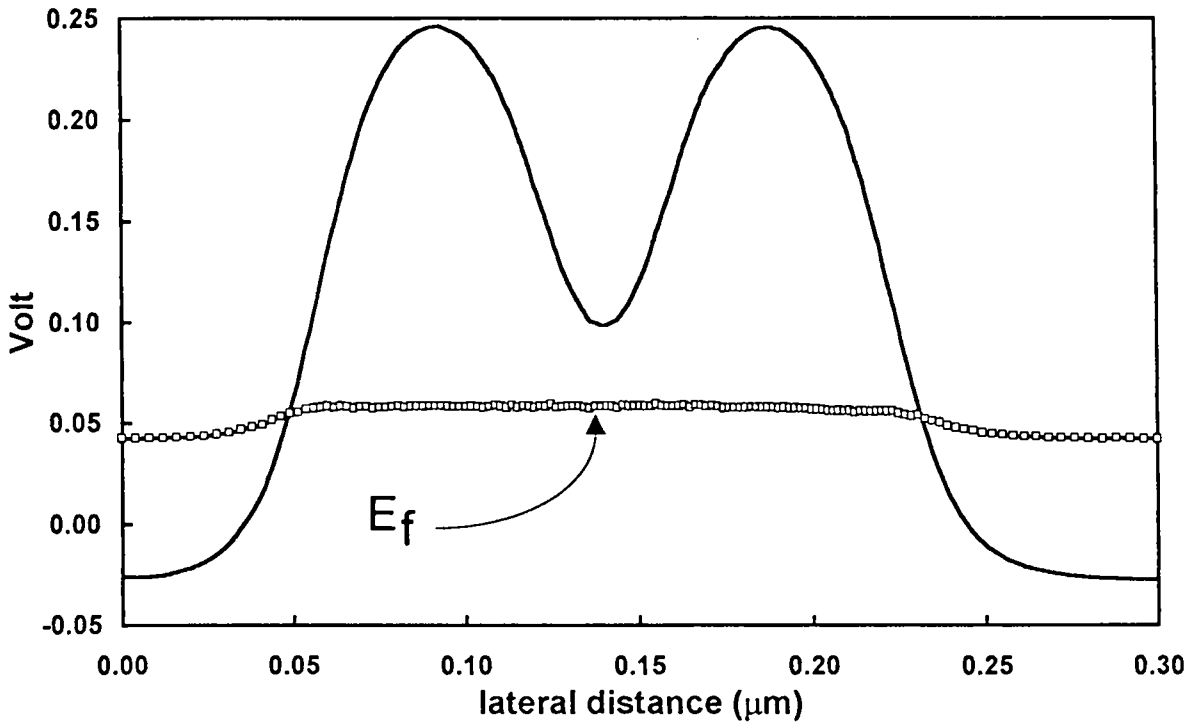


Figure 4.11 lateral potential profile in the channel in close proximity the heterojunction, the Fermi level is calculated using Fermi-Dirac statistics

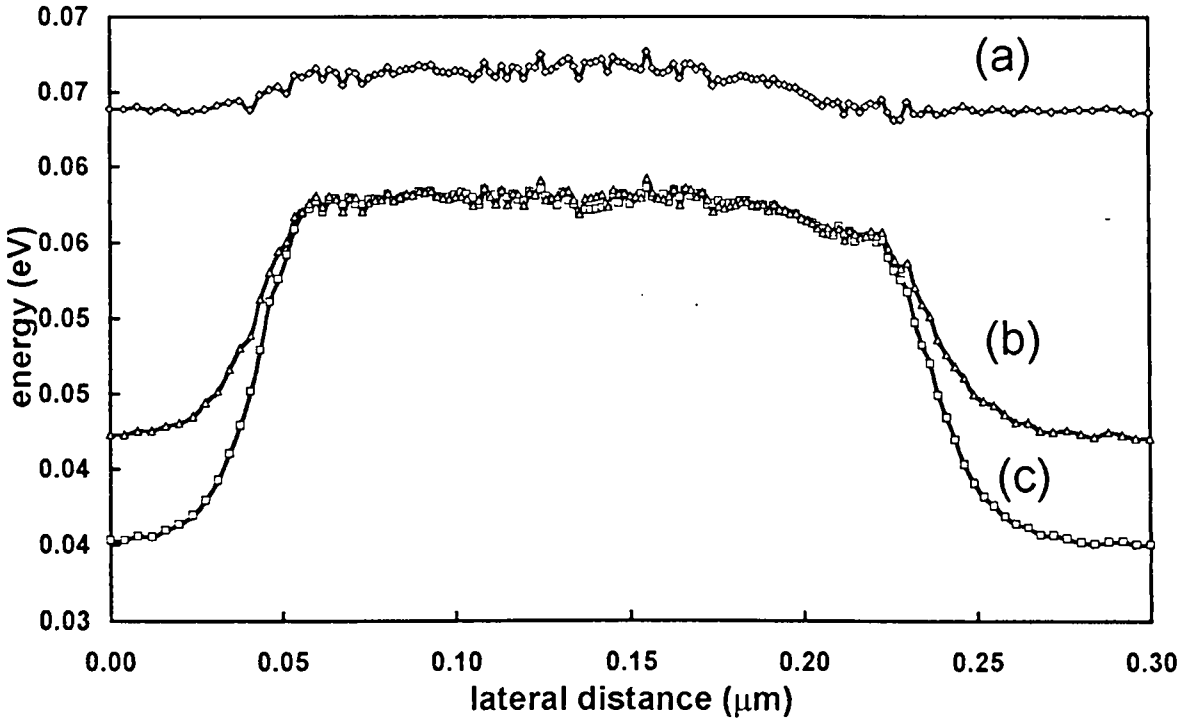


Figure 4.12 lateral Fermi-level profile at different depths, (a) in the channel,(b) in the channel close to heterojunction,(c) in the spacer layer. Fermi level is calculated using Fermi-Dirac statistics

Also it can be noticed that, the discontinuity of Fermi-level (the difference between a-case and c-case) does not maintain a constant value in the x-direction. In contrast at the region under the gates where the electron concentrations are low enough to apply Boltzman statistics, this discontinuity is kept constant equal the same value depicted in Figure 4.9 ($\approx 8.3\ meV$).

Now, we will investigate the degeneracy effect in the vertical or growth direction (y-direction). For this purpose the vertical potential profile along with the Fermi-level are depicted in Figure 4.13. From this figure, one can also notice the discontinuity in the Fermi-level profile at the heterojunction. Now the discontinuity is not only a function of the effective mass jump but is also a function of the degree of the abrupt change in the electron concentration. This gives rise to an additional complication to the interface problem that has to be solved to achieve a flat Fermi-level profile.

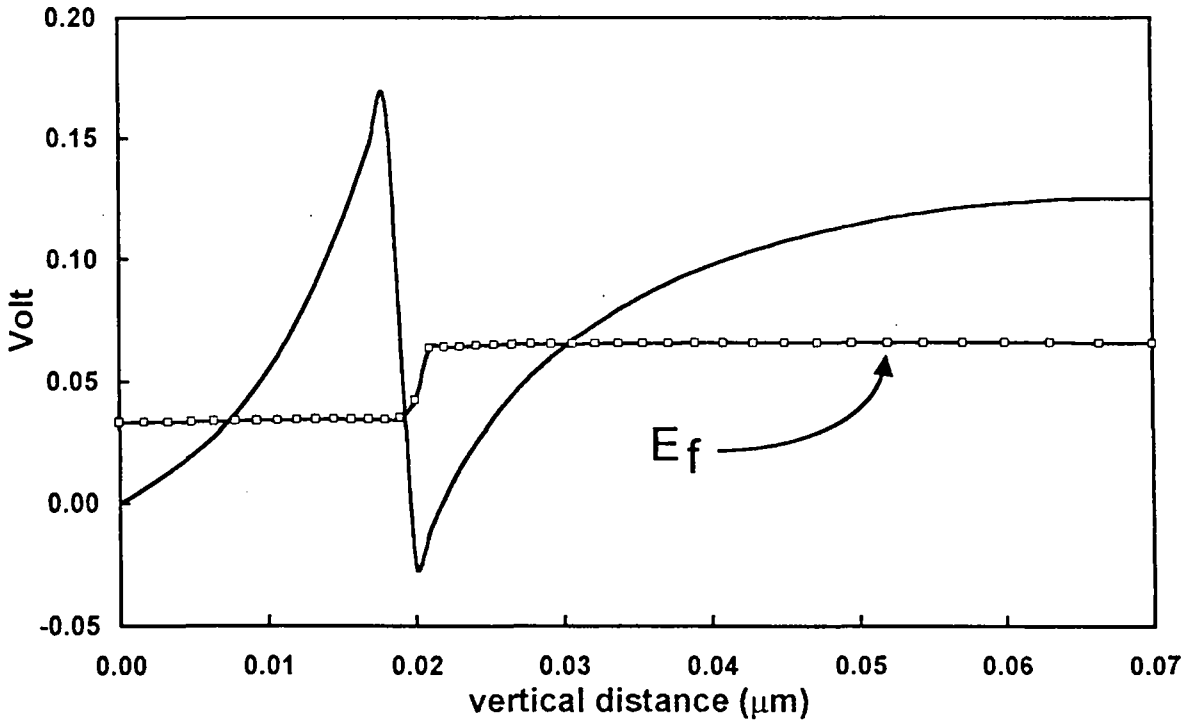


Figure 4.13 vertical potential profile in the region under the source, the Fermi level is calculated using Fermi-Dirac statistics

At last, to give an overall picture of the Fermi-level distribution in the cross section of the device, the 2D-contoures (labelled in meV) of the Fermi-level are depicted in Figure 4.14.

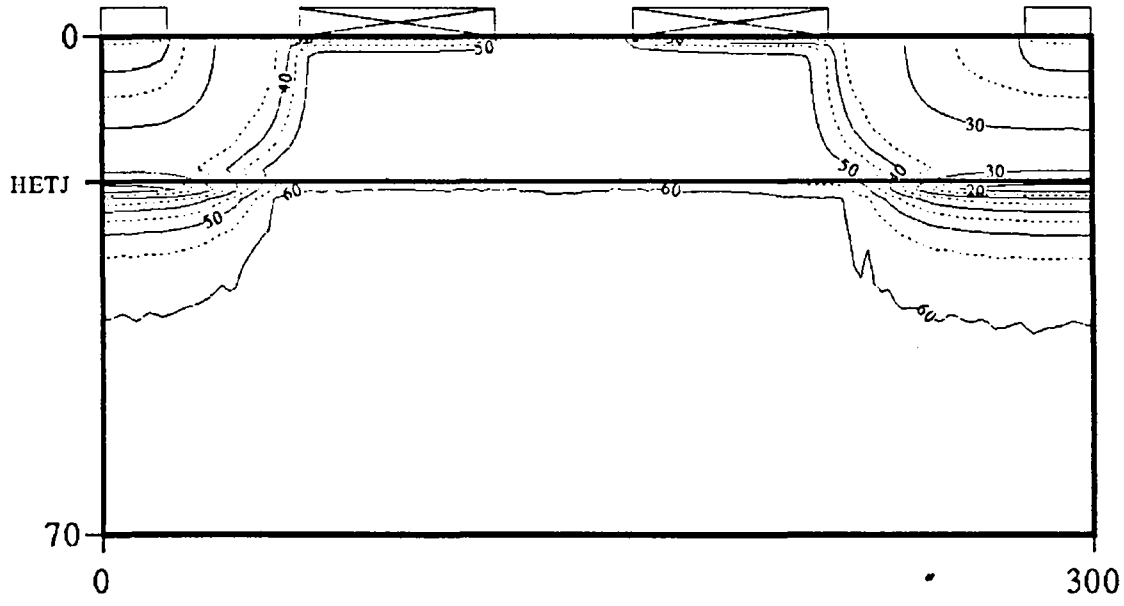


Figure 4.14 Fermi-level contours labelled in meV.

One can deduce that the Fermi-level contours (Figure 4.14) follow the charge contours depicted in Figure 4.10b.

Now we are proposing an original model that able to give a flat profile based on a physical picture of thermal energy.

4.11 Novel degeneracy model :

In Azoff model, the average energy defined mathematically by Eq. 4-73 is considered totally as a thermal energy. However, it is quite physical to assume that the random motion of electrons at a finite temperature involve a transition between occupied and empty states [33]. In other words, it is necessary to include the degree of occupancy of states. Under this basis one can expect an effective threshold level of energy for the onset of a particle thermal motion. In the following we will denote this threshold as the degenerate potential.

Mathematically, the thermal energy ($3/2\ k_B\theta$) can be calculated in the framework of the Fermi-Dirac statistics as follows,

$$3/2\,k_B\theta = \frac{\int_0^{\infty} G_{3D} f_{1/2}\left(1 - f_{1/2}\right) \varepsilon\, d\varepsilon}{\int_0^{\infty} G_{3D} f_{1/2}\ \ d\varepsilon} \qquad \textit{Eq. 4-75}$$

where the term $(1 - f_{1/2})$ is introduced for describing the available empty states needed for the thermal agitation. The above formula is quite general. When there is a full occupancy ($f_{1/2} = 1$) below the degenerate potential the contribution in the integral vanishes. This energy term can be considered as a potential energy. Its average value reads,

$$V_{d-average} = 3/2\ k_B\ \theta\ (\gamma_d - 1) \qquad \textit{Eq. 4-76}$$

Now, we have to derive this degenerate potential considered as a correction factor. For this purpose we will consider the limit of zero temperature. In that case the step-like Fermi-Dirac profile gives a straightforward relation between the maximum threshold value and the average degenerate potential. This results in a factor of two such that,

$$V_{d-\max.} = 2 \times 3/2\ k_B\ \theta\ (\gamma_d - 1) \qquad \textit{Eq. 4-77}$$

At last, to treat the maximum degenerate potential (Eq. 4-77) with an analogous way to that of the thermal part of energy in the conservation equations, we have to divide Eq. 4-77 by the factor (3/2). This results in the degeneracy additional potential to be,

$$V_d = 2\ k_B\ \theta\ (\gamma_d - 1) \qquad \textit{Eq. 4-78}$$

Finally, it can be shown numerically that,

$$V_d = \varepsilon_f|_{(Fermi-Dirac)} - \varepsilon_f|_{(Botzmann)} \qquad \textit{Eq. 4-79}$$

It is worth mentioning that the effect of the above degeneracy correction is included in Eq. 4-23 and Eq. 4-24 by including an additional electric field defined by the spatial gradient of V_d depicted in Eq. 4-78 or Eq. 4-79.The above formalism is systematically applied to the same structure described in Figure 4.5 and the two dimensional simulation steady state results are shown in Figure 4.15. From Figure 4.15 (a) one can deduce that the potential contours under the gate are shifted downwards with respect to those plotted both in Figure 4.6 and Figure 4.10. Although the channel is still practically depleted (see Figure 4.15 b) the device exhibits a less gate control on the 2DEG with respect to the above simulations of Figure 4.6 and Figure 4.10. For the illustration purpose, we showed in Figure 4.10 (c) the two dimensional contour representation of the degenerate potential V_d. The correlation between the charge distribution and V_d is quite clear.

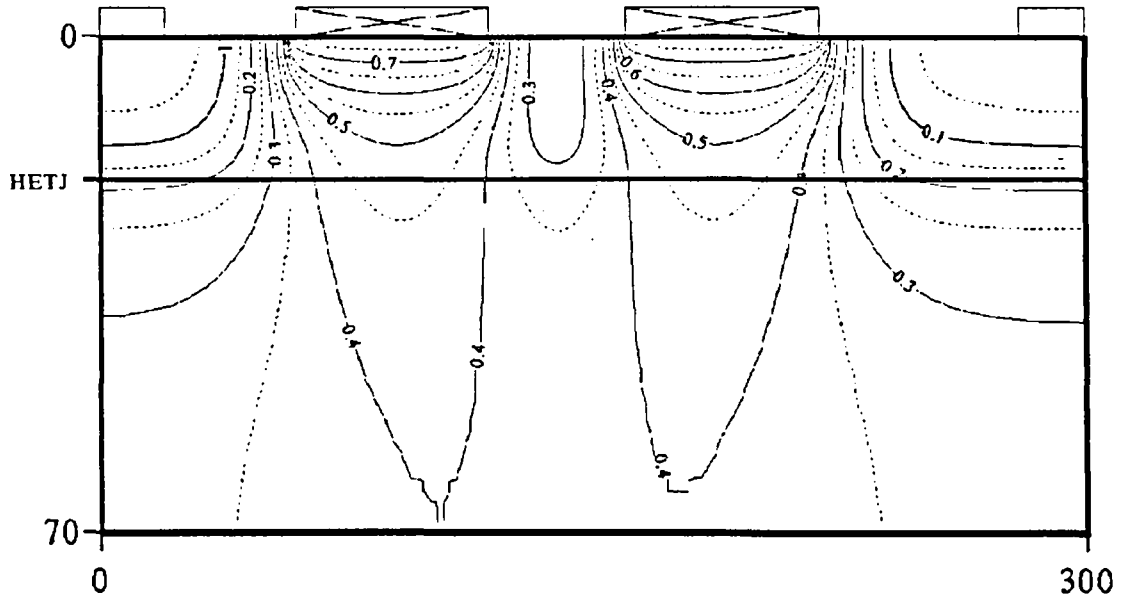


Figure 4.15 (a) steady state equi-potential lines resulting from the new generate model.

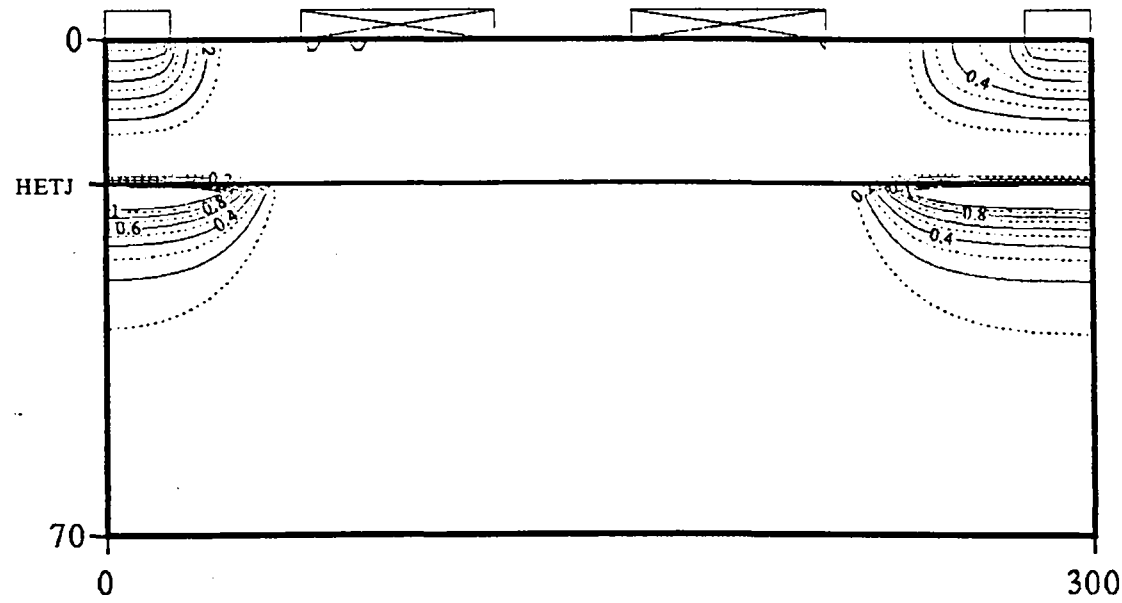


Figure 4.15 (b) steady state charge contours resulting from the new degenerate formalism.

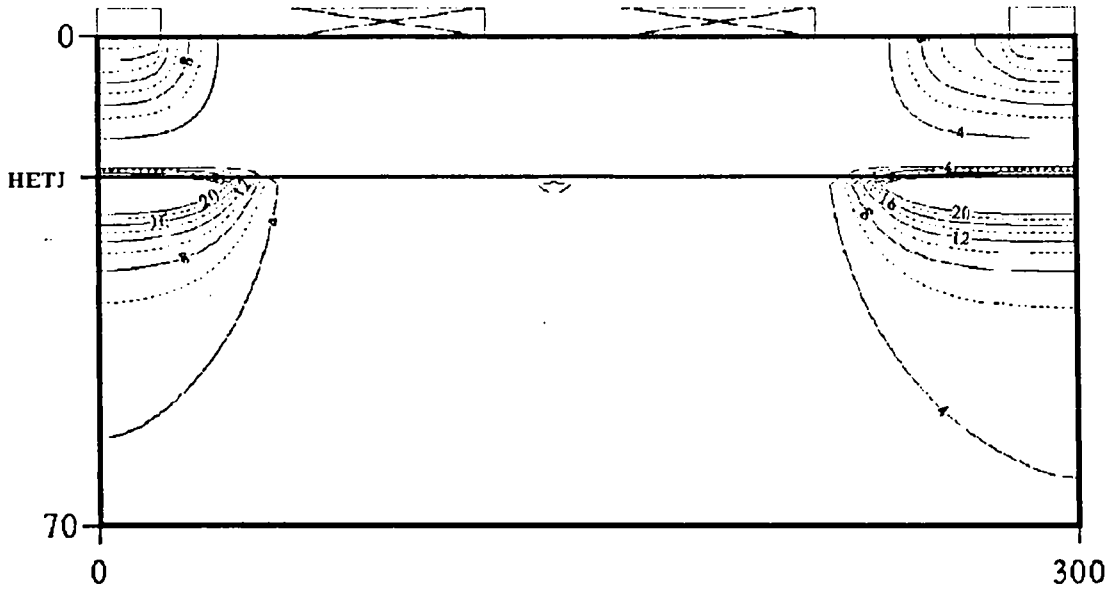


Figure 4.15 (c) the degenerate potential contours (V_d) in meV associated with the above shown charge contours

On the other hand, to illustrate the advantage of the new formalism over Azoff model discussed in section 4.10, we plotted the lateral potential distribution along with the Fermi-level profile in the channel just under the heterojunction in Figure 4.16. the degenerate potential V_d are also illustrated.

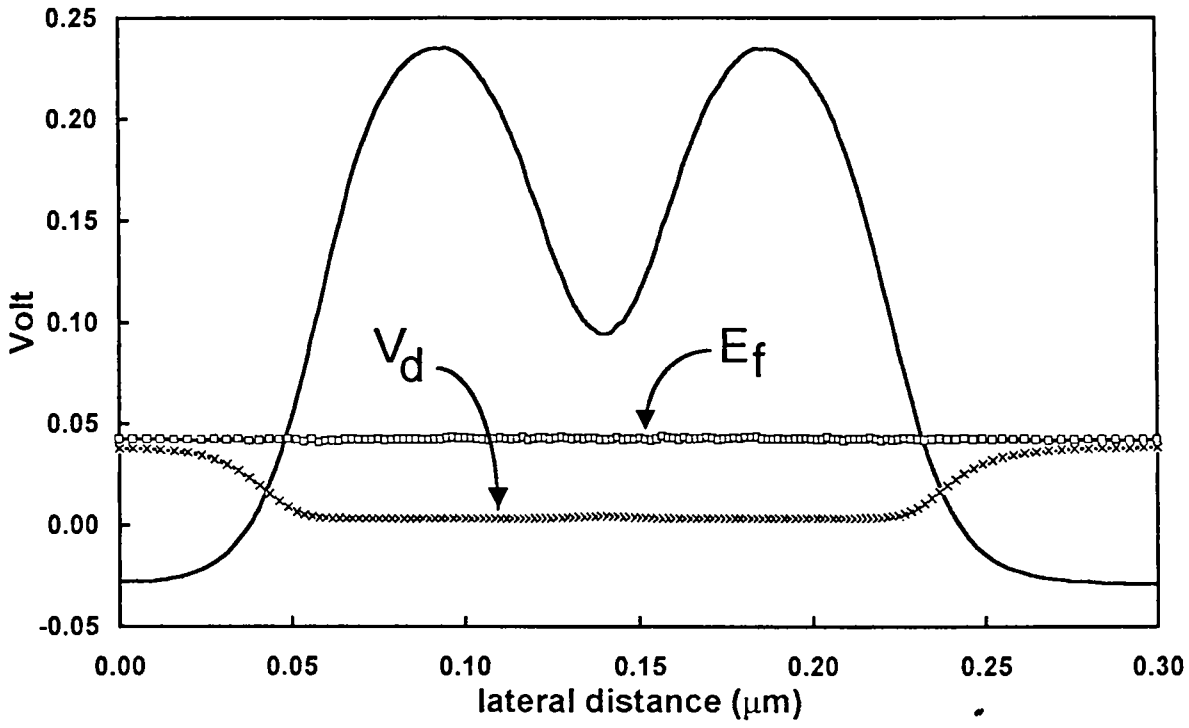


Figure 4.16 the lateral potential distribution along with the Fermi-level profile and the degenerate potential V_d in the channel just under the heterojunction

The flat Fermi-level in the above figure is remarkable as compared to the space variation obtained by means of the model of Azoff illustrated previously in Figure 4.11 and Figure 4.12. Also plotted is the degenerate potential V_d along the lateral axis.

Nevertheless, in the growth direction there still subsists a slight jump in the Fermi-level profile at the heterointerface. This is illustrated in Figure 4.17. At first glance, the discontinuity in the Fermi-level in this figure ($\approx 10\ meV$) is identical to that obtained in section 4.9 (Figure 4.8 and Figure 4.9). Therefor the last important issue to discuss is the continuity of Fermi-level across the heterojunction considered in the next section.

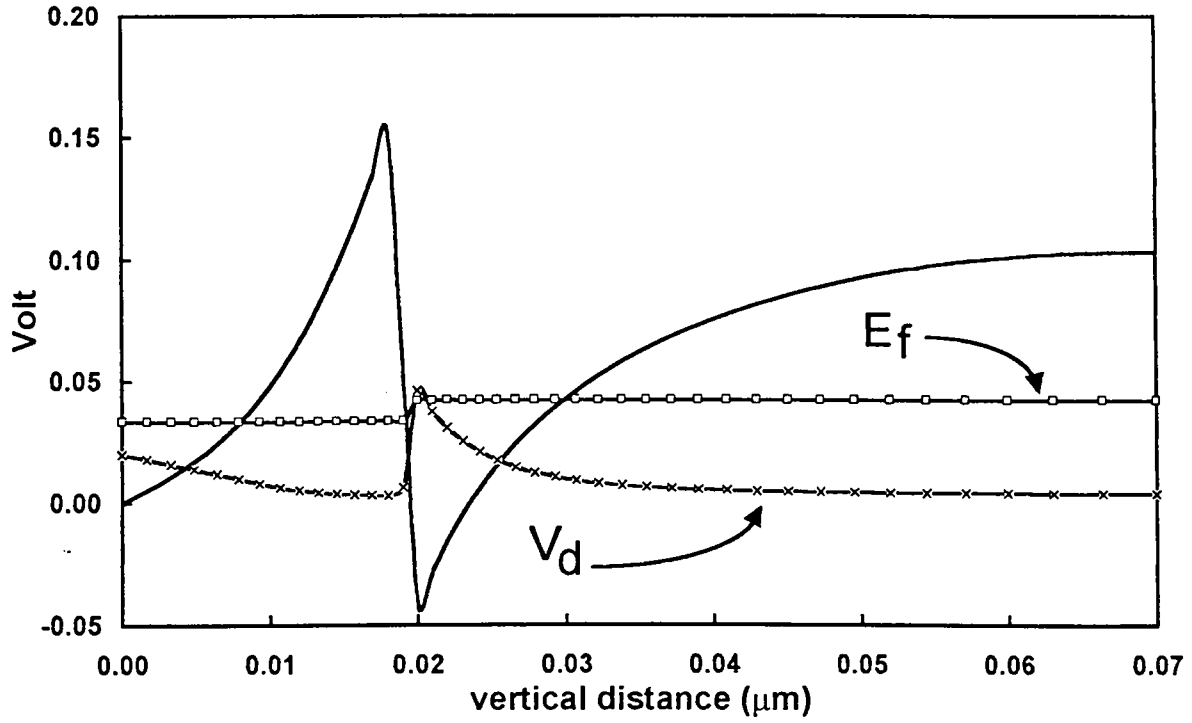


Figure 4.17 the vertical potential distribution along with the Fermi-level profile and the degenerate potential V_d in the region under the source

4.12 Fermi-level continuity across a heterointerface :

From the above discussion one can conclude that the Fermi-level discontinuity is function of two variables. The first one is the discontinuity of the effective mass at both sides of the heterojunction (see also Figure 4.8, Figure 4.9 and Figure 4.17). The second is the electronic temperature at the interface between the two semiconductors (see Figure 4.11 and Figure 4.12). Taking advantage of the independence of this discontinuity on the charge concentration, we will deduce the necessary correction based on the Boltzman statistics distribution for sake of simplicity.

$$n = N_c \exp(\varepsilon_f - \varepsilon_c / k_B\theta) \qquad \textit{Eq. 4-80}$$

where,

$$N_c = 2\left(\frac{2\pi m^* k_B \theta}{h^2}\right) \qquad \textit{Eq. 4-81}$$

By simple algebra we can prove that the additional potential V_{add} required to eliminate the Fermi-level discontinuity reads,

$$V_{add} = \frac{3}{2}k_B\theta \ln\left(\frac{m_1^*}{m_2^*}\right) + \frac{3}{2}k_B(\theta_1 - \theta_2)\ln\left(\frac{m_1^* + m_2^*}{2}\right) \qquad \textit{Eq. 4-82}$$

where θ_1 and θ_2 are the electron temperatures on each side of the heterojunction.

We introduce a difference in the temperature so that this model can be generalised to the model of Azoff discussed in section 4.10. Let us recall that in this model an effective temperature ($\gamma_d\,\theta$) is introduced. In the present case of isothermal analysis the second term vanishes. Given m_1^*=.0836, m_2^*=.067 and $\theta = 292$ K one can calculate 8.364 meV as the discontinuity additional potential. At last we include an additional potential gradient across the heterointerface in Eq. 4-23, the results are represented in Figure 4.18. Both the lateral and the vertical Fermi-level profiles are now flat with a full analysis based on Fermi-Dirac statistics. For completeness we also plotted the two dimensional contours of both the potential and the charge density in Figure 4.19.

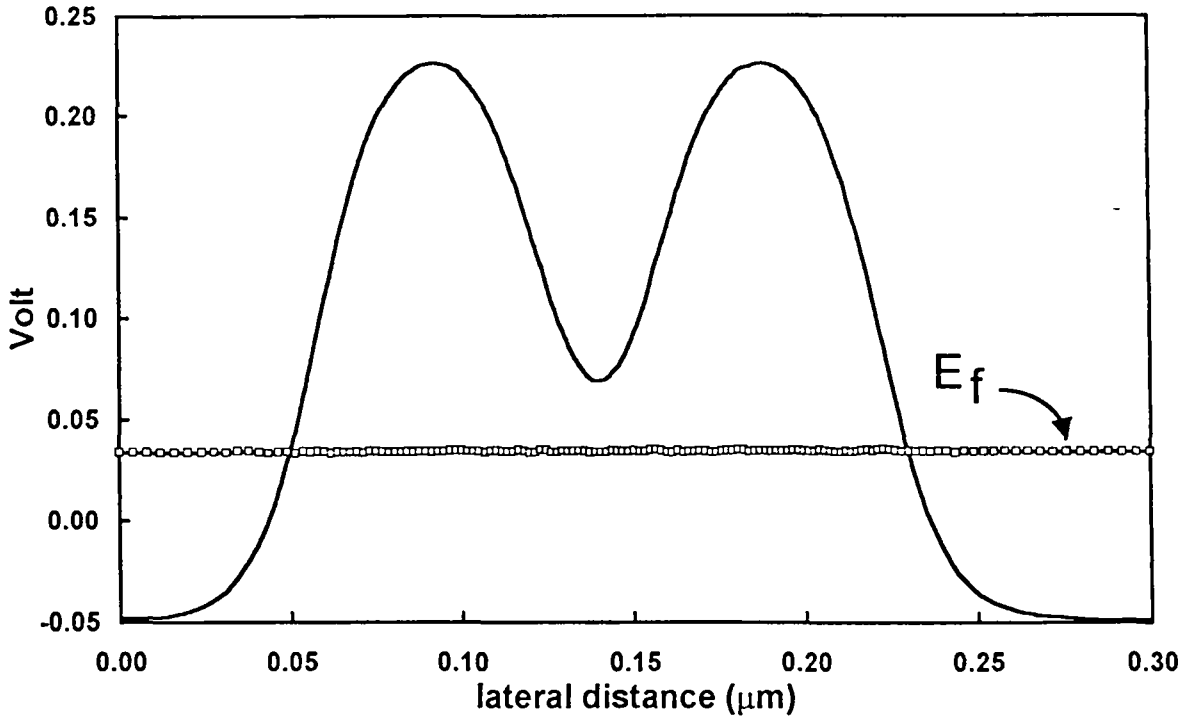


Figure 4.18 (a) the lateral potential distribution along with the Fermi-level profile in the channel just under the heterojunction

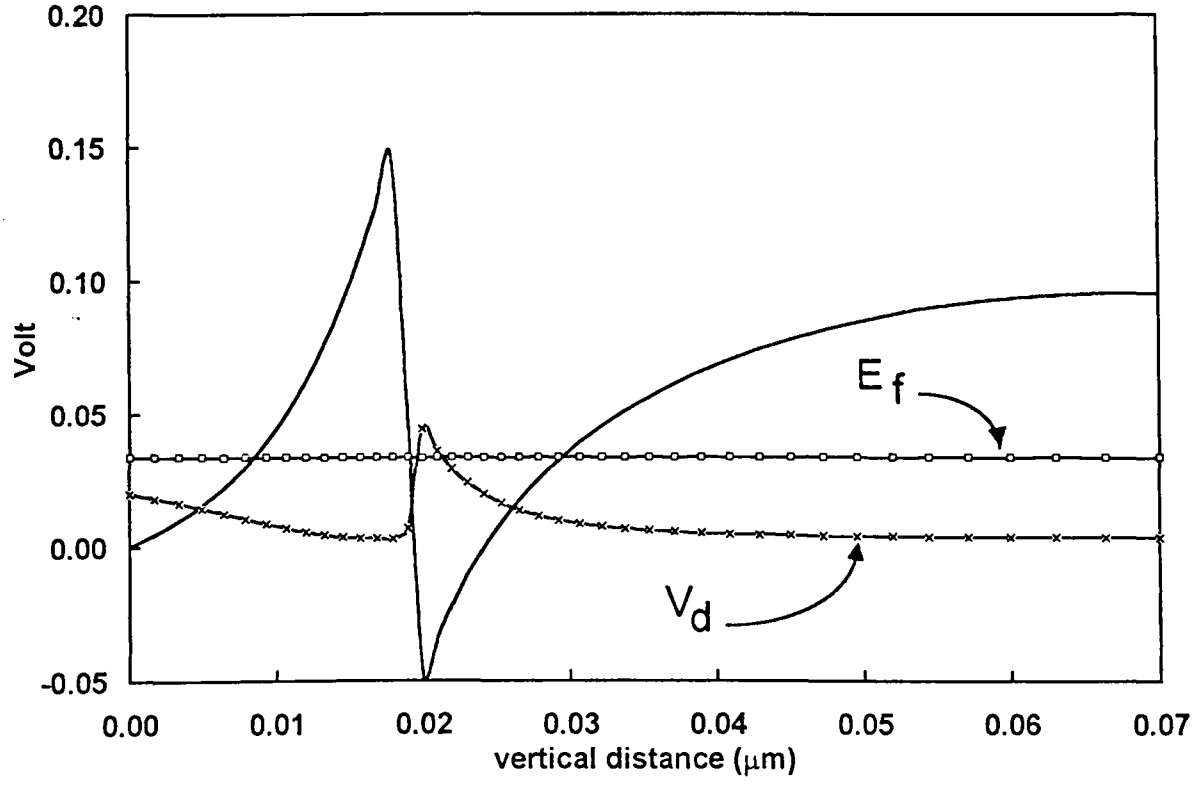


Figure 4.18 (b) the vertical potential distribution along with the Fermi-level profile and the degenerate potential V_d in the region under the source

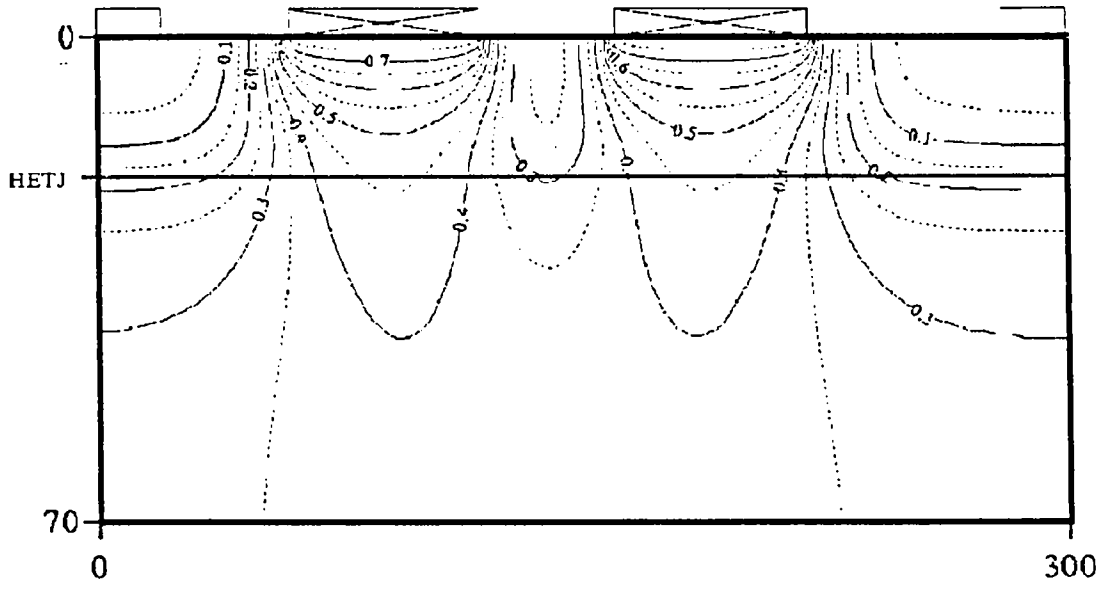


Figure 4.19 (a) steady state equi-potential lines resulting from the new generate model.

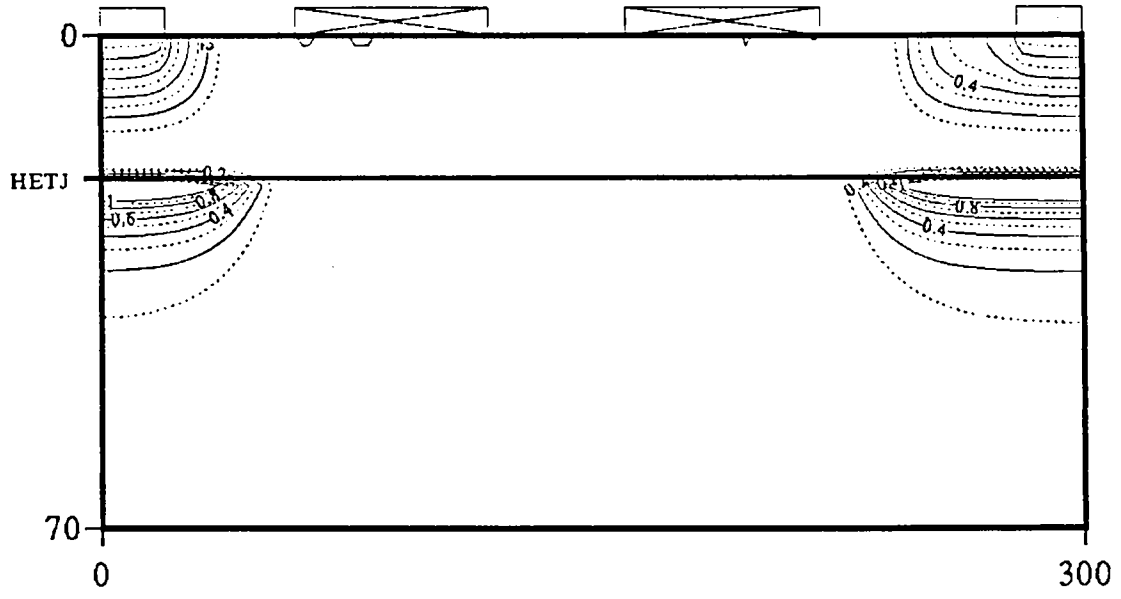


Figure 4.19 (b) steady state charge contours resulting from the new degenerate formalism

4.13 Model application at low temperature (4.2 K) :

From the previous sections, we are now able to simulate degeneracy along with heterojunction devices while conserving the validity of the Fermi-Dirac statistics. At this stage, it remains to study the limitations of applying the hydrodynamic model discussed above under very low temperature conditions. This will concern two key points : (i) the usually used initial condition, (ii) the validity of shallow Ohmic contacts for the source and drain. Also we shall discuss the scaled solution of the above conservation equations.

The first problem we are facing by assuming a low temperature operation concerns the usually used initial conditions :

$$n(x,y) = N_d(x,y) \text{ and } \varepsilon(x,y) = 0 \qquad \textit{Eq. 4-83}$$

This gives rise to numerical instabilities due to the relatively high value of the mobility at 4.2 K. This motivates us to apply a scaled solution for the above conservation equations. The basic idea is to slow down the instantaneous motion of mobile carriers by means of a scale down factor M. All the key relevant figures in the conservation equations are affected by this factor notably the mobility and energy relaxation time. The conservation equations now reads,

$$\frac{\partial n}{\partial t} + \nabla_r.(n\,\vec{v}) = 0 \qquad \textit{Eq. 4-84}$$

$$n\,\vec{v} = n M \mu \vec{F} - M \mu \nabla_r \left(n \frac{K_B \theta}{q} \right) \qquad \textit{Eq. 4-85}$$

$$\frac{\partial \varepsilon}{\partial t} + \vec{v}.\nabla_r(\varepsilon) = q\,\vec{v}.\vec{F} - \frac{1}{n}\nabla_r.(n K_B \theta\,\vec{v}) - \frac{(\varepsilon - w_o)}{\tau_\varepsilon(\varepsilon)/M} \qquad \textit{Eq. 4-86}$$

Where,

- M is a dimensional-less temporal scale factor smaller than unity.

- $t' = t \times M$, is a slowed down time with respect to the real time t.
- all the other symbols have their conventional meanings.

At last, the output data are expressed by suppressing the scale factor.

The above idea is systematically used for simulating the device described previously in Figure 4.5 at 4.2 K. The resulting channel potential profile along with the Fermi-level are depicted in Figure 4.20. The free mobile carrier concentration in the inter-gate zone is so low that it is impossible to define the relevant Fermi-energy. In addition, one can note a slight offset between the left and right hand sides of the figure despite the fact that the device is unbiased. This reflects the drawback resulting from the initial conditions stated in Eq. 4-83. In fact, the built-in gate potential (0.8 V) is seen by this initial charge distribution as an abrupt voltage step. In the beginning of the time iterations, this voltage step depletes the inter-gate zone, finally the electrostatic barriers are formed before the charges reoccupy the inter-gate region. As the temperature is very low (4.2 K) the thermally activated current is not able to surmount the formerly generated barriers.

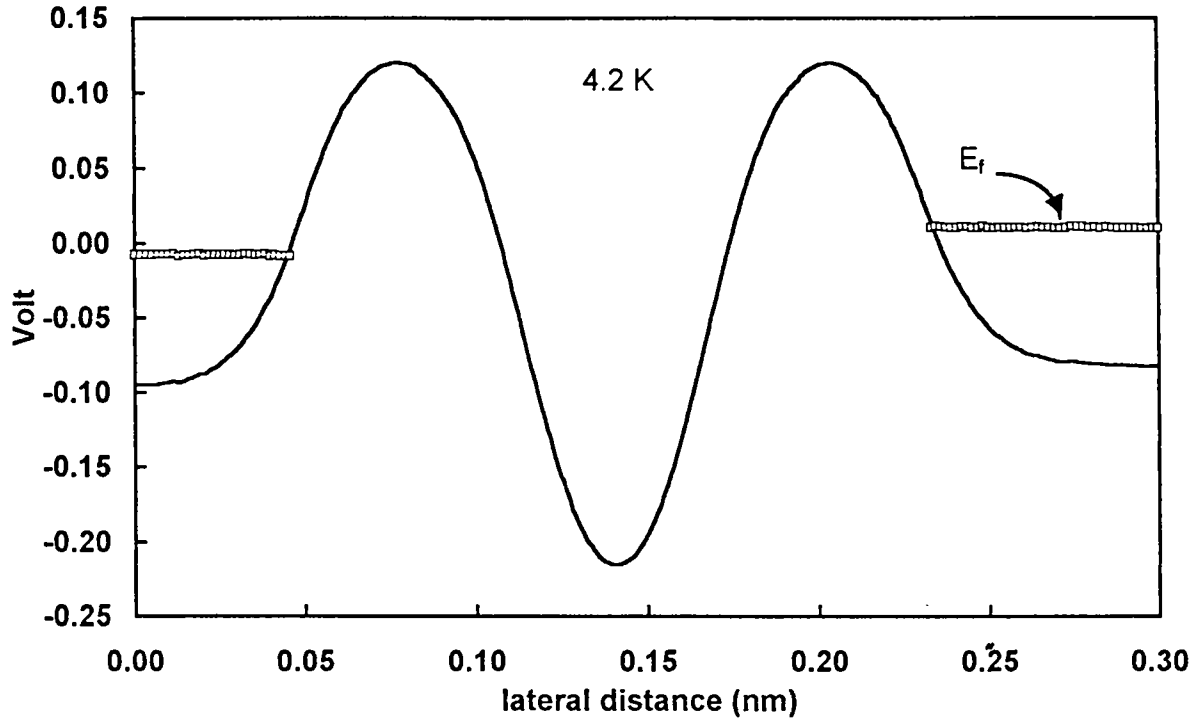


Figure 4.20 potential profile in the channel along with the Fermi-level.

From the other hand, the vertical potential profile along with the Fermi-level in the region under the drain shallow Ohmic contact are depicted in Figure 4.21. In this figure, one can note that, although the temperature is very low to minimise any discontinuity in the Fermi-level across the heterointerface (see also Eq. 4-82), the later level exhibits a non-acceptable discontinuity at the heterojunction. The reason is also that, the heterojunction barrier is created before the Fermi-level reaches its attempted flat profile. At this instant, the thermally activated carrier concentration is not sufficient to recover this unwanted discontinuity. This means that the shallow Ohmic contact in this specific situation is not able to access the 2DEG conduction channel. The solution is to simulate (as in the fabricated device of reference [1]) the deep Ohmic contact.

As a matter of illustration, the two dimensional contours of both, the potential and the carrier concentration are displayed in Figure 4.22.

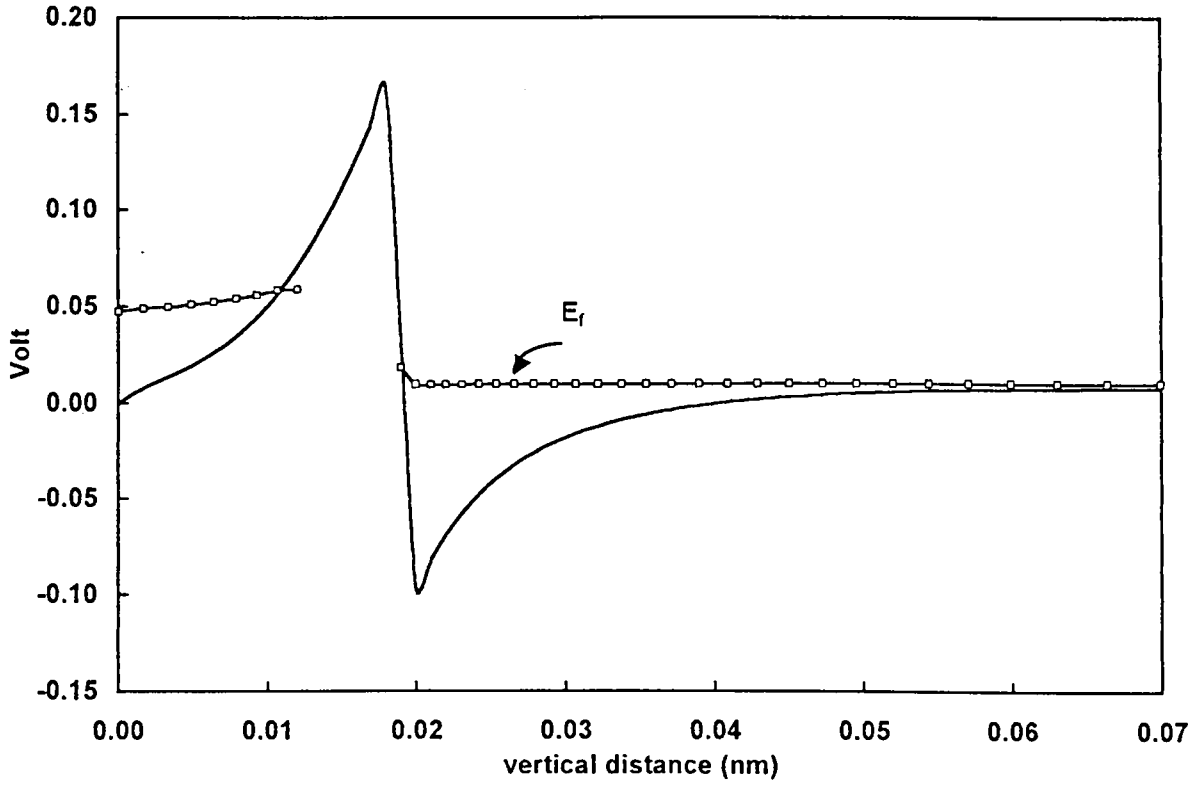


Figure 4.21 vertical potential profile along with Fermi-level in the region under the source.

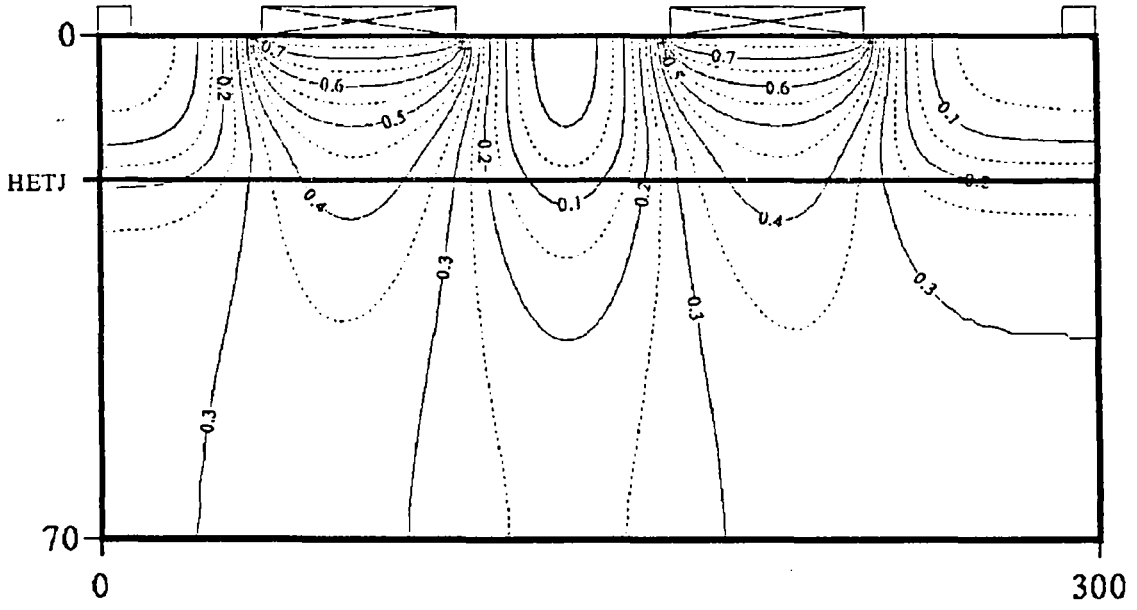


Figure 4.22 (a) steady state contours of potential.

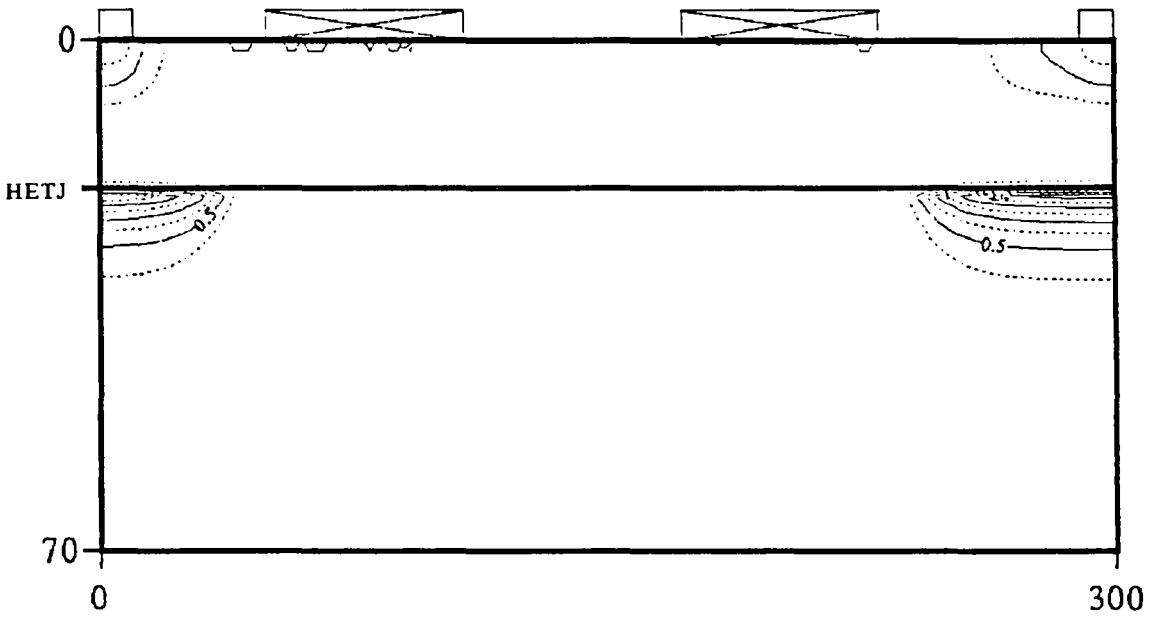


Figure 4.22 (b) steady state contours of charge density.

4.14 Simulation of a deep Ohmic contact :

Previously, we have conventionally considered a shallow Ohmic contact. However, under very low temperature conditions, it appears extremely difficult to contact by this means the 2D-gas. Indeed, the potential barrier at the heterointerface oppose very efficiently to the electron conduction with no means to take advantage of thermal or tunnelling assisted charge transfer.

In reality, the assumption of a shallow Ohmic contact is questionable in the GaAs/AlGaAs technology which makes use of annealing process for alloying the metal-semiconductor interface. Such a thermal treatment induces inter-diffusion of Gold and Gallium. A deep Ohmic contact is thus formed (Figure 4.23). In that particular case, it seems difficult to distinguish between the wide and the narrow band-gap materials. In addition it is unreasonable to assume that the material characteristics will be preserved in the vicinity of contact region with such a thermal treatment. On this basis, it is acceptable to assume that there exists a finite transition layer between the metal characteristics and semiconductor properties.

On the other hand, the way in which the structural properties of this transition layer are varying in the contact growth direction could be simulated by means of any monotonic function. Moreover, we assumed that the thickness of this transition layer is much less than the distance between the active zone and the Ohmic contact itself.

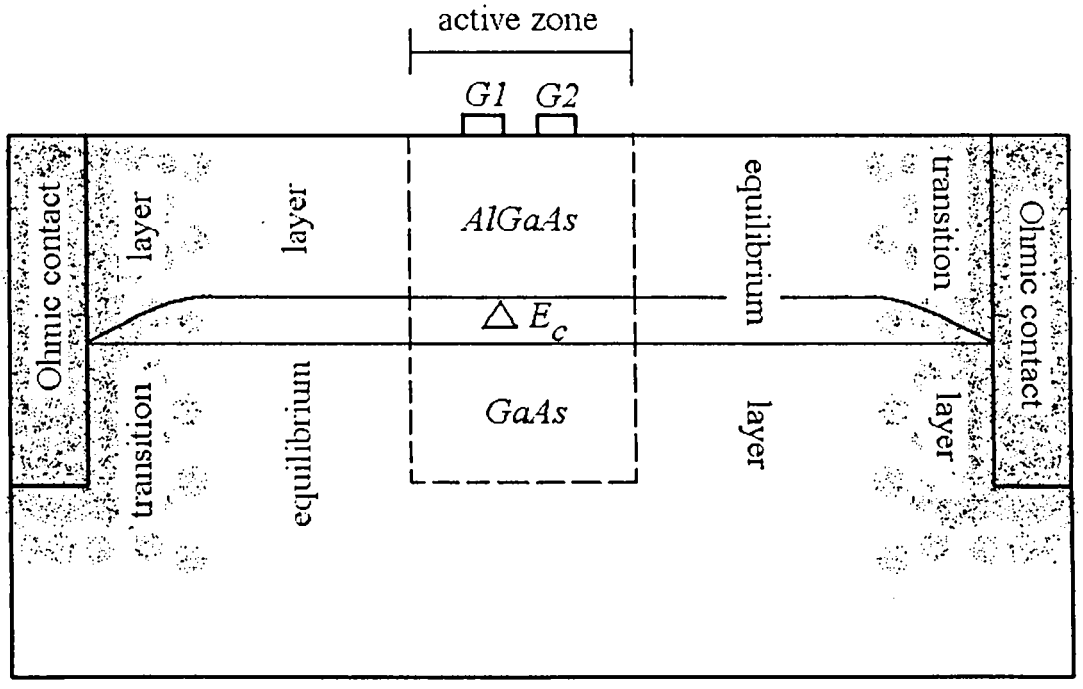


Figure 4.23 schematic diagram illustrating the vertical transition zones associated with the vertical Ohmic contacts.
The dashed lines represents the boundaries of the simulated region.

Let us recall that in the test vehicle considered in the present work, the Ohmic contact are located very far from the active zone. Hence it seems useless to include the whole structure in the simulation code. A first method to reduce the simulated zone is to bring artificially closer the Ohmic contact and the active region. However, it is worth noting that the transition region has to be conserved in order to avoid any artefact between the contact region and the heterojunction. Therefore a second approach is to simply model the boundaries between the active zone and the equilibrium zone as it is illustrated by means of dashed lines in Figure 4.23.

In this situation, we assumed that a vertically constant Fermi-level exists at the vertical boundaries of the active zone rather than a constant equi-potential value. On the other hand the uniform constant charge distribution is replaced by the charge distribution associated with this constant Fermi-level. In the present work this equilibrium distribution was obtained using Thomas-Fermi approximation, but it can be calculated using any other distribution function statistics.

4.15 Notes on the thermal part in the energy equation :

Let us recall the energy equation and rewrite it while separating the thermal and the drift energies as follows,

$$\varepsilon = \varepsilon_{th} + \varepsilon_{drift} \qquad \textit{Eq. 4-87}$$

$$\frac{\partial(\varepsilon_{th}+\varepsilon_d)}{\partial t} + \vec{v}.\nabla_r(\varepsilon_{th}+\varepsilon_d) = q\,\vec{v}.\vec{F} - \frac{1}{n}\nabla_r.(n K_B\theta\,\vec{v}) - \frac{(\varepsilon_{th}+\varepsilon_d - w_o)}{\tau_\varepsilon(\varepsilon)} \qquad \textit{Eq. 4-88}$$

by simple algebra we get,

$$\frac{\partial(\varepsilon_{th}+\varepsilon_d)}{\partial t} + \vec{v}.\nabla_r(\varepsilon_{th}+\varepsilon_d) = q\,\vec{v}.\vec{F}_{net} - K_B\theta\ \nabla_r.(\vec{v}) - \frac{(\varepsilon_{th}+\varepsilon_d - w_o)}{\tau_\varepsilon(\varepsilon)} \qquad \textit{Eq. 4-89}$$

where $\vec{F}_{net}$ is the net deriving force (drift - diffusion) reads,

$$\vec{F}_{net} = \vec{F} - \frac{1}{n}\nabla_r(n K_B\theta) \qquad \textit{Eq. 4-90}$$

Thus Eq. 4-89 can be separated into two equations as follows :

$$\frac{\partial(\varepsilon_{th})}{\partial t} + \vec{v}.\nabla_r(\varepsilon_{th}) = -K_B\theta\,\nabla_r.(\vec{v}) - \frac{(\varepsilon_{th} - w_o)}{\tau_\varepsilon(\varepsilon)} \qquad \textit{Eq. 4-91}$$

$$\frac{\partial(\varepsilon_d)}{\partial t} + \vec{v}.\nabla_r(\varepsilon_d) = q\,\vec{v}.\vec{F}_{net} - \frac{(\varepsilon_d)}{\tau_\varepsilon(\varepsilon)} \qquad \textit{Eq. 4-92}$$

Where now Eq. 4-91 is written only for the thermal energy and Eq. 4-92 for the drift one. This means that the thermal energy directly derived from the Monte-Carlo data is not sufficient to calculate the thermal part of the electron energy. This is because the thermal energy of Mont-Carlo data is solely the part of thermal energy associated with the drift one. On the other hand, the remainder of the thermal energy, resulting from any local decrease or increase in electronic pressure $K_B\theta\,\nabla_r.(\vec{v})$, is not included in Monte-Carlo data. In contrast, one can calculate this thermal energy remainder by

means of Eq. 4-92. Not to mention that the term $K_B\theta\,\nabla_r.(\bar{v})$ with a uniform average velocity in Monte-Carlo simulations vanishes.

This helps us to correctly describe the thermal cooling problem and the time analysis of the simulations, not to mention the effect of the calculated thermal energy (drift associated + pressure associated) on the counterbalance between the drift and diffusion forces. The correction outlined above could be theoretically applied to a strong cooling process between room temperature and liquid Helium temperature. Such a procedure is directly linked to the general issue of feeding [34] the central inter-gate parabolic-like quantum-well

Previously, we have shown that the electron density within this region is practically zero at low temperature with no means to supply this region by means of tunnelling or thermal supplying processes. However, such a situation of complete depletion seems questionable with a strong influence of initial condition and/or the bias history. Now, when we use Thomas-Fermi approximation as the initial condition we obtained the results displayed in Figure 4.22. In the following we will start from this potential distribution which seems more realistic to address the nature of quantum transport.

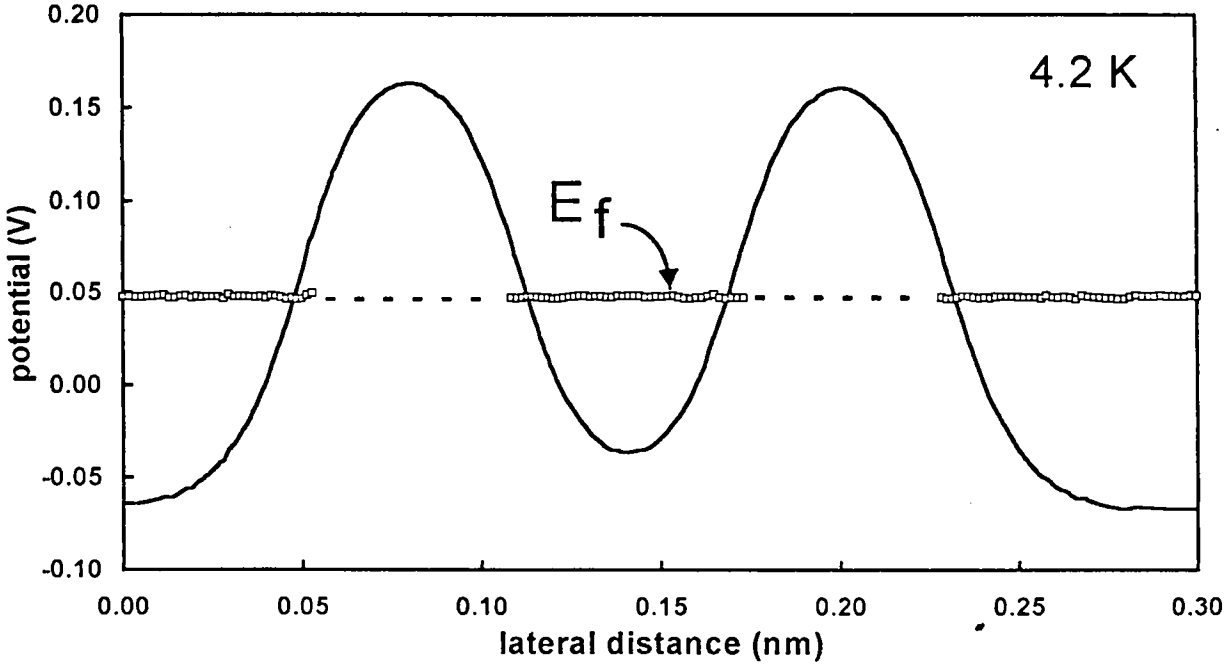


Figure 4.24 lateral channel potential profile along with the Fermi-energy.

As a general rule, we believe that the Thomas-Fermi approximation has to be used as the initial condition in simulation processes specially at low temperatures.

4.16 Leburton's model :

In this section, we will investigate the model proposed by Leburton et al. [35]. In this model the adiabatic approximation is assumed to be valid. Under this condition the authors decoupled the 2D-dependence of Schroedinger equation into vertically and laterally solutions. The vertical one gives the ground energy position in the quasi-triangular quantum well. The lateral one gives the information about the lateral transmission probability. To investigate the validity of such an approximation, we calculated the potential along with the charge concentration firstly under the assumption of full donor ionisation. Figure 4.25 depicts the two dimensional distribution of both the potential and the charge concentration for the device under test considered before. From this figure one can detect that the device of a 50 nm-thick wide-band-gap semiconductor, doped 10^{18} cm^{-3}, is normally-on under the assumption of full ionisation which contradicts the experimental data. Such a disagreement can be overcame either by assuming first a shrinking in the doped epilayer thickness or by revisiting the assumption of full ionisation.

Let us consider now these assumptions respectively with firstly the simulations of a 20 nm-thick wide-band-gap layer. This results in the potential and charge distributions depicted in Figure 4.26.

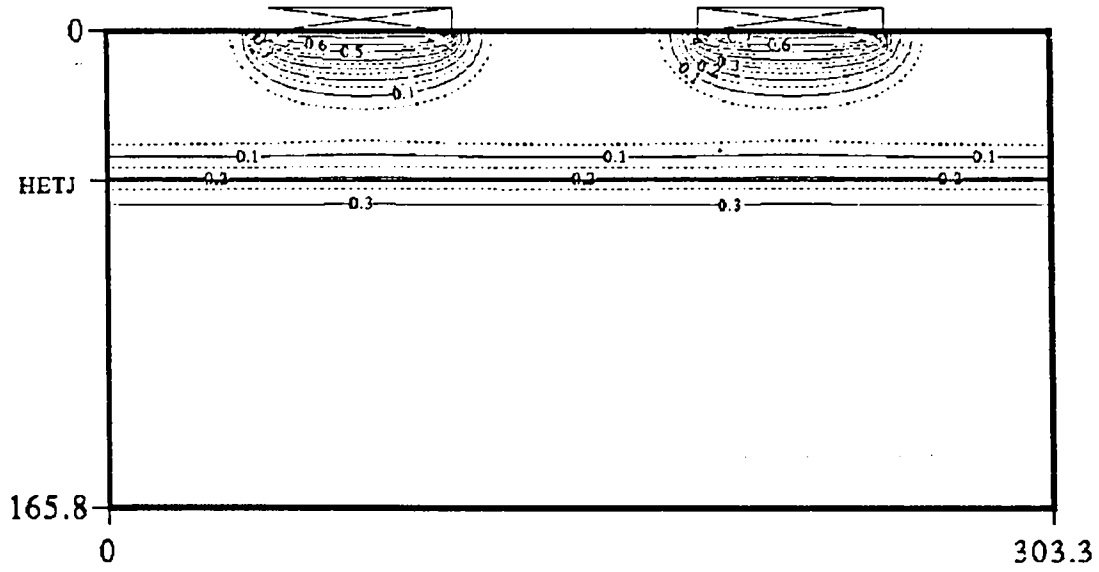


Figure 4.25 (a) potential contours in case of full ionised donors of 10^{18} cm^{-3} in the wide band gap material of about 50 nm-thick.

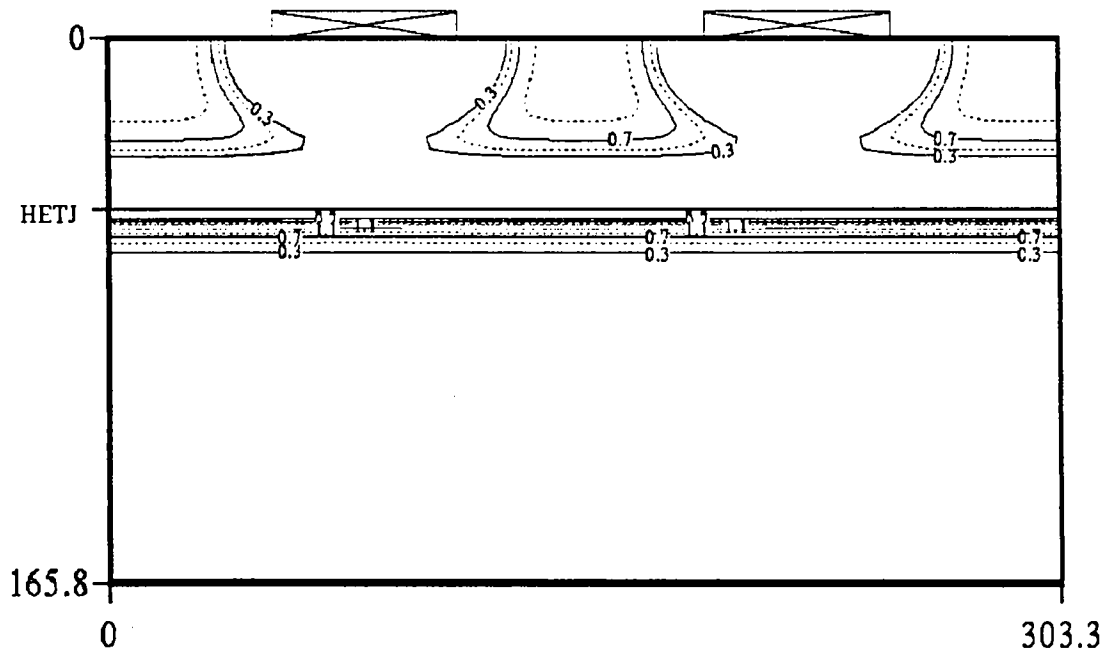


Figure 4.25 (b) charge contours in case of full ionised donors of 10^{18} cm^{-3} in the wide band gap material of about 50 nm-thick.

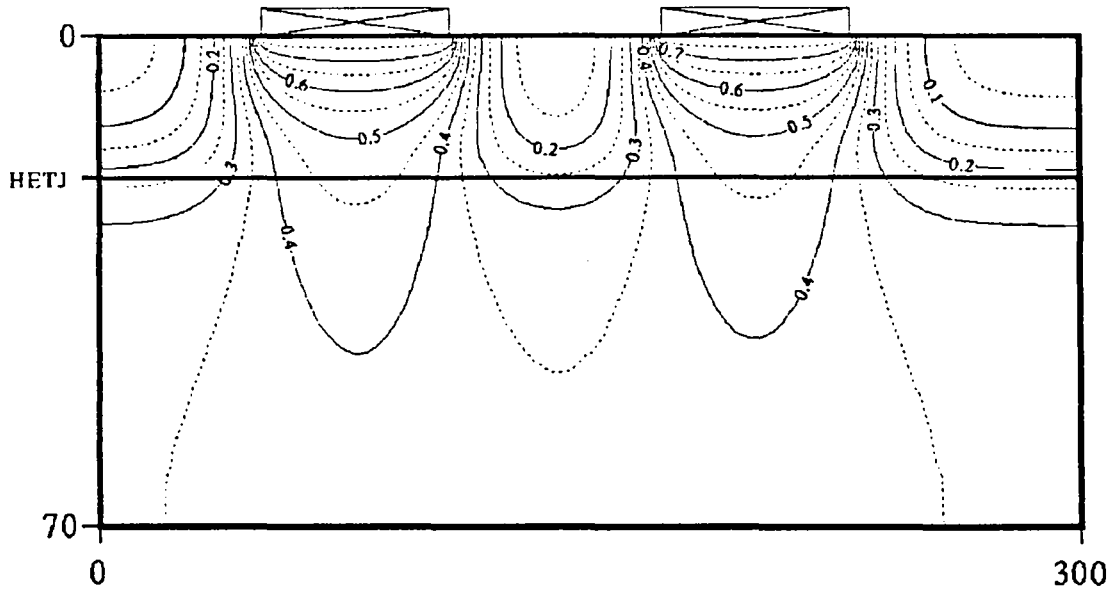


Figure 4.26 (a) potential contours in case of full ionised donors of 10^{18} cm^{-3} in the wide band gap material of about 20 nm thick

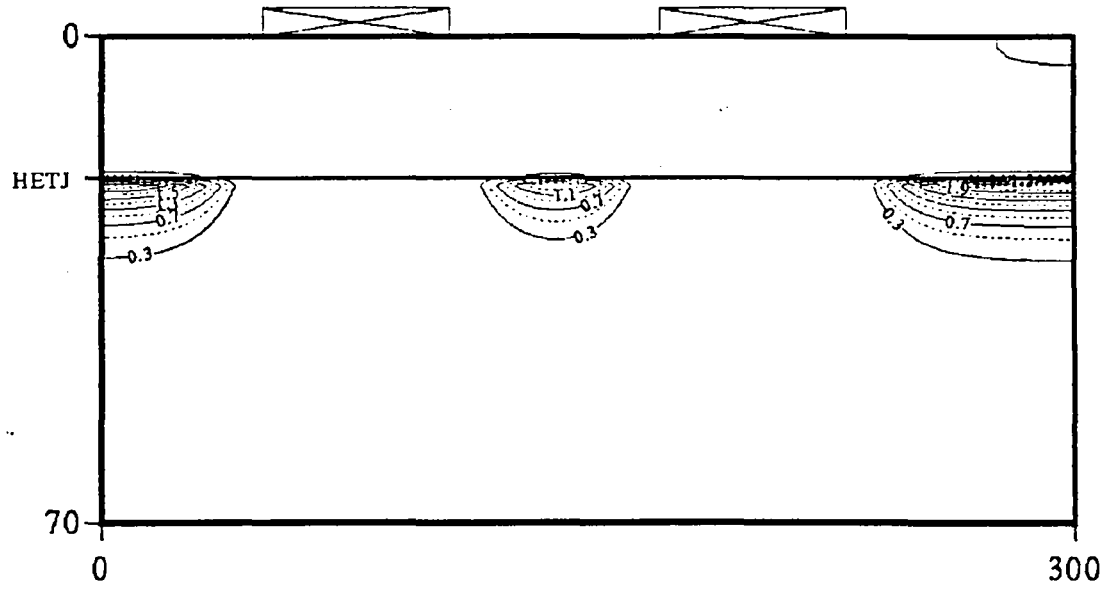


Figure 4.26 (b) charge contours in case of full ionised donors of 10^{18} cm^{-3} in the wide band gap material of about 20 nm thick

In Figure 4.26 one can note the formation of a confined charge in the inter-gate region between completely depleted zones under the two gates. Figure 4.27 illustrates the potential profile in the 2DEG channel just under the heterointerface. In this figure, the extensions of Fermi-level under the gates are plotted in dashed lines. Indeed, in the channel under the gates, the Fermi-level is not determined because the charge concentration in these zones is practically zero. Also it can be seen that the inter-gate zone is isolated from the adjacent zones. Nevertheless, we assume that the Fermi-level is constant everywhere.

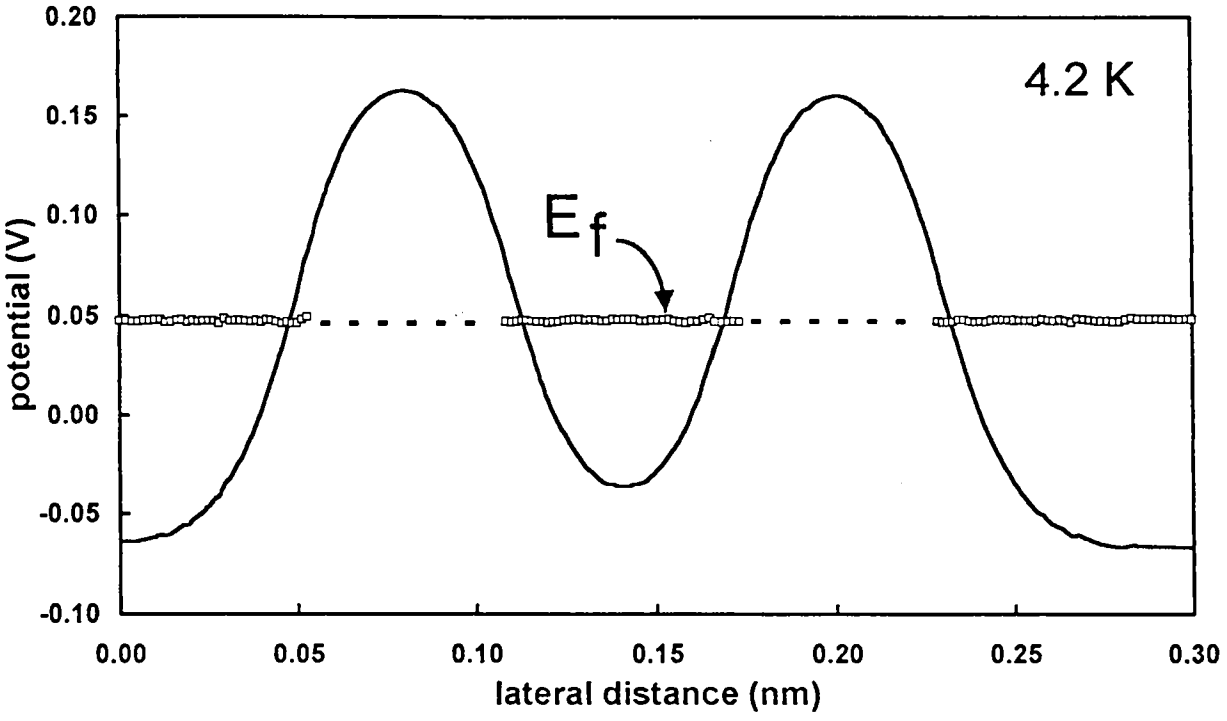


Figure 4.27 lateral channel potential profile detected from the two dimensional distribution of Figure 4.26 (a).

To calculate the ground state of the quasi-triangular quantum well at heterointerface, one can either restrict the calculation to this quantum well itself or take all the vertical diminution into account.

In the former, we assumed a closed system with an infinite barrier height at the heterointerface. In the latter, the simulated domain of the closed system was defined with an infinity barrier height at the lateral free surface. As expected, an enlargement in the simulated domain results in a decrease in the energy position of the ground state. Figure 4.28 illustrates the ground state calculated using the two methods.

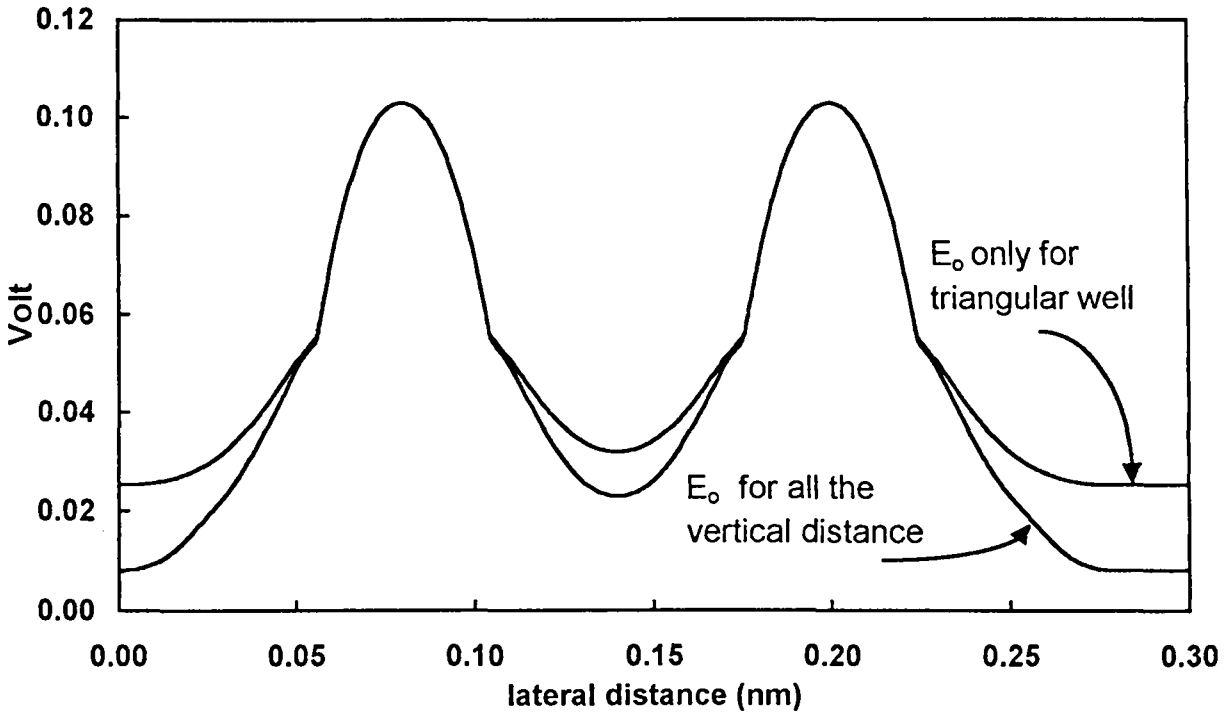


Figure 4.28 ground level profile calculated using different methods.

Also depicted in Figure 4.29(a) are the family of lateral profile of ground state (based on the vertical dimension taking as a whole) for various gate voltages. In this figure (b) the corresponding transmission probabilities are also shown. Also plotted is the Fermi-level (E_f) position in order to investigate the crossover condition between E_f and the resonant peaks.

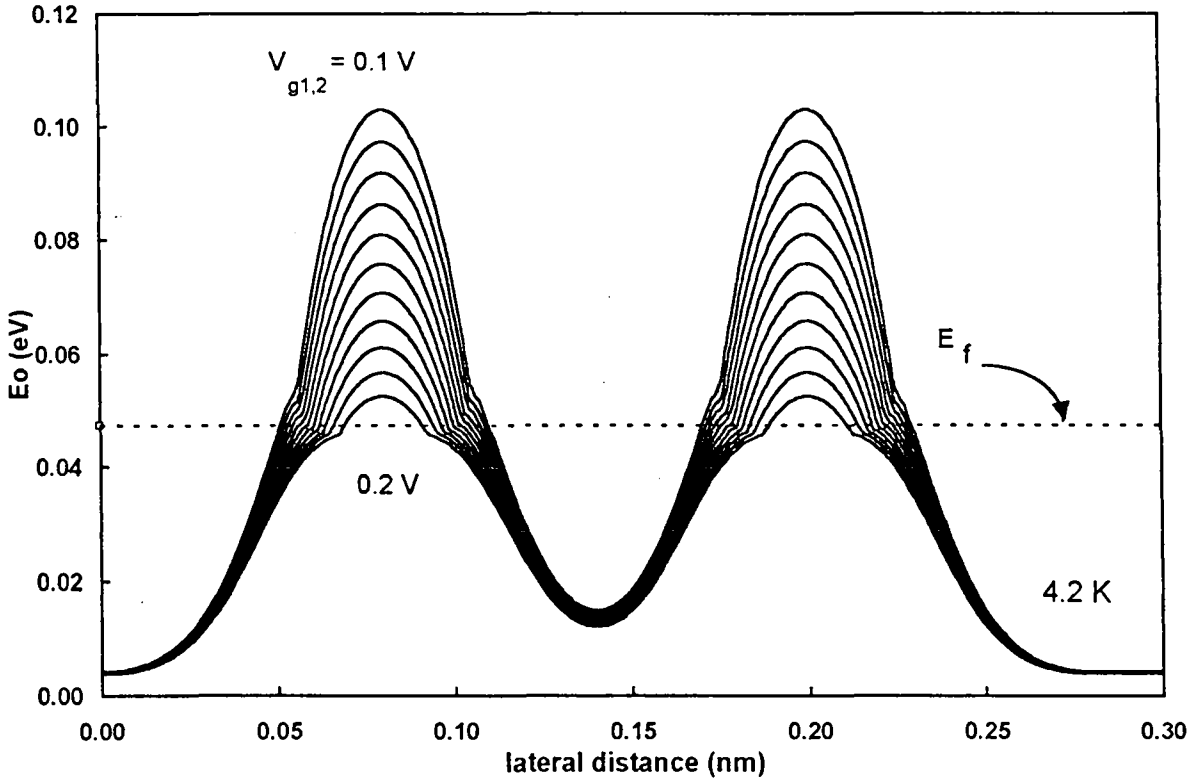


Figure 4.29 (a) lateral profile of heterointerface ground level for various gate voltages.

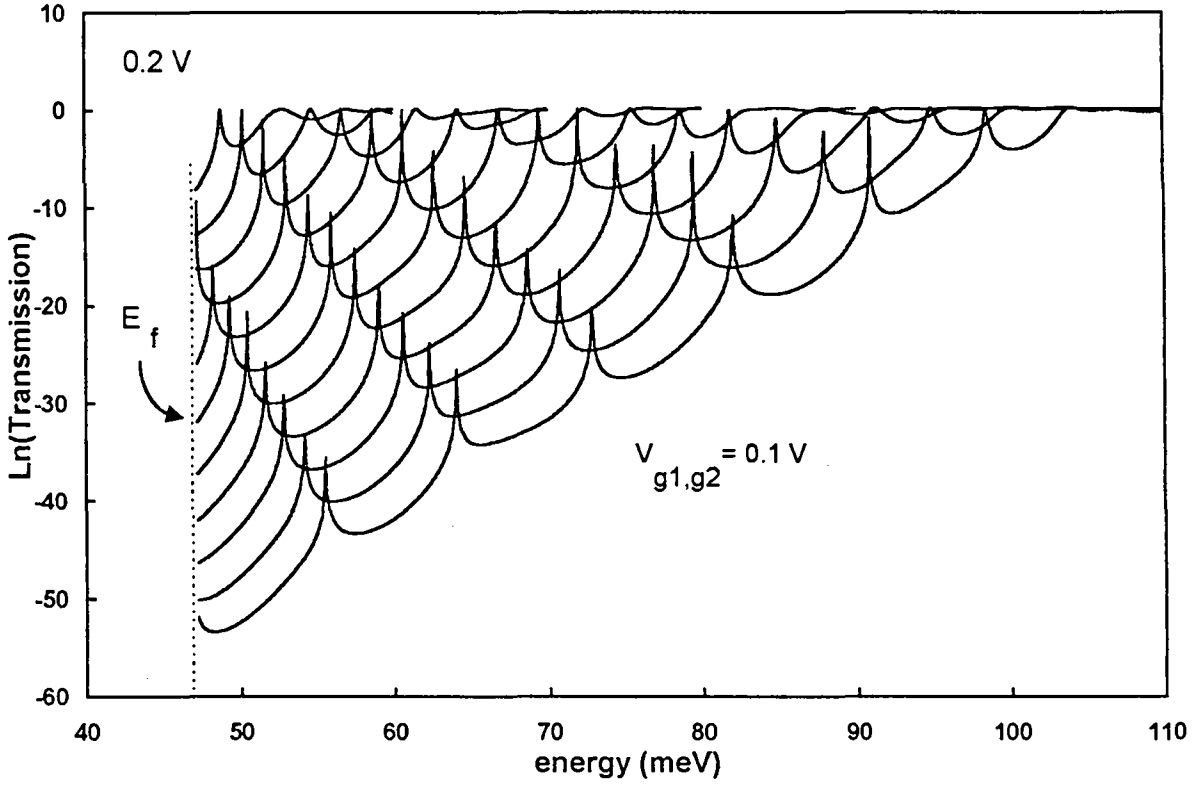


Figure 4.29 (b) transmission probability calculated for the potential profiles of (a).

At $V_{g1} = V_{g2} = 0.1$ Volt, six resonant peaks are clearly apparent in the energy window from E_f ($\approx 50\ meV$) up to the barrier height ($\approx 100\ meV$). When the gate bias is increased to 0.2 Volt only one sharp resonant peak can be detected. The resonant features in the I-V characteristics are the reflect of the crossover between the resonant peaks and E_f. On this basis it can be seen from Figure 4.29(b) that only the nearest quantum level with respect to E_f satisfy this condition. From the current-voltage characteristic point of view this means that only one peak should be measured in contradiction with the experiment. Nevertheless, it seems interesting to quantitatively calculate this resonant tunnelling contribution with the following procedure. As a first remark, it is important to distinguish between the lateral 2D injection described by the lateral 2D supply function (Eq. 4-93 [35]) and the vertical 2D injection discussed in the previous chapters.

$$SF_{lateral} = qL_z v(\varepsilon)\frac{m^*}{\pi\hbar^2}\left[f_{1/2}(\varepsilon - qV_{ds}/2) - f_{1/2}(\varepsilon - qV_{ds}/2)\right] \qquad \textit{Eq. 4-93}$$

where now $v(\varepsilon)$ is the lateral velocity in function of lateral energy ε and all the other variables have their usual meaning. V_{ds} is equal 0.2 mV. The I-V characteristics, we calculated, is plotted in logarithmic scale in Figure 4.30. As expected the tunnelling current peak value is very low in agreement with experiment due to the wide electrostatic tunnelling barriers.

The fact that the resonant states of the transmission probability disappeared before reaching the crossing point with Fermi-level is due to the tapered shape of the electrostatic barriers (see also Figure 4.29). Although the calculated current value is acceptable as compared with experiment in [1] (1nA), the broadening of the calculated current peak is extremely under estimated. In addition, the number of peaks in the device under test was three whereas only one peak was calculated at this stage of work. At last, the peak voltage is also highly over estimated with respect to the experiment.

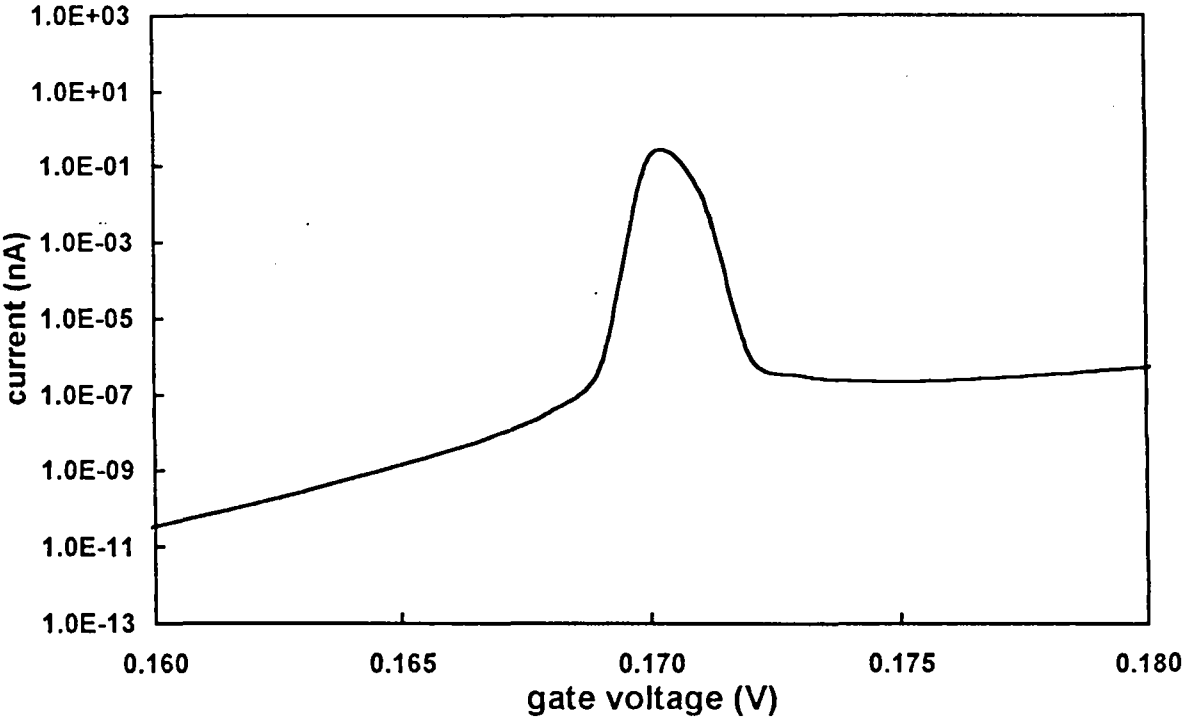


Figure 4.30 I-V characteristics resulting from the integration with respect to energy of the transmission probabilities of Figure 4.29 (b) times the lateral injection supply function.

In an attempt to get a closer agreement with experiment Leburton et al. [35] proposed a purely phenomenological approach, which takes the interface disorders into account. A Fermi-like distribution profile was assumed with an interface disorder potential which reads,

$$qV_{DO}(x,y)=\left[V(x,y)-V_{\min}(y)\right]\times\left[\frac{2}{1+\exp\left(\frac{\alpha_{DO}\left[V_{\max}(y)-V(x,y)\right]}{V_{\max}(y)-V_{\min}(y)}\right)}-1\right] \qquad \textit{Eq. 4-94}$$

Where α_{DO} is a fitting parameter. (in the analysis of [35] α_{DO} is chosen to be 15).

The main goal of introducing of interface disorders was to sharpen the shape of the barrier. In our case, we tried to introduce an equivalent effect by means of different magnitudes namely 10^{18} cm^{-3} and 1.5 10^{18} cm^{-3} respectively.

The main difference between Leburton's correction and our approach is that we introduce a disorder on the charge level whereas Leburton et al. tailored the potential by means of a disorder potential. Nevertheless, it seems that the assumption of non perfect heterointerface in the present status of technology is the earlier idea to be invoked. In the following, however, we shall prove that this has a second order effect.

It is worth noting that the gradual interface-energy discontinuity is formerly assumed in our model. The ionisation probability of these interface donor-states are described by a Fermi-Dirac like distribution described in Eq. 4-95.

$$N^{+}_{ionised} = \frac{N^{+}_{o}}{1 + 2\exp(\varepsilon_{f} - \varepsilon_{ion})} \qquad \textit{Eq. 4-95}$$

Where N^{+}_{o} is the maximum interface disorder states, $N^{+}_{ionised}$ is the effectively ionised interface disorder states and ε_{ion} is the ionisation energy.

The iso-potential patterns are shown in Figure 4.31(a) for $N_o=10^{18}$ cm^{-3} and Figure 4.31(b) for $N_o=1.5 \times 10^{18}$ cm^{-3}.

Comparing the two cases represented in this figure, one can deduce that the effect of interface disorders is to increase the effective gate bias voltage as seen by the 2DEG in the conduction channel. This can be directly pointed out by noticing the 0.4 V voltage contour in the above figure. For seek of clarity, we plotted in Figure 4.32 the potential profile for different interface disorder states concentrations. The general trends of the laterally varying potential profile are similar to that displayed in Figure 4.29(a).

This means that the introduction of various interface charge concentrations has overall the same effect of varying the gate bias voltages. Under the same scope, let us recall that experimentally the authors of reference [1] obtained the same current profile under optical control conditions (see Figure 4.3).

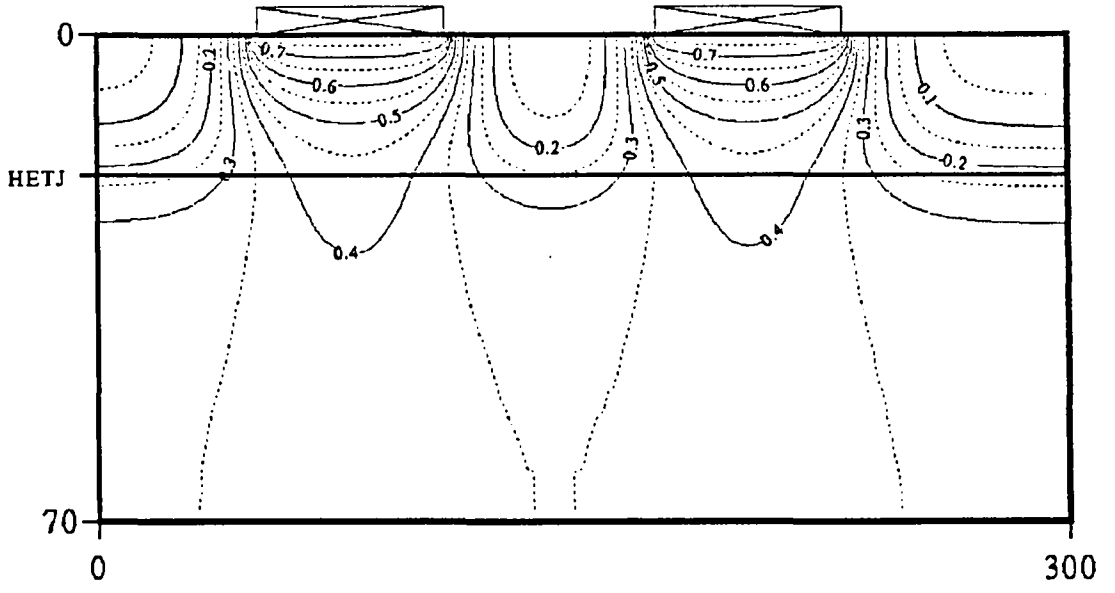


Figure 4.31 (a) potential contour representation when the interface density of disorder states are $N_o=10^{18}\ cm^{-3}$

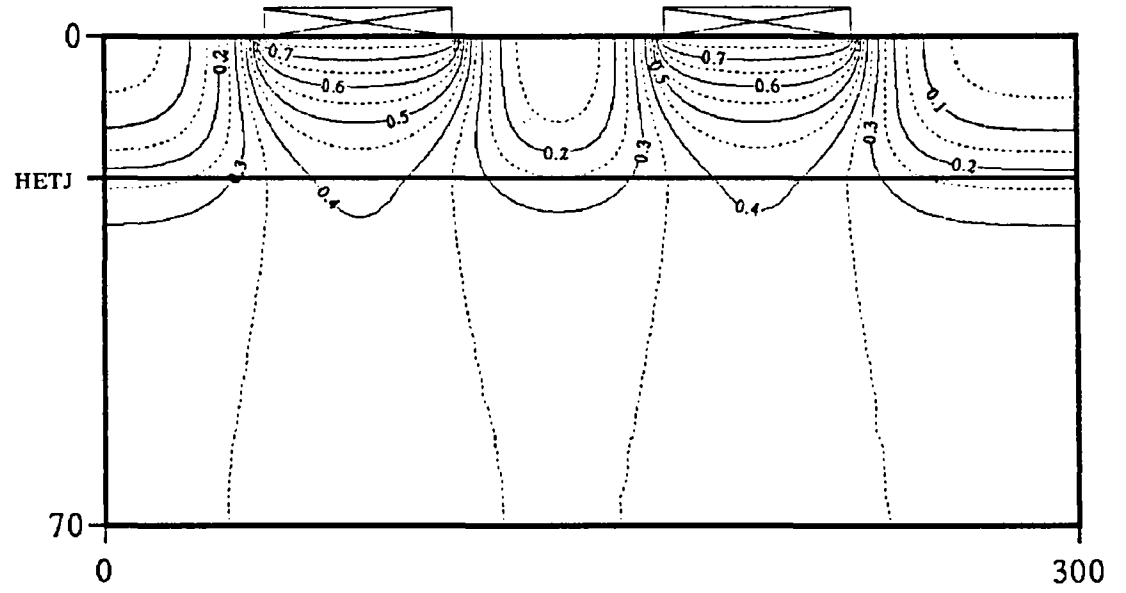


Figure 4.31 (b) potential contour representation when the interface density of disorder states are $N_o=1.5\ 10^{18}\ cm^{-3}$

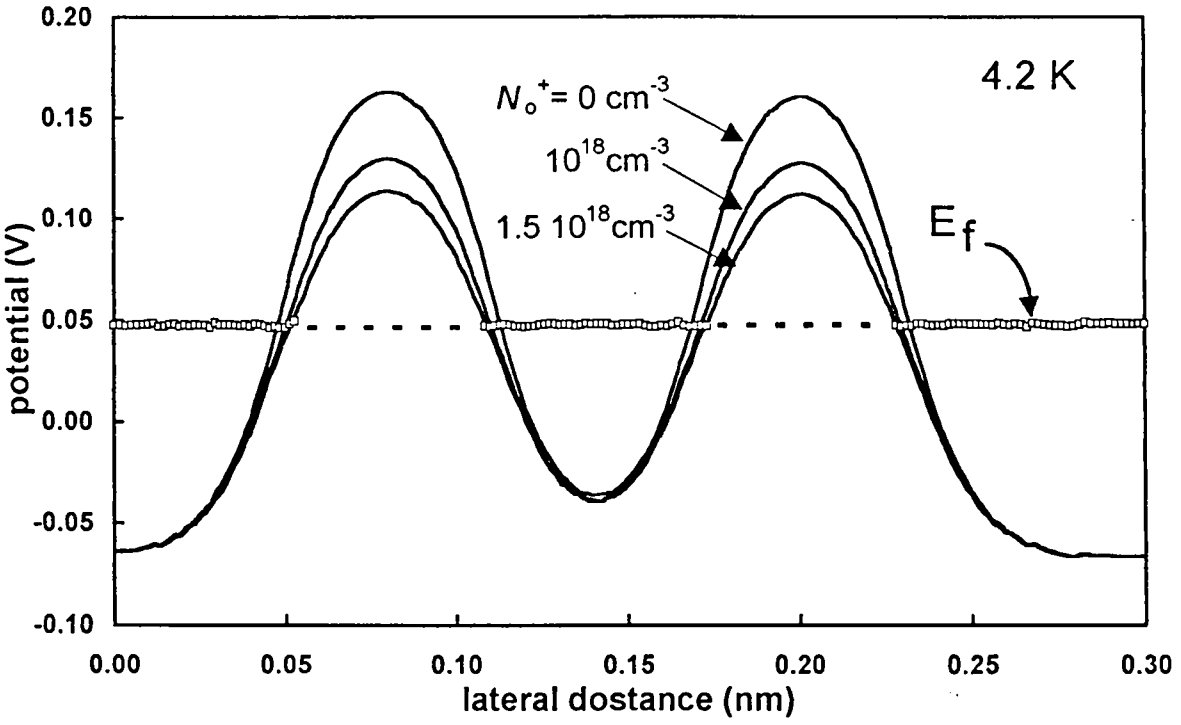


Figure 4.32 lateral potential- profile for different interface disorder conditions.

In conclusion, the impact of introducing additional sheet charge concentration or resulting potential modifies only the barrier height but failed to sharpen the potential profile. Therefore, it seems that such an approach is too phenomenological and it was decided to test other work assumption while conserving the real device topological dimensions.

In contrast to the previous section, we will investigate the effect of including a partial ionisation probability while maintaining the original vertical dimensions of the measured device. At very low temperature (4.2 K) it is expected that the donors will be partially ionised. Also the optical control measurements in [1] support this assumption. The optimisation criterion is that the device should be at the onset of conduction at gate bias of 0.1 V [1]. To met this condition, 13 % of the donors are assumed to be effectively ionised. Besides the inter-gate distance was slightly increased up to 77 nm rather than 60 nm which was the nominal inter-gate distance of the device experimentally characterised. It is believed that this difference in geometry is within the technological tolerance.

The resulting potential profiles, displayed for various bias conditions are depicted in Figure 4.33. Let us recall that previously in Figure 4.29(a) only the electrostatic barrier height is influenced by the gate bias. In contrast, here, not only the barrier height but also the depth of the electrostatic quantum-well is decreased with the gate potential. The consequences are directly detected when we compare the family of transmission plotted in Figure 4.29(b) with Figure 4.33(b). The former demonstrates that only one resonant transmission peak crosses the injection Fermi-level whereas in Figure 4.33(b) three peaks satisfy the cross-over condition.

However, these studies illustrate clearly the trade off between the trends of experimental data versus the bias control and the quantitative agreement notably for the current magnitude, the inherent broadening and the voltage peak positions. In particular, the present numerical experiment, we succeed to represent the voltage positions but loose the feature of good agreement in current peak values. Also, the broadening of the measured current peak is extremely under estimated in both theoretical approaches. Before ending this discussion, it seems interesting to consider further improvements in theory on the basis of what we learned through out this thesis. There exist two important issues to be discussed, notably, the voltage broadening of current peaks and the peak current value as compared with experiment.

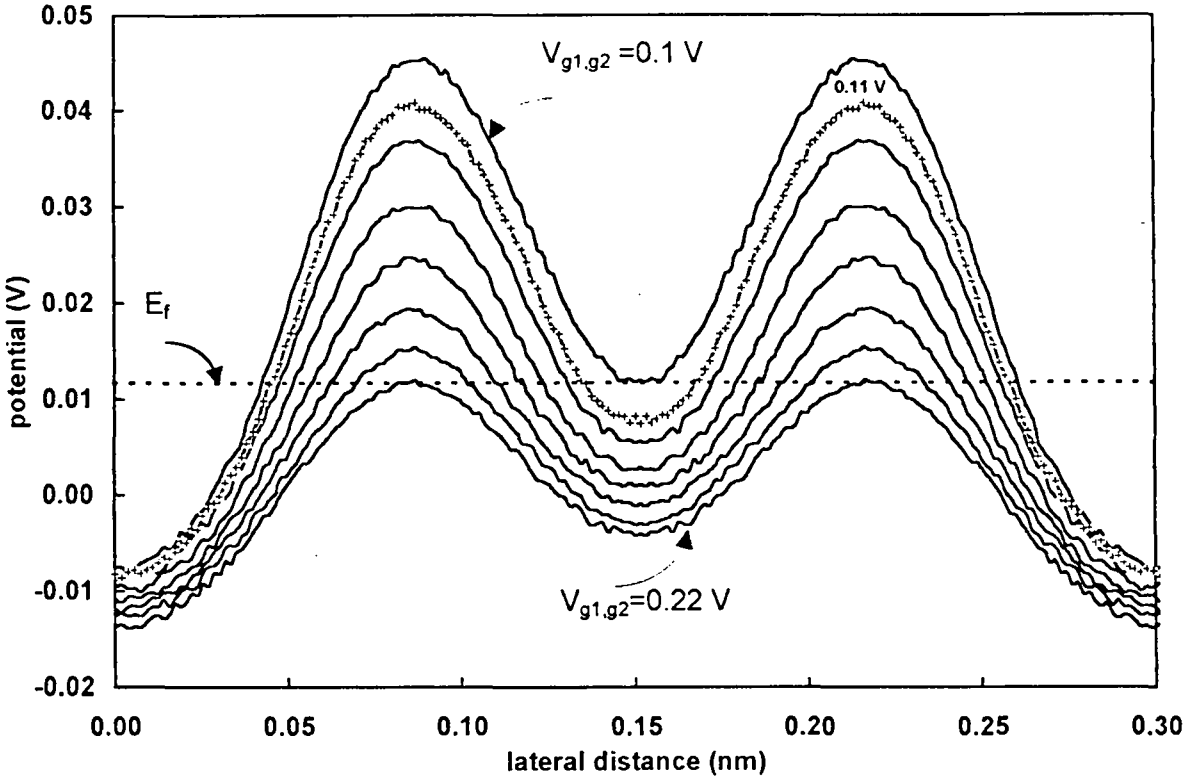


Figure 4.33 lateral profile of channel potential for various bias conditions.

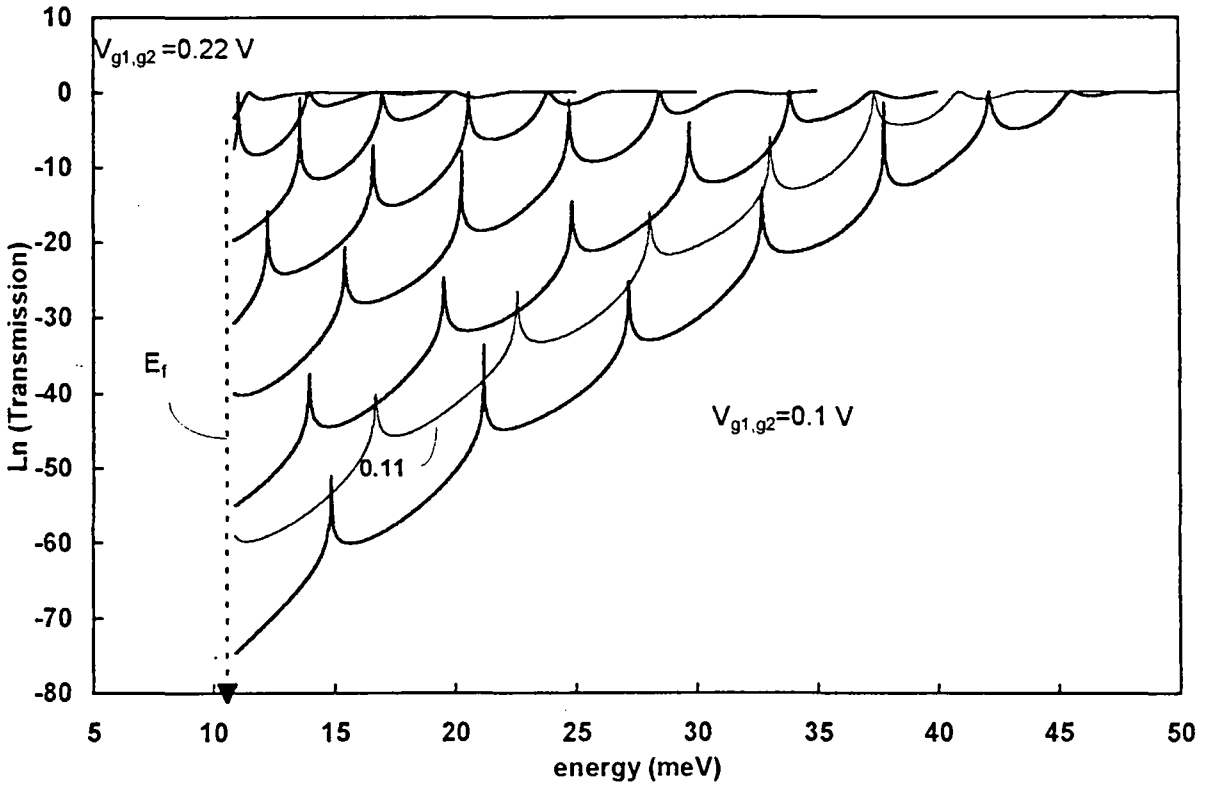


Figure 4.33 (b) transmission probabilities calculated for the potential profiles of (a).

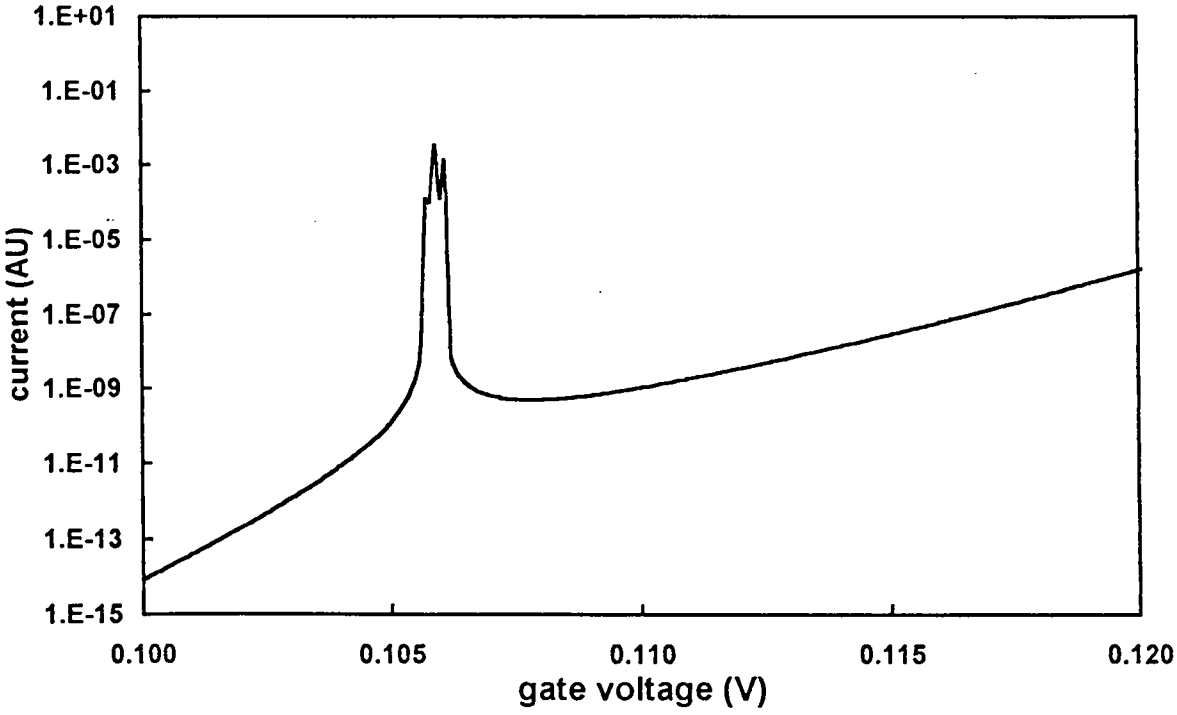


Figure 4.34 I-V characteristics for the first resonant peak.

In our opinion, one of the kernel issues is that we described the resonant tunnelling process in a two-dimensional potential function using simply the one dimensional tunnelling lateral model. This is in strong contradiction with the simple view of current lines normal to the equipotentials which exhibit a pronounced 2D real space character. In the following, we are discussing the expected impact and proposing original tracks for modelling a real two dimensional real space tunnelling process.

4.17 Proposal for 2D tunnelling model :

Let us redraw a schematic diagram dividing the device cross-section into classically allowed propagation regions and classically forbidden barrier zones. This is shown in Figure 4.35, where the hatched areas represent the propagation regions while the clean areas correspond to the depleted ones.

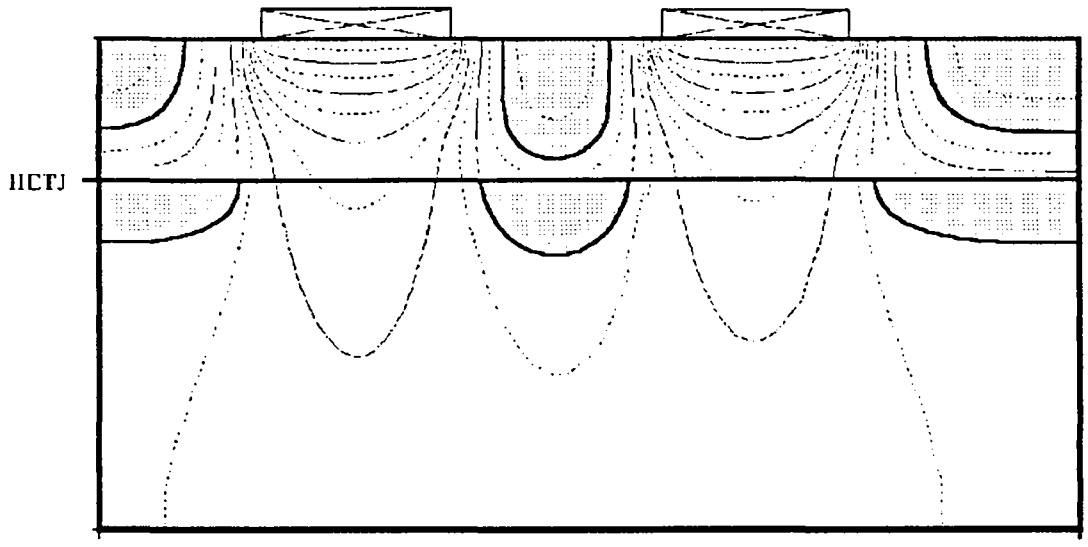


Figure 4.35 schematic diagram illustrating the classically allowed regions (hatched area) and the classically forbidden zones (white area).

Let us now investigate the source injection zone. Indeed, the vertical confinement is much more pronounced as compared with the lateral one. This gives rise to the vertical ballistic velocity to be much higher than the lateral one. This gives rise to a vertical normal-injection analogous to the 2D injection extensively studied in the previous chapters. The result is an appreciable increase in the injection efficiency with respect to the lateral one described above by Eq. 4-93. It remains to discuss the most probable tunnelling path. In fact, intuitively, an electron will chose to cross the 2D real space barrier at tracks where the path integral attenuation is minimised. This is the first guide-line from the quantum transport probability point of view. Indeed, the direct lateral tunnelling path which is used all over the analysis before is not the one which minimises the attenuation experienced by the tunnelling electrons. This in turn, will consequently increase the transmissivity of each electrostatic two-dimensional barrier and hence the overall transmission probability. The electron real space trajectories are schematically illustrated in Figure 4.36 along with the generic electron motion in the classically allowed regions with a very efficient vertical injection.

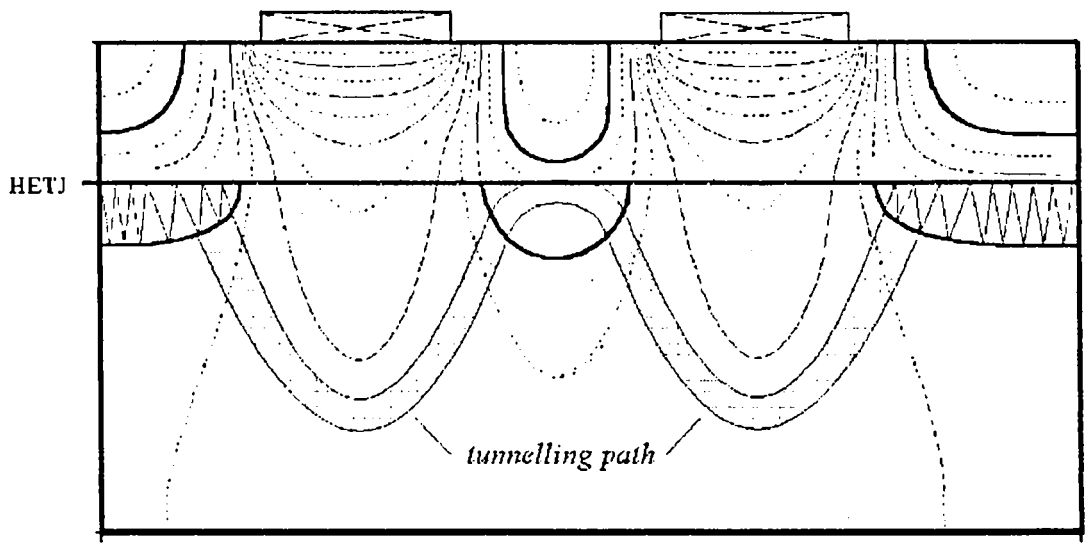


Figure 4.36 schematic diagram illustrating the electron trajectories both in the propagation and the tunnelling regions (hatched aria).

In partial conclusion, it is expected that both the improvement in the quantum transmission probability and in the injection efficiency would permit to have a better agreement in current peak values.

Finally, we are expecting that the global current peak values will approach the experimental data to a very large extent.

With respect to the shape and broadening of the current peaks, let us now discuss schematically the resulting transmission probability. Figure 4.37 is a schematic plot of the expected transmission probabilities. In fact, we have various resonant paths in the inter-gate quantum well which give rise to different degrees of freedom for the corresponding resonant energy. This results in a family of transmission probabilities rather than unambiguously defined single-path transmissions. In other words, the single-path highly peaked single transmission probability should be replaced by a multi-path transmission probability resulting in an envelop transmission function (dashed line in Figure 4.37). The relevant peak current profile is now following the envelop shape of the transmission probability which could give rise to a more realistic peak current profile, in an attempted agreement with experimental data.

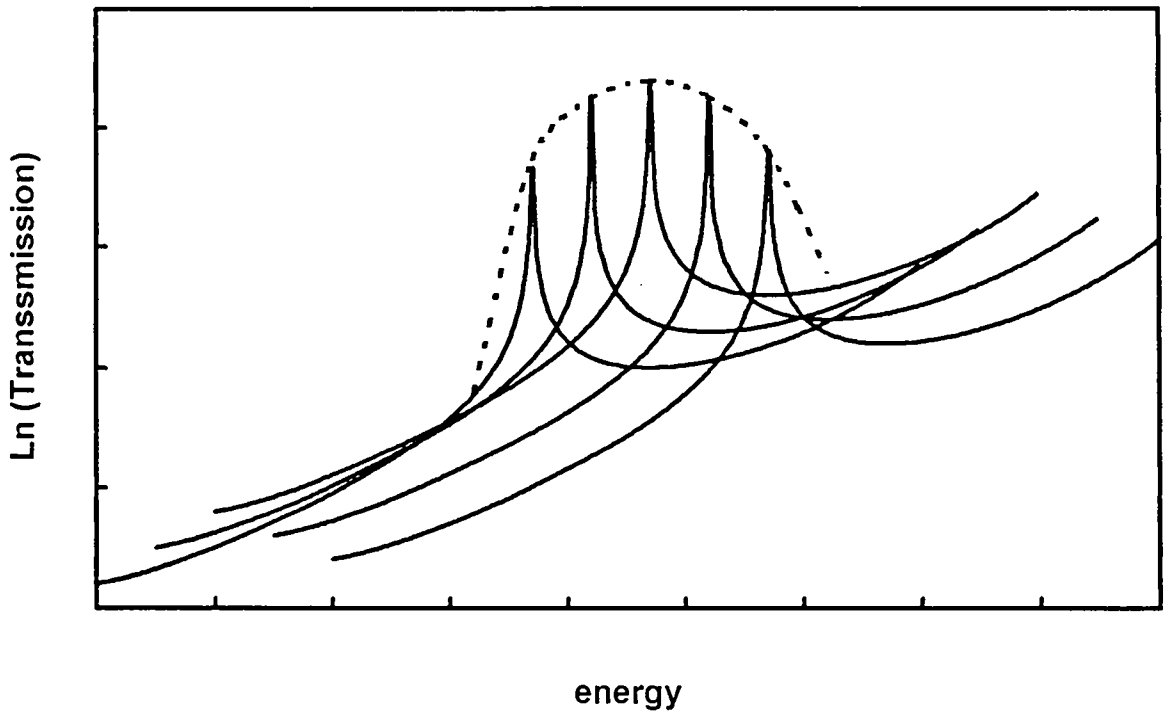


Figure 4.37 schematic diagram illustrating the possible resonate transmission peaks resulting from the flexibility achieved by different paths in the inter-gate region.

4.18 Conclusion :

In this last chapter, the lateral induced 2D real space resonant tunnelling electrostatic induced potential function is investigated. This was done under the scope of hydrodynamic energy model. The latter is subjected into three novel key modifications : firstly, the modelling of a deep Ohmic contact, the second is the additional thermal part of the energy equation which influences the counterbalance between the electric field and diffusion force. The last important modification concerns the introduction of the degeneracy effect into the non-degenerate hydrodynamic energy model. The scaled time domain analysis is investigated and used to alleviate the numerical instability. The latter was provoked by the high mobility value along with the pronounced increase in the electron concentration as calculated under high degeneracy conditions.

The tunnelling process was extensively investigated under the adiabatic assumption made by Leburton and co-workers. Finally we have shown that such an assumption is not able to describe quantitatively all the relevant features of the tunnelling process in the 2D real space potential.

As a last conclusion, we have shown that it is not entirely correct to describe the resonant tunnelling process in a two dimensional potential function using simply the one dimensional lateral tunnelling model. Some guide-lines were discussed for accounting for two dimensional tunnelling mechanisms. In particular a strong unbalance between the lateral and transverse velocities was recognised with the associated benefit of pronounced vertical injection. On the other hand we proposed to consider a multi-path tunnelling mechanisms rather than a well defined resonant tunnelling condition. The direct expected consequences are a broadening of the transmission probability spread now across a transmission window. The net effect should be an enhancement in the current peak value along with a voltage broadening in the voltage dependence current relation in closer agreement with experiment.

4.19 References :

[1] Khalid Ismail, Ph.D. thesis Massachusetts Institute of Technology (MIT), May, 1989.

[2] K. Ismail, D. A. Antoniadis, H. I. Smith, Lateral resonant tunneling in a double-barrier transistor, Appl. Phys. Lett., Vol.55 (6), p.589, 1989.

[3] S. Y. Chou, D. R. Allee, R. F. W. Pease, J. S. Harris, Observation of electron resonant tunneling in a lateral dual gate resonant tunneling field-effect transistor, Appl. Phys. Lett., Vol.55 (2), p.176, 1989.

[4] A. C. Seabaugh, J. N. Randall, Y. C. Kao, J. H. Luscombe, A. M. Bouchard, $In_{0.52}Al_{0.48}As/In_{0.53}Ga_{0.47}As$ lateral resonant tunneling transistor, Electronics Letters, Vol.27 (20), 1991.

[5] K. Ismail, W. Chu, D. A. Antoniadis, H. I. Smith, Negative transconductance and negative differential resistance in a grid-gate modulation doped field effect transistor, Appl. Phys. Lett., Vol.54 (5), p.460, 1989.

[6] F. Stern and S. Das Sarma, Electron energy levels in GaAs-GaAlAs heterojunctions, Phys. Rev. B. Vol.30, p. 840, 1984.

[7] T. J. Drummond, H. Morkoc, K. Lee and M. Shur, Model for modulation-doped field effect transistors, IEEE Elctron Device Letters, Vol. EDL-3, p.338, 1982.

[8] R.Dingle, H. L. Stromer, A. C. Gossard, and W. Wiegman, Appl. Phys. Lett. 33, 665 ,1978.

[9] W. Schokley, a unipolar field effect transistor, Proc. IRE, 40, p 1365, 1952.

[10] H. Statz, H. A. Haus and R. A. Pucel, noise characteristics of gallium arsenide field effect transistor, IEEE Trans. Electron Devices, Vol. ED21, No.9, p.549, September 1974.

[11] A. Greben and S. Ghandi, general theory for pinched operation of the junction-gate FET, Solid State Electronics, Vol. 12, p. 573, December 1969.

[12] M. Shur, analytical model of GaAs MESFET's, IEEE Trans. Electron Device, Vol. ED25, No.6, p.612, June 1978.

[13] D. Delagebeaudeuf and N. T. Linh, Metal-(n)GaAlAs-GaAs two-dimensional electron gas FET, IEEE, Trans. Elctron Devices, Vol, Ed-29, p. 955, 1982.

[14] K. Lee, M. S. Shur, T. J. Drummond and H. Morkoc, electron density of the two-dimensional electron gas in modulation doped layers, J. Appl. Phys., Vol. 54, p. 2093, 1983

[15] M. Reiser, two-dimentional numerical FET model for DC AC and large-signal analysis, IEEE Trans. Electron Device, Vol. ED20, no.1, p.35, 1973.

[16] K. Yamaguchi et al., two-dimensional numerical analysis of stability criteria of GaAs FET's, IEEE Trans. Electron Devices, Vol. ED23, no.12, p. 1283, 1976.

[17] J. Barnes, R. J. Lomax et al., finite element simulation of GaAs MESFET's with lateral doping profiles and submicro gates, IEEE Trans. Elctron Devices, Vol. ED23, no.9 p.1042, 1976.

[18] Kjell Blotekjaer, transport equations for electrons in two valley semiconductors, IEEE, Trans. Electron Devices, Vol.ED17, no.1, p.38, 1970.

[19] A. Cappy, sur un nouveau modèle de transistor a effet de champ a grille sub-micronique, Thèse de Troisième Cycle, Université des Science et Techniques de Lille, France, 1981.

[20] R. Cook and J. Fery, two-dimensional numerical simulation of energy transport effects in si and GaAs MESFET's, IEEE Trans. Electron Devices, Vol.ED29, no.6, p.970, 1982.

[21] M. Ibrahim, two-dimensional simulation of microwave GaAs submicronic-gate FET, M.Sc. Thsis, Faculty of Engineering, Cairo University, 1983.

[22] S. El-Azhary, effect of substrate and gate length on the GaAs MESFET performance, M.Sc. Thesis, Faculty of Engineering, Cairo University, 1983.

[23] K. Ismail, design and simulation of structurally optimised GaAs and GaInAs FET's, M.Sc. Thesis, Faculty of Engineering, Cairo University, 1985.

[24] T. Shawki, performance analysis of $Al_{0.3}Ga_{0.7}As$/GaAs MODFET transistors, M.Sc. Thesis, Faculty of Engineering, Cairo University, 1987.

[25] T. Shawki, conception d'un modèle hydrodynamique bidimensionel de transistor a effet de champ a heterostrcuture : application a l'analyse physique et a l'optimisation des composants submicronique, Thèse de docteur de l'Université, l'Université des Science et Techniques de Lille, 1990

[26] T.Shawki, G. Salmer and Osman L. El-Sayed, MODFET 2-D hydrodynamic energy modelling : optimisation of subquarter-micron-gate structures, IEEE Trans. Electron Devices. Vol.37, p.21, 1990.

[27] Sherif, G. Salmer and O. L. El Sayed, in Simulation of Semiconductor Devices and Process, vol. 5. Springer, Berlin, 1993

[28] A. R. Mikhail, performance analysis of dual-gate MODFET's, M.Sc. Thesis, Faculty of Engineering, Cairo University, Mars, 1993

[29] K. Sherif, A. Refky Mikhail, T. Shawki, O. El-Sayed and G. Salmer, two-dimensional hydrodynamic simulation of sub-micrometer dual-gate MODFET's, Solid State Electronics, Vol.38, no.4, p.917-929, 1995.

[30] R. S. Varga, Matrix iterative analysis, Englewood cliffs, N. J. : Printice Hall, 1962.

[31] Jing-Rong Zhou and David K. Ferry, Modelling of quantum effects in ultrasmall HEMT devices, IEEE Trans. Electron Devices Vol.40, No.2, 1993.

[32] E. M. Azoff, Generalised energy-momentum conservation equations in the relaxation time approximation, Solid State Electronics Vol.30, No.9, p.913-917,1987.

[33] D. L. Rode, Semiconductor and semimetals, Vol.10

[34] Dejan Jovanovic and Jean-Pier Leburton, self-consistent analysis of single-electron charging effects in quantum-dot nanostructures, Phys. Rev. B. vol.49, 1994.

[35] Jeffry M. Bigelow and J.-P. Leburton, Self-consistent simulation of quantum transport in dual-gate field-effect transistors, J. Appl. Ohys. 76 (5), 1994.

4.20 Table of contents :

CONCLUSION

General conclusion :

The common denominator of the studies reported in this work is the resonant tunnelling effect. This effect can be observed under the condition of charge transfer by tunnelling and localisation by means of crystalline and/or electrostatic potential barriers. With respect to the general context out lined in chapter one, two kinds of structures in two terminal and three terminal configuration respectively have been considered. To summarise the main conclusion we have drown through out this thesis, let us recall that it was found that the injection process plays an unexpected role in the negative differential conductance voltage dependence. The impact of such a finding is beyond the sole framework of double barrier heterostructure and was generalised to the open quantum systems. On the other hand, the scattering assisted effect was systematically integrated in order to successfully describe the current contrast. At last, from the transport properties point of view, we succeeded to combine a classical description of transport properties based on a hydrodynamic approach with a quantum resonant tunnelling transport properties based on what we have learned about resonant tunnelling through out this thesis.

In chapter two, we have addressed the general problem of charge transfer between two systems of different dimensionality. This was applied to the tunnelling transition between the accumulation layer and the collector region through the resonant path afforded by a Double Barrier Heterostructure (DBH). We found by this means novel effects, in this work referred to as signature effect, resulting from the coupling between the quantum well and the injection zone. In short, beyond the bias point for anti-crossing of quantum states attached to the accumulation and quantum well regions respectively the tail of the wave function is sufficiently high to induce a finite density of states. To our knowledge this is the first time that a such result is mentioned with a dramatic consequence on the conduction mechanisms under out-of-resonance conditions.

On the basis of local density of states calculations, the current-voltage characteristics have been calculated in details under various assumptions and shows a very broad voltage range for NDC effect with a plateau-like shape in a very good agreement with experimental data. It is believed that with the recent advances in epitaxial growth allowing in fact to grow extremely thin and highly doped epi-layers, such a general formalism could find a very general application. The recent characterisation in our group of high current derivability DBH with high doped cladding layer shows I-V relationship with the same shoulder type behaviour predicted in this work for the first time to our knowledge.

In chapter three, we have studied the scattering assisted tunnelling. The effect of scattering is included not only in the transmission probability but also into the injection conditions through a novel probabilistic model. In connection to the new formalism described in details in chapter two, we have succeeded to define four probability rates associated with each possible scattering assisted supply condition. The transmission probability is self-consistently solved along with the equation defining the finite life time in the well. In addition, the transmission probability is calculated using a new method we called AEES based on Airy function solutions.

The above model, is systematically applied to simulate the resonant tunnelling diode fabricated in our group of research and which was characterised in the present work as a function of temperature. The comparison between the simulated results and the measured data are surprisingly in a very good agreement. This good agreement is also shown to be sustained for a broad range of operating temperature conditions. Using the above novel model, we have described successfully the main anomalies in the I-V characteristics notably, the onset-bump, the main resonant current and the post-resonant phonon replica.

In chapter four, the 2D electrostatic induced potential function is investigated. This was done under the scope of hydrodynamic energy model. Concerning this specific feature three novel modifications were proposed (i) the modelling of a deep Ohmic contact, (ii) introduction of the additional thermal part of the energy equation which influences the counterbalance between the electric field and diffusion force, (iii) the last important modification concerns the introduction of the degeneracy effect. In order to alleviate the numerical instability the scaled time domain analysis is investigated and convergence was satisfactory reached.

Starting from the 2D potential pattern, the tunnelling process was extensively investigated under the assumption of separation of variables. Such a model is able to describe the main features of the measured current-voltage characteristics notably the three anomalies due to resonant tunnelling through the camel-shape electrostatic potential. It is believed that such a basic study of quantum effect will find application in context of nano-devices. To mention a few, one can expect that at ultra short gate length the conventional MOSFET will face the problem of a quantum behaviour superimposed or dominant with respect to the conventional field effect control scheme. Based on tunnelling and temporary localisation effect, we thought that the devices designed for single electron electronics could also be addressed by means of the numerical code carried out in the present work.

As a last conclusion, we have shown that it is not acceptable to describe the resonant tunnelling process in a two dimensional potential function using simply the one dimensional lateral tunnelling model. A novel two dimensional tunnelling model depends on the vertical 2D injection regime is proposed. It is also shown that this regime is analogues to that discussed in both chapters one and two for vertical heterostructural devices. It is expected that by using the new 2D physical approach we can get a better agreement with the experimental data not only with respect to the current peak values but also as regards the voltage broadening of the resonant current peaks.